INTRODUCTION TO ENVIRONMENTAL FORENSICS

SECOND EDITION



EDITED BY

BRIAN L. MURPHY
ROBERT D. MORRISON

INTRODUCTION TO ENVIRONMENTAL FORENSICS

SECOND EDITION

Visit the companion website for
Introduction to Environmental Forensics:
http://books.elsevier.com/companions/9780123695222

The website for *Introduction to Environmental Forensics, Second Edition* contains dozens of color images from the book. These images are meant to enhance the student's understanding and comprehension of the material by illustrating key concepts covered in the text.

INTRODUCTION TO ENVIRONMENTAL FORENSICS
SECOND EDITION

Edited by

Brian L. Murphy and Robert D. Morrison



AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier



Acquisitions Editor: Jennifer Soucy
Signing Editor: Pamela Chester
Assistant Editor: Kelly Weaver
Project Manager: Christie Jozwiak
Marketing Manager: Chris Nolin
Cover Designer: Gene Harris

Elsevier Academic Press
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper.

Copyright © 2007, Elsevier Inc. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone: (+44) 1865 843830, fax: (+44) 1865 853333, E-mail: permissions@elsevier.co.uk. You may also complete your request on-line via the Elsevier homepage (http://elsevier.com), by selecting "Customer Support" and then "Obtaining Permissions."

Library of Congress Cataloging-in-Publication Data
APPLICATION SUBMITTED

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 13: 978-0-12-369522-2
ISBN 10: 0-12-369522-8

For all information on all Elsevier Academic Press publications
visit our Web site at www.books.elsevier.com
Printed in China
07 08 09 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

TABLE OF CONTENTS

INTRODUCTION TO THE SECOND EDITION vii
CONTRIBUTORS xi

CHAPTER 1 APPLICATIONS OF ENVIRONMENTAL FORENSICS 1
Brian L. Murphy

2 SITE HISTORY: THE FIRST TOOL OF THE ENVIRONMENTAL FORENSICS TEAM 23
Shelley Bookspan, A.J. Gravel, and Julie Corley

3 PHOTOGRAMMETRY, PHOTOINTERPRETATION, AND DIGITAL IMAGING AND MAPPING IN ENVIRONMENTAL FORENSICS 49
James I. Ebert

4 THE MEASUREMENT PROCESS 83
Dallas Wait and Charles Ramsey

5 STATISTICAL METHODS 129
Thomas D. Gauthier and Mark Hawley

6 STATISTICAL TOOLS FOR RATIO DATA 185
Michael E. Ginevan

7 PRINCIPAL COMPONENTS ANALYSIS AND RECEPTOR MODELS IN ENVIRONMENTAL FORENSICS 207
Glenn W. Johnson, Robert Ehrlich, William Full, and Scott Ramos

8 RECEPTOR MODELS FOR SOURCE APPORTIONMENT OF SUSPENDED PARTICLES 273
John G. Watson and Judith C. Chow

9 CHEMICAL FINGERPRINTING METHODS 311
Gregory S. Douglas, Stephen D. Emsbo-Mattingly, Scott A. Stout, Allen D. Uhler, and Kevin J. McCarthy

10 APPLICATION OF STABLE ISOTOPES AND RADIOISOTOPES IN ENVIRONMENTAL FORENSICS 455
R. Paul Philp and Emilie Jardé

11 FORENSIC APPLICATIONS OF CONTAMINANT TRANSPORT MODELS IN THE SUBSURFACE 513
Ashok Katyal and Robert D. Morrison

12 FORENSIC AIR DISPERSION MODELING AND ANALYSIS 577
Bruce A. Egan and Brian L. Murphy

13 ENVIRONMENTAL FORENSIC MICROSCOPY 611
James R. Millette and Richard S. Brown

14 APPLICATIONS OF LASER ABLATION INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY (LA-ICP-MS) IN ENVIRONMENTAL FORENSIC STUDIES 637
A. Mohamad Ghazi

15 EMERGING FORENSIC TECHNIQUES 671
Jean-Christophe Balouet, Gil Oudijk, Ioana Petrisor, and Robert D. Morrison

INDEX 733

INTRODUCTION TO THE SECOND EDITION

Since the publication of the first edition of *Introduction to Environmental Forensics* in 2002, the science of environmental forensics has matured appreciably. In 2005, *Environmental Forensics: A Contaminant Specific Guide* was published in the forensic series by Elsevier with the intent of providing the user with a means to access the forensic methodologies on a contaminant-specific basis. In contrast, this edition of *Introduction to Environmental Forensics* is designed to provide the reader with a methodological organization of the forensic tools available and as a complementary reference to the *Contaminant Specific Guide*. Additional forensic methods in this second edition include chapters on laser ablation inductively coupled mass spectrometry (LA-ICPMS), manual- and computer-controlled scanning electron microscope (SEM) techniques and x-ray diffraction, pattern recognition methodologies, expanded chapters on sampling techniques and statistical methods, and a presentation of several emerging forensic techniques. In this edition and subsequent editions of *Introduction to Environmental Forensics*, methods of general applicability will be emphasized and the *Contaminant Specific Guide* will provide forensic approaches for specific contaminants. We would be grateful to readers for suggestions for improvement.

"Forensic" is related to "forum" and refers to any public discussion or debate. In the United States "forensic" most often refers to courtroom or litigation proceedings. However, environmental forensics may also provide the fact basis for mediated or negotiated transactions or for any public inquiry related to environmental matters. Questions that environmental forensics seeks to answer are:

- Who caused the contamination?
- When did the contamination occur?
- How did the contamination occur? (For example, was it an accidental spill or a series of routine operating releases?)
- How extensive is the contamination?
- Are the test results valid? Is there evidence of fraud?
- What levels of contamination have people been exposed to?
- Can environmental forensics assist in allocating remediation costs?

The contexts of environmental forensic investigations include liability allocation at hazardous waste sites where multiple parties are involved, site assessments for property transfers, insurance litigation, toxic torts, and cost allocation among multiple parties found liable for releasing contaminants into the environment.

Environmental forensic investigations frequently deal with the historical release of contaminants. Generally there are two sources of information in conducting an investigation, namely:

- The documentary record, including statements by witnesses or other knowledgeable individuals, aerial photographs, insurance maps, and electronic information copied from computer hard drives, and
- Measurement or sampling data.

Once the historical information has been acquired and evaluated, one can then identify which forensic technique is most suited for answering the forensic question(s) of concern and how to appropriately use the selected methodology. For example, chemical or isotope concentration data can be used in different ways to answer forensic questions, including

- Tracer techniques based on the presence or absence of a particular chemical.
- Ratio techniques where the relative amounts of two or more chemicals are compared.
- Trend techniques where the spatial or temporal variation of a concentration, or a ratio, is of interest.
- Quantity techniques that depend on the integrated concentration over space or time, i.e., the mass of a chemical, to provide forensic information.

A forensic investigation may involve multiple forensic techniques and applications that are evaluated to answer the forensic question of interest. For example, identifying the source of and age dating a hydrocarbon spill may be of interest. The presence or absence of lead, methyl tertiary butyl ether (MTBE), or other additives, for example, may provide this crucial information that provides insight regarding the source or age of the release. The ratio of different hydrocarbon componants or octane readings may provide the basis to distinguish different fuels or brands. The spatial variation of a contaminant plume or its growth over time may assist in both source identification and age dating. Finally, the total mass or volume of petroleum hydrocarbons in the environment may be compared to inventory or leak detection records for source identification.

The most successful forensic investigations rely on the approach of selecting the most applicable techniques from numerous methodologies. An investigation relying on the results of a single forensic technique, exclusive of other available tools, is frequently successfully challenged when contrary evidence based on multiple forensic approaches is introduced. When forensic evidence is arrayed as multiple, but independent lines of evidence, a stronger scientific case, less susceptible to scientific challenge, emerges. This book is intended to provide you with your own toolbox of forensic techniques.

Brian L. Murphy

Robert D. Morrison

CONTRIBUTORS

EDITORS



Robert D. Morrison, Ph.D. is a soil physicist with over 36 years of environmental consulting experience. Dr. Morrison has a B.S. in Geology, an M.S. in Environmental Studies, an M.S. in Environmental Engineering and a Ph.D. in Soil Physics from the University of Wisconsin at Madison. Dr. Morrison published the first book on environmental forensics in 1999, is the editor of the Environmental Forensics Journal and is a co-founder of the International Society of Environmental Forensics (ISEF). Dr. Morrison specializes in the use of environmental forensics for contaminant age dating and source identification.



Brian L. Murphy, Ph.D. is a Principal Scientist at Exponent. He is the coeditor of Environmental Forensics: Contaminant Specific Guide, also with Robert Morrison. He is the author of more than thirty journal publications and is on the editorial board of the journal *Environmental Forensics*. Dr. Murphy has an Sc.B. from Brown University and MS and Ph.D. degrees from Yale University. All his degrees are in physics. He has spent most of his career as an environmental consultant, including as General Manager for Physical Sciences at what is now ENSR, founding President of Gradient Corporation, and Vice President of Sciences International. He also has had Visiting Instructor positions at Harvard School of Public Health and the University of South Florida. His practice focuses on using environmental forensics to reconstruct

past contaminating events, for a wide variety of contaminants, either for the purpose of remedial cost allocation or in order to determine doses in toxic torts.

CHAPTER AUTHORS



Jean Christophe Balouet holds a Ph.D. in Earth Sciences, and is manager, principal owner of Environment International, Le Perreux-sur-Marne. He also serves Université, Paris VII, France. He has developed dendroecological applications to environmental forensics, as well as adapted spectroscopic methods such as EDXRF or LIBS for the measure of elemental tracers in such proxy-recorders and successfully applied it to 27 different sites. As a research scientist, he has served the Smithsonian Institution, the Paris Museum of Natural History and several major research institutions worldwide. He has served and chaired several new international environmental standards and is the author of about 100 peer-reviewed publications and books. Dr. Balouet has over 25 years of international experience as a consultant/expert scientist serving governmental and intergovernmental authorities, including the United Nations Environment Program and its Technology, Industry and Economy Division. He also serves NGOs and private parties, including major industries in environmental issues, including consultancy, expertise and environmental forensics.



Shelley Bookspan, Ph.D. is currently a consultant in Santa Barbara, California. In 1982, she was among the founders of PHR Environmental Consultants, Inc. She is past president of PHR, which pioneered the application of historical methodology to pollution liability matters. Dr. Bookspan's doctorate, from the University of California, Santa Barbara, is in history, with emphasis on American history, history of technology, and urban history. She also holds a master's degree in city planning from the University of Pennsylvania and a master's degree in history from the University of Arizona. Dr. Bookspan is widely published both in the fields of history and environmental consulting, and she is now working in multimedia presentation and video production.



Richard Brown, M.S. is a microscopist and an Executive Director at MVA Scientific Consultants in Atlanta, Georgia. Since 1989 Rich has applied light and electron microscopy to small particle problems in the Environmental Forensics field to characterize and identify the particles and determine their source(s). Rich received his Master's degree in Forensic Chemistry and worked as a Criminalist doing crime scene investigation and trace evidence examination at the Orange County Sheriff-Coroner Department in Santa Ana California. Rich has testified in court as an expert witness and is a member of the California Association of Criminalists, the American Academy of Forensic Sciences and is a diplomate of the American Board of Criminalistics.



Judith C. Chow is a Research Professor in the Desert Research Institute's (DRI's) Division of Atmospheric Sciences specializing in aerosol measurement, method development, and applications. At DRI, Dr. Chow established an Environmental Analysis Facility that quantifies mass, elemental, ionic, and carbon concentrations on filter samples. She has prepared guidance documents on aerosol measurement for the U.S. EPA, authored the 1995 Air & Waste Management Association's Critical Review of $PM_{2.5}$ measurement methods, and has completed more than 100 research publications on sampling and analysis methods and interpretation of results. She is a member of the National Academy of Engineering's U.S. Committee on Energy Futures and Air Pollution in Urban China and the United States and was the 1999 recipient of DRI's Alessandro Dandini Medal of Science. Dr. Chow received her Sc.D. degree in Environmental Science from the Harvard School of Public Health in 1985.



Julie Corley is a Senior Managing Consultant with LECG, LLC. Prior to joining LECG, Ms. Corley was the Director of Research for PHR Environmental Consultants, Inc., a consulting firm specializing in professional historical research conducted in support of environmental, product liability, and other legal disputes. She

has worked for PRP groups as well as private and public sector clients in support of historical fact-finding endeavors to answer questions about the effects of past operations on current conditions.



Gregory S. Douglas received M.S. and Ph.D. degrees in Oceanography from the University of Rhode Island and has over 25 years of experience in the field of environmental chemistry. He has designed, implemented, managed, audited, and defended a wide range of environmental forensic chemistry studies for government and industry concerning complex petroleum and fuel contamination issues in marine and soil/ground water systems. Dr. Douglas has performed extensive research concerning the fate of gasoline NAPL and gasoline additives in groundwater. He has prepared interpretive reports on more than 100 site or incident investigations involving petroleum product source and age dating issues, and has published and presented on the development and application of environmental forensics analytical methods, source identification and allocation within complex contaminated environments.



James I. Ebert, Ph.D. is an archaeologist, anthropologist, and forensic scientist. He is Chief Scientist at Ebert & Associates, Inc., an Albuquerque, New Mexico firm specializing in forensic, environmental, and archaeological applications of photogrammetry, photo analysis, remote sensing, and digital mapping and imaging technologies. He is a Certified Photogrammetrist (American Society of Photogrammetry and Remote Sensing), a fellow of the American Academy of Forensic Sciences, and a member of the New York State Police Medicolegal Investigations Unit. Since 1989 Dr. Ebert has conducted ongoing archaeological and environmental research at Olduvai Gorge in Tanzania with Rutgers University's Olduvai Landscape Palaeoanthropological Project.



Bruce Egan, Sc.D., CCM is President of Egan Environmental, Inc., a small business that specializes in providing air quality consulting services to both public and private sectors. Prior to forming EEI, he was Senior Vice President and Chief Scientist at the ENSR Corporation and Vice President and Technical Director at Woodward Clyde Consultants. A graduate of Harvard College and University, Dr. Egan's formal training is in Fluid Mechanics, Thermodynamics, Meteorology, and in Environmental Health Sciences. He has been a Visiting Lecturer at the Harvard School of Public Health. His specialties include air toxics, hazard assessments, air dispersion modeling and permitting in complex terrain settings, and compliance demonstrations for facilities subject to state and federal regulations. Dr. Egan is a Certified Consulting Meteorologist, and an elected Fellow of the American Meteorological Society. He is also a Fellow of the Air and Waste Management Association. He has managed projects for the U.S. EPA and for trade associations on the development of air quality dispersion models that are used for New Source Review permitting and has consulted with a large number of clients in the power production, energy, chemical and paper industry sectors on risk management plans and industrial hazard litigation cases. He has provided expert testimony and depositions in a number of litigated cases. He is an author of over 80 technical papers and has served on many professional organization committees and government sponsored panels on air quality and hazard assessment matters.



Robert Ehrlich, Ph.D. For the past 30 years, Dr. Robert Ehrlich has been a leader in applying image analysis and pattern recognition procedures in the earth and environmental sciences. Dr. Ehrlich has been editor of the journal Mathematical Geology and has published more than 100 papers on the application of numeric techniques including development of new algorithms. He was a Professor at the University of South Carolina, Department of Geological Sciences from 1974 to 1997. His M.S. and Ph.D. degrees are from Louisiana State University. He is presently President of Residuum Energy Incorporated and heads Residuum's Salt Lake City office.



Stephen Emsbo-Mattingly holds an MS in Environmental Science from the University of Massachusetts and more than 18 years of experience in environmental chemistry and forensics. He specializes in the identification of tar, creosote, petroleum, PAHs, PCB, PCP, solvents, and other compounds of concern in various media. Mr. Emsbo-Mattingly's recent research and publications focus on the formation and weathering of PAHs and PCBs with an emphasis on the transfer of source signatures between sediment, tissue, and vapor phases. His investigations often involve the simultaneous identification of various other contaminant sources, like non-point releases from storm sewer discharges and wastewater treatment plants. Most of these projects use historical laboratory data, multivariate statistics, and GIS-based spatial analysis.



William E. Full, Ph.D. has spent his career in academia and industry focusing on computer applications and algorithm development in the geosciences. He has B.S. in mathematics from the University of Notre Dame, and M.S. and Ph.D. degrees in Geology from University of Illinois, Chicago and the University of South Carolina, respectively. In 1982 he received the Andre Borisovich Vistelius Research Award for Young Geomathematicians from the International Association of Mathematical Geology. He has published many papers on information entropy and multivariate analysis, and he was the principal developer of the Polytopic Vector Analysis procedure. Dr. Full was on the faculty of the Department of Geology at Wichita State University for 15 years, and is currently a Distinguished Professor at the University of Rome, Italy, and President of WEF consulting in Wichita, Kansas.



Thomas D. Gauthier, Ph.D. is a Senior Science Advisor at ENVIRON International Corp. in Tampa, Florida. He has over 15 years of consulting experience and works on projects involving the transport and fate of chemicals in the environment, historical dose reconstruction, source identification and the statistical analysis of environmental chemistry data. Dr. Gauthier received his BS in chemistry from Merrimack College and Ph.D. in analytical chemistry from the University of New Hampshire. He is a member of the American Chemical Society and is a frequent reviewer for the *Journal of Environmental Forensics.*



A. Mohamad Ghazi, Ph.D. is an analytical geochemist. His major area of research is in environmental geochemistry with special focus in environmental forensics involving the application of chemical concepts to interpretation, distribution, speciation, impacts and bioavailability of chemicals in the environment and wide range of other media. With more than 20 years of laboratory, field and classroom experience, Dr. Ghazi has made significant contributions in the developing and application of laser ablation ICP-mass spectrometry for *in-situ* analysis of material. Currently he is a geologist with the DoD Remediation Unit Georgia Environmental Protection Division. Previously he was a research scientist and an associated professor in Geology and the Director of the Laser Ablation Plasma Mass Spectrometry at Georgia Sate University. Dr. Ghazi is the author and co-author of over 70 peer reviewed articles and book chapters. He has been the principal investigator on 15 successful research proposals to National Science Foundations and other private funding agencies. He is an honorary member of Sigma Xi International Scientific Society. In addition, he is an associate editor for the International Journal of Environmental Forensics and a member of Geological Society of America, American Geophysical Union and American Geochemical Society.



Michael E. Ginevan, Ph.D. a Principal Scientist at Exponent Inc., is an internationally recognized expert in health and environmental statistics, with more than 25 years experience in the application of statistics and computer modeling to problems in public health and the environment, and in the conduct of environmental, epidemiologic, and risk assessment studies. He is the author of "Statistical Tools for Environmental Quality Measurement," and over 60 other publications in the areas of statistics, computer modeling, environmental studies, and epidemiology. He is a former Deputy Director of the Office of Epidemiology and Health Surveillance at the U.S. Department of Energy, is a founder and past Secretary of the American Statistical Association (ASA) Section on Statistics and the Environment, a recipient of the Section's Distinguished Achievement Medal, and is a Charter Member of the Society for Risk Analysis. He has served on numerous review and program committees for ASA, the U.S. Department of Energy, The U.S. Nuclear

Regulatory Commission, the National Institute for Occupational Safety and Health, the National Cancer Institute, and the U.S. Environmental Protection Agency.



A.J. Gravel is the Managing Director of LECG's Forensic History and Analysis Group. Mr. Gravel has managed the execution of hundreds of domestic and international environmental, products liability, and other investigations, and litigation support projects. He specializes in matters involving retrospective analysis and environmental cost recovery. His environmental projects have dealt with numerous CERCLA, RCRA, State Voluntary Cleanup Program and toxic tort related issues such as Potentially Responsible Party (PRP) identification; corporate succession and asset searches; land use and business operation histories; past industrial chemical generation, usage, and disposal analysis; cost allocation analysis; insurance investigations and government involvement studies. His work has involved the examination of historical contamination and/or product liability issues for petroleum, chemical, shipbuilding, tire manufacturing, railroad and utility industry clients.



Mark Hawley, Ph.D. is a Senior Science Advisor at ENVIRON Corporation in Arlington, Virginia. His technical specialties include hydrology, statistics, modeling, and interpretation of environmental data. Dr. Hawley received a B.S. degree in Geology from Rensselaer Polytechnic Institute and M.S. and Ph.D. degrees in Civil Engineering (Water Resources/Hydrology) from the University of Maryland, and was formerly an Assistant Professor at the University of Virginia. Since joining ENVIRON in 1989 he has managed and participated in a wide variety of projects related to site investigation and remediation, risk assessment, product safety, and management of industrial and hazardous wastes. He has served as an expert witness or provided litigation support in cases involving the fate and transport of potentially hazardous substances including petroleum products, chlorinated solvents, metals, and PCBs.



Emilie Jardé, Ph.D. received her Ph.D. in organic geochemistry from the University of Nancy (France) in 2002. Her Ph.D. dealt with the organic composition of various sewage sludges originating from Lorraine (North-East of France): molecular characterisation and biodegradation effects. She then spent two years as a post-doctoral fellow with Professor G. Gruau at the University of Rennes (France) still working in organic environmental geochemistry and more specifically on the identification of the origin of organic matter in superficial waters from Brittany. After that, she spent a few months at the University of Oklahoma, Norman (USA), as a post-doctoral fellow with Professor R.P. Philp to learn more about the use of stable isotope in monitoring organic pollutants in the environment. In July 2006, she joined the group of G. Gruau at the University of Rennes (France) as a CNRS-researcher. The major theme of her research is the molecular analysis of organic matter from sediments, water, soils, sewage sludge, and animal slurry in environmental studies. To achieve that, she uses the correlation between organic compounds and sources by the identification of specific molecular marker or distribution from typical origins with gas chromatography-mass spectrometry or pyrolysis methods.



Glenn W. Johnson, Ph.D. received his M.S. from the University of Delaware in 1988 and his Ph.D. from the University of South Carolina in 1997. Both graduate research programs involved application of multivariate techniques to geological and environmental chemical data. In the seven years between degrees, he worked as an environmental consultant with Roux Associates, Inc. and McLaren/Hart Environmental Engineering Corp., where much of his work focused on environmental forensics and litigation support. Dr. Johnson is currently a Research Associate Professor at the Energy and Geoscience Institute at the University of Utah. His work at EGI includes research in chemometrics and environmental forensics, as well as teaching within the Department of Civil and Environmental Engineering. Dr. Johnson is a registered professional geologist, and a member of the Society of Environmental Toxicologists and Chemists and the International Society of Environmental Forensics.



Ashok K. Katyal, Ph.D. is Principal Engineer and President of Resources & Systems International, Inc. Ashok received his Ph.D. in Engineering Science from Washington State University. He is recognized as an expert in multiphase flow, surface-groundwater interaction, and watershed management. He is an author of BIOFT3D, BIOSLURP, MARS, MOVER, MOFAT, SVE_3D, and several more software programs currently used in the environmental industry in many countries. Ashok is a consultant to several national and multinational companies.



Kevin J. McCarthy has over 20 years experience in the field of petroleum environmental chemistry. His experience is in the detailed chemical analysis and chromatographic interpretation of petroleum products and wastes and their characterization using advanced instrumental methods and chemometric data analysis techniques (forensic chemical fingerprinting). His expertise includes the molecular-level characterization of petroleum and petroleum products and investigations of the chemical alteration of petroleum due to physical and biological weathering. He has participated in developing specialized methodologies for the analysis and characterization of oil and petroleum products.



James R. Millette, Ph.D. is an Executive Director of MVA Scientific Consultants in Atlanta, Georgia. Since 1972, Dr. Millette's research focus has been the investigation of environmental particles using microscopy techniques and has published over 60 peer-reviewed articles in the area. Dr. Millette's previous work included 11 years as a research scientist at the United States Environmental Protection Agency Research Center in Cincinnati, Ohio, and 5 years at McCrone Environmental Services performing and supervising analysis of particulates in various media. Dr. Millette is a full member of the American Academy of Forensic Scientists and has testified in court on several occasions concerning environmental science and particulate analysis.



Gil Oudijk is the owner of Triassic Technology, Inc., a consulting firm located in Hopewell, New Jersey (USA) and founded in 1994. He is a graduate of the Pennsylvania State University and has close to 25 years experience in the hydrogeological field with special emphasis on ground-water pollution problems and forensic techniques such as the age-dating and fingerprinting of contaminant releases. He is presently an Associate Editor for the *Journal of Environmental Forensics* and a task leader for the American Society of Testing & Material's (ASTM's) standards committee on environmental forensics.



Ioana Petrisor, Ph.D. has a PhD in Biology/Environmental Biotechnology from Romanian Academy of Sciences (awarded in 2000) and a Bachelor in Chemistry with a Major in Biochemistry from Bucharest University in Romania, Faculty of Chemistry (awarded in 1992). In December 1999 she has completed an UNESCO training program (3 months) on Plant Molecular Genetics at the University of Queensland, Department of Botany in Brisbane, Australia. She is co-author of 63 scientific papers published in peer-review Journals and Proceedings and of one invention patent. She has more than 12 years of experience (both academic and industry) in the Environmental Engineering/Biotechnology field. She is Managing Editor for the Environmental Forensics Journal and a member of the Editorial board of several other peer-review Journals. She was recently elected as Vice-Chairman of the newly formed ASTM committee on Forensic Environmental Investigations (subcommittee of Environmental Assessment, Risk Management and Corrective Actions main E50 committee). She is a member of ACE, AEHS, ITRC (Vapor Intrusion and UXO teams) and other professional organizations. Her work experience includes managing and conducting innovative research (at lab, field and pilot scales) for U.S. DOE and the European Community on the topics of bioremediation and phytoremediation, environmental characterization and risk assessment. She is now involved in forensics studies aiming to identify the source and age of contamination, and subsequently allocating responsibilities between different potential responsible parties for case studies involving gasoline and other petroleum products releases, PCB contamination, as well as chlorinated solvents and heavy metals.



R. Paul Philp, Ph.D. is Professor of Petroleum and Environmental Geochemistry at the University of Oklahoma. He received his Ph.D. from the University of Sydney, Australia in 1972 and a D.Sc. from the same University in 1998 on the basis of his research in geochemistry over the past 20 years. Prior to starting at the University of Oklahoma in 1984 Dr. Philp was a Principal Research Scientist, C.S.I.R.O., Sydney, Australia. His current research interests center around petroleum, environmental and forensic geochemistry with an emphasis on molecular and isotopic characterization of oils, gases, rock extracts and contaminants for the purposes of source determination, characterization of depositional environments, maturity, biodegradation and for correlation purposes. Much of the current research activity in the area of forensic geochemistry involves the use of stable isotopes for the purposes of fingerprinting contaminants in the environment for correlation purposes; source determinations and evaluating whether or not natural attenuation is active. This approach is particularly valuable in the case of refined products or single component contaminants when the more traditional GC and GCMS techniques are of little or reduced use. He has authored or co-authored over 340 articles and books and has lectured extensively on petroleum and environmental geochemistry in SE Asia, South America, Europe and Africa.



Scott Ramos, Ph.D. is an analytical chemist with experience in environmental and trace hydrocarbons, natural products and chemometrics. He has a B.S. degree in chemistry from MIT, an M.S. in environmental science from Washington State University, and a Ph.D. in analytical chemistry and chemometrics from the University of Washington. Scott has worked for NOAA's National Marine Fisheries Service, in Seattle, at FEEMA, the state pollution control agency in Rio de Janeiro, and the federal Amazon Research Institute (INPA) in Manaus, both in Brazil. During the last 22 years, he has been the chief scientist at Infometrix, Inc, in Bothell, WA. Publications include studies on contamination by polycyclic aromatic hydrocarbons, essential oil characterization, and various applications of chemometrics.



Charles Ramsey is the founder and President of EnviroStat, Inc., a company that provides training in sampling and laboratory subsampling for defensible environmental decisions to federal and state agencies as well as private companies. He has a BA in Chemistry from the University of Denver and a MS in Environmental Science from the Colorado School of Mines. Mr. Ramsey has 20 years sampling experience including seven years with the National Enforcement Investigations Center (NEIC) of the USEPA. While with the EPA, Mr. Ramsey provided sampling and statistical expertise on all major environmental regulations. Mr. Ramsey is involved with the development of guidance documents for sampling and statistics.



Scott A. Stout is an organic geochemist with nineteen years of petroleum and coal industry experience. He has extensive knowledge of the chemical compositions of coal-, petroleum-, gasoline-, and other fuel-derived sources of contamination in terrestrial and marine environments. Dr. Stout has written interpretive reports on more than 250 site or incident investigations and has authored or coauthored nearly 100 papers published in scientific journals and books. He has conducted environmental research while employed at Unocal Corporation, Battelle Memorial Institute, and is currently a partner at NewFields Environmental Forensics Practice, Rockland, Massachusetts.



Allen D. Uhler has over 25 years experience in environmental chemistry. He has developed advanced analytical methods for petroleum-, coal-derived and anthropogenic hydrocarbons and other man-made organic compounds in waters, soils, and sediments, vapor and air. He has conducted assessments of the occurrence, sources, and fate of fugitive petroleum at refineries, offshore oil and gas production platforms, bulk petroleum storage facilities, along petroleum pipe-

lines, at varied industrial facilities, and in sedimentary environments. He has studied coal-derived wastes at former manufactured gas plants, wood-treating facilities, and in nearby sedimentary environments. His experience includes expertise in the measurement and environmental chemistry of man-made industrial chemicals including PCB congeners and Aroclors, persistent pesticides, dioxins and furans, metals, and organometallic compounds.



Dallas Wait, Ph.D. is a chemistry expert at Gradient Corporation with over 28 years of experience evaluating the source and fate of chemicals in the environment, characterizing consumer products, designing test method and quality assurance programs, interpreting data, and determining the reliability of chemistry measurements and sampling procedures. Dr. Wait's consultations often resolve data quality issues, aid in agency negotiations concerning data usability, and provide pivotal chemistry testimony. He serves on the editorial board for two peer-reviewed journals and coauthored the second edition of EPA's SW 846 Test Method Manual. Dr. Wait is a member of the Scientific Advisory Board for the International Conference on Soils, Sediments and Water and for eight years was the Chairperson for either the *Risk*, *Forensic* or *Analysis* sessions of the conference. He is a member of numerous scientific work groups involved with developing and evaluating test methods and quality assurance programs, such as ASTM and AOAC. Before joining Gradient in 1989, he was Technical Director, Vice President and cofounder of ENSECO's ERCO Laboratory, a nationally prominent environmental laboratory involved, in part, with oil spill research, agency method development studies, aquatic toxicology GLP testing support, and site investigations. Dr. Wait received his BS and Ph.D. degrees in Chemistry from the University of Rhode Island.



John G. Watson, Ph.D. is a Research Professor in the Desert Research Institute's (DRI's) Division of Atmospheric Sciences specializing in the characterization, source apportionment, and control of suspended particles that cause adverse health effects and regional haze. Dr. Watson has developed theoretical and empirical models that, when coupled with appropriate measurements, quantify contributions for pollution from different sources. With his colleagues, he has applied these

methods in urban and regional aerosol studies to solve problems of excessive concentrations, visibility impairment, and deposition. Dr. Watson obtained his Ph.D. degree in Environmental Science from the Oregon Graduate Institute in 1979 and was awarded the Howard Vollum Prize for Distinguished Achievement in Science and Technology in 1989, DRI's Alessandro Dandini Medal of Science in 1992, and the Air & Waste Management Association's 2000 Frank A. Chambers Award for major contributions to the science and art of air pollution control.

CHAPTER 1

APPLICATIONS OF ENVIRONMENTAL FORENSICS

Brian L. Murphy

1.1 Introduction 1
1.2 Liability Allocation at Superfund Sites 2
1.2.1 Equivalence of Harm and Risk 4
1.2.2 Allocation Principles 5
1.3 Environmental Site Assessment 7
1.4 Insurance Litigation 8
1.4.1 Imminence of Off-site Migration 9
1.4.2 Trigger of Coverage 9
1.4.3 Expected and Intended 10
1.4.4 Sudden and Accidental 11
1.4.5 Equitable Cost Sharing 12
1.5 Toxic Torts 12
1.5.1 Epidemiology 13
1.5.1.1 Association and Causation 13
1.5.1.2 Texas Sharpshooter Effect 14
1.5.1.3 Statistical Significance 15
1.5.2 Differential Diagnosis 16
1.5.3 Risk Assessment 17
1.6 Natural Resource Damage Assessment 18
1.7 Marine Oil Pollution 19
Acknowledgments 20
References 20

1.1 INTRODUCTION

There are two separate motivations for environmental forensic studies. First, such studies may be performed for the sake of obtaining knowledge of historical emissions to the environment or historical environmental processes and for no other reason, what might be termed purely research or academic studies. Second, such studies are carried out to determine liability in a variety of contexts. This latter purpose is the focus of this chapter.

Table 1.1
Liability context and related forensic issues.

Context	Forensic Issues
Cost allocation at Superfund sites	Responsibility for waste streams or areas of contamination Contribution of waste streams to remedial cost or to the need for a remedy
Site investigation for property transfer	Existence and extent of contamination Cost of remediation Existence of other responsible parties
Insurance litigation	Policies triggered by a release Characteristics of a release or of contamination relative to contract language Equity of liability allocation
Toxic tort	Probability that chemical exposure caused manifest or latent injury
Natural resource damage	Distinguishing anthropogenic from natural effects
Assessment	Monetizing anthropogenic damages
Marine oil pollution	Determining spill source

Our discussion of liability-driven forensic studies is based on liability under U.S. laws. However, we do not focus on the law itself, although some of the references given discuss various legal issues. Rather we focus on how the legal requirements concerning liability translate into technical issues and questions, which can be answered using forensic methods.

We discuss liability in six different contexts, as shown in Table 1.1. This table also lists key forensic issues for each context. In the remainder of this chapter we describe how measurements of chemical concentrations or other properties combined with the forensic techniques described later in this book can be used to illuminate these issues.

These six contexts probably represent a good proportion of the situations in which the tools of environmental forensics are employed to allocate liability. However, they do not represent all such situations. For example, environmental forensics techniques also are used to identify air pollution sources, including in international or transboundary air pollution situations. Techniques relevant to air pollution sources are discussed elsewhere in this text, particularly Chapters 8 and 12. We have selected these six contexts because they are the most structured and universally applicable in the United States.

1.2 LIABILITY ALLOCATION AT SUPERFUND SITES

The Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), commonly referred to as Superfund, prescribes specific procedures for dealing with chemical release or disposal sites that are considered

to be the most hazardous in the United States. States also have hazardous waste site remediation programs patterned to varying degrees after the federal program. Thus state lead occurs at many sites deemed less hazardous than those in the federal Superfund program, for example at dry cleaning facilities across the country. Also, because the federal Superfund legislation excludes petroleum release sites, state lead occurs at facilities such as gasoline stations and former manufactured gas plant sites.

State laws often are modeled after CERCLA, although as discussed later, CERCLA does not provide much guidance on liability allocation. In any case, the discussion in this chapter of methods of liability allocation at federal Superfund sites is still relevant but the details may vary from state to state.

Potentially responsible parties (PRPs) at Superfund sites include present owners and operators, past owners and operators, waste generators, and transporters or arrangers for transport of waste. Costs borne by PRPs at CERCLA sites may be for site remediation or in payment for past, present, and future damage to natural resources. Payment for future damages arises when the site cannot be totally remediated so that the habitat is restored.

Superfund liability includes for actions that predate the CERCLA legislation. Furthermore, liability is perpetual; it cannot be circumvented by being assigned to someone else. Liability does not depend on fault but simply on being a member of one of the classes of PRP just described. Finally, Superfund liability can be *joint and several*; that is, in principle all liability may be borne by a single PRP irrespective of the relative degree of fault.

Two sections of CERCLA touch on allocation of liability among the PRPs. Section 107 provides for recovery of remediation costs. Plaintiffs in a recovery action may be the U.S. Environmental Protection Agency (EPA) or states. Courts are divided on whether a PRP may be a plaintiff but the recent trend has been to deny Section 107 to PRP plaintiffs (Aronovsky, 2000). Although Section 107 specifies joint and several liability, this is discretionary with the court. In particular, where a PRP can demonstrate distinct harm or divisibility of harm, that party may be responsible just for their contribution to harm. A distinct harm arises, for example, when there are separate groundwater plumes or areas of surface soil contamination. A divisible harm might be where there are successive site owners conducting the same operation. The basis for divisibility in that case might be the relative number of years of operation.

Section 113 of CERCLA allows a party who has incurred response costs to seek contribution from other PRPs.[1] This section also provides contribution protection for parties that have settled with the United States. Under Section 113 the liability of nonsettling PRPs is limited to their proportionate share. The nonsettling PRP's liability may be determined in either of two ways. It may be determined by subtracting out the amount of prior settlements or by

[1]In 2004 the United States Supreme Court in *Cooper Industries, Inc. v. Aviall Services, Inc.* that the right of Superfund contribution is only available to PRPs who have been subject to a federal "civil action" under CERCLA Sections 106 or 107. Settlement with the government does not qualify as a civil action and does not permit a PRP to seek contribution from other PRPs. The Court left open the issue of whether contribution from other PRPs could still be sought in that case under Section 107. However, only a few lower courts have recognized a right of PRP cost recovery under this section. The net effect has been to provide on incentive for PRPs to wait until they are sued before incurring response costs.

subtracting out the proportionate share of the harm for the settling PRPs. As Ferrey (1994) points out, the results of these two approaches may be quite different. No guidance in determining proportionate shares, beyond citing equitable factors, is found in CERCLA.

The equitable factors most often cited are the Gore factors proposed by then-Representative Albert Gore in 1980 but not enacted. They are: (1) the ability to distinguish the party's contribution to the nature and extent of the problem, (2) the degree of the party's involvement in the activities that caused the problem, (3) the degree of care exercised by the party, (4) the degree of cooperation of the party with governmental agencies, (5) the quantity of the hazardous waste involved, and (6) the toxicity of the waste. These factors have been found to be far from sufficient and in some cases may not be applicable at all. Furthermore, they are simply a list; they provide no conceptual framework for allocation.

Other factors that have been suggested include: (7) existing contracts between the parties, (8) the owner's acquiescence in the operator's activities, and (9) the benefit to the owner from the increase in land value due to remediation.

Thus CERCLA and its legislative history are not particularly helpful in specifying how liability is to be allocated. Courts generally have determined that there is a presumption of joint and several liability unless harm is distinct or there is a reasonable basis for its division.

1.2.1 EQUIVALENCE OF HARM AND RISK

What is harm at a Superfund site? It seems logical that it be closely identified with the concept of risk. A baseline risk assessment is conducted at all Superfund sites. Removal actions may precede completion of the risk assessment for urgent matters but the continued remediation of a site is based on a finding that the computed baseline risks, either human or ecological, are unacceptable. Therefore, it is logical to identify the risks at a site, including those requiring removal action with harm. As discussed by Rockwood and Harrison (1993) several federal circuit court rulings also support this notion.

Thus, an argument can be made that risk assessment is the appropriate tool to use in apportioning liability. However, in fact risk is often not a consideration in apportioning liability. If the PRPs at a Superfund site agree on an allocation scheme, then that scheme is by definition satisfactory, assuming that there is no second guessing by other parties such as insurers. PRPs often decide to allocate liability based on the contribution of each to the cost of the remedy. Of course, surrogate measures for estimating contribution to costs may be used, such as counting barrels or estimating plume areas.

There are several common situations where the harm due to multiple PRPs is not distinct but requires division; for example, (1) commingled groundwater plumes, (2) hazardous waste disposal sites with multiple users, and (3) successive site ownership. If these situations result in contamination by similar chemicals, a straightforward allocation based on contribution to the cost of a remedy may make sense. However, when one or more PRPs' wastes differ significantly from the others in the risk they pose, those PRPs may wish to consider a risk-based approach.

1.2.2 ALLOCATION PRINCIPLES

Economic principles of cost allocation, including the stand-alone cost method, have been discussed by Butler *et al.* (1993) and by Wise *et al.* (1997). The cost allocation matrix approach of Hall *et al.* (1994) also is based on determining contribution to the cost of a remedy. Marryott *et al.* (2000) present a stand-alone cost type model in which a weighted sum of contaminant mass in the plume and plume volume serves a surrogate for remediation costs. The basic equation for calculating stand-alone costs is:

$$f_i = \frac{SAC_i}{\sum_i SAC_i} \tag{1.1}$$

where SAC_i is the stand-alone cost for the waste stream due to the ith PRP generator/transporter.

This equation does not address how liability is to be allocated between the generator, transporter, and site owner, nor does it address orphan shares, such as from unidentified or defunct parties. It is solely an allocation by waste stream. Equation 1.1 states that each PRP pays in proportion to the cost that would have been incurred if there were no other PRPs at the site. Because of redundancy of cost items and economies of scale the total cost of a remedy will generally be less than the denominator of Equation 1.1 and hence each PRP actually will pay less than their computed stand-alone cost.

Risk-based allocation methods have been discussed by Murphy (1996, 2000) and by Mink *et al.* (1997). The risk contribution analogue to Equation 1.1 is:

$$g_i = \frac{SAR_i}{\sum_i SAR_i} \tag{1.2}$$

where g_i is the cost fraction for the ith generator/transporter based on stand-alone contribution to risk SAR_i. The analogy with stand-alone costs is

incomplete; however, the total risk is equal to the sum of the individual PRP-caused risks rather than generally being less.

Of course cost allocation may be a mixture of cost-based and risk-based methods:

$$h_i = \alpha f_i + (1-\alpha) g_i \tag{1.3}$$

where α is a constant. As α decreases from 1, a "contribution to the need for a remedy" component is mixed in with the "contribution to the cost of a remedy."

The kind of information needed to calculate f_i or g_i differs. For example, in computing stand-alone costs, well installation costs may vary as plume area and groundwater treatment costs may vary as contaminant mass in the plume. How long a pump and treat remedy needs to be maintained will depend on the ratio of individual chemical concentrations to acceptable levels in groundwater and on chemical properties that determine partitioning to soil. In computing stand-alone risks, concentrations and toxicities of specific chemicals will be required. Of course, as indicated earlier, it may be to the advantage of all PRPs to lower transaction costs by using surrogate quantities rather than attempting to collect the additional information necessary for refined or precise calculations.

For a cost-based allocation, typical forensic issues are:

- Attributing different groundwater plumes to individual parties or where plumes are inextricably commingled to two or more parties.
- For successive site owners, determining when major releases occurred, or for contamination by chronic operating discharges determining relative production amounts or years of operation.
- At hazardous waste sites accepting waste from multiple parties, determining waste stream volumes attributable to individual generators or transporters.

The additional information needed for a risk-based allocation is concentrations of specific chemicals in groundwater plumes, waste streams, or historical releases.

Time is a missing factor in many allocations whether by risk or by cost. For example, a PRP's wastes in groundwater might not arrive at an extraction point for many years because of a slow groundwater velocity or retardation effects. If the remedy will not be relevant to that PRP's wastes until some possibly distant future time, it can be argued that that PRP's contribution should be discounted to a smaller present value.

1.3 ENVIRONMENTAL SITE ASSESSMENT

As the term is used in this chapter, an environmental site assessment is conducted as a preliminary to a real estate transfer. (Similar tasks may be conducted as part of an internal management assessment, a process generally known as an *environmental audit*. An audit may be concerned solely with compliance with applicable laws and regulations or it may include a more management-oriented review of responsibilities, organization, communications, and measurement of progress.) The main purposes of an environmental assessment are to determine:

- Whether contaminants are present on site
- If present, the extent of contamination so that likely remediation requirements and costs can be estimated

The American Society for Testing and Materials (ASTM) has published two Standard Practices for conducting Phase I Assessments. Phase I is intended to assess the likelihood of site contamination. As the term is used in these standard practices, a Phase I Assessment does not include any environmental sampling. These Standard Practices originally were developed to satisfy one of the requirements for the innocent landowner defense under CERCLA.

ASTM Standard Practice E 1527 describes the four components of a Phase I site assessment as on-site reconnaissance, interview of site owners and occupants as well as local government officials, records review, and report preparation. This Standard Practice is intended to be conducted by an environmental professional. ASTM Standard Practice E 1528 on the other hand may be conducted by any of the parties to a real estate transaction as well as an environmental professional. This Standard Practice is based on a transaction screen process, consisting of the same three components prior to report preparation: a site visit, questions for the owner/occupants, and a records review. The difference is that the questions or issues to be addressed during the conduct of these components are all prescripted.

In the ASTM description, sampling of soils, groundwater, or other media would be a Phase II Assessment. ASTM has published a framework for the Phase II Assessment as Standard Guide 1903–97, and has published a number of standards dealing with sampling methods. These have been collected in the document ASTM Standards Related to the Phase II Environmental Site Assessment Process. A Phase II Assessment is generally necessary in order to determine the extent of contamination and hence the likely remedial requirements and associated costs. The Phase II Assessment would be guided by the results of Phase I.

The ASTM descriptions provide a framework but not one that should be followed slavishly. For example, at some sites the necessity of sampling certain locations and media may be evident and it may make the most sense to conduct sampling simultaneously with the components of a Phase I Assessment. Similarly, if one or more potential fatal flaws are obvious, the Phase I or Phase II Assessments may focus solely on those areas.

As noted earlier, if contamination is found, an understanding of the extent and options for remediation to regulatory acceptable limits becomes important. If the cost of remediation and the uncertainties are determined, this may become the basis for structuring a deal by allocating risks between the parties. Ideally, one would like a description of the complete spectrum of cost possibilities and their associated probabilities. An estimate of the expected time to regulatory closure and the associated uncertainties may also be factored in.

Environmental forensics enters into the site assessment process in several ways. First, in Phase I the site use history, as revealed by interviews and records, and visual clues during reconnaissance are combined with the analyst's knowledge of specific industrial operations to develop expectations of the presence and type of contamination. Second, in Phase II this information is augmented by sampling data to determine the extent of contamination. Finally, determining who is responsible for the contamination may involve other parties and hence introduce other remediation cost-sharing options. For example, groundwater contamination under a site in fact may originate from off-site sources.

1.4 INSURANCE LITIGATION

Insurance claims are based on the contract language between insurer and insured. Contracts until the mid-1980s were based on comprehensive general liability (CGL) policies. Subsequently, environmental impairment liability (EIL) policies were introduced to deal specifically with contamination and other environmental issues. Interpretation of the language is governed by state law and can vary greatly. However, the same phrases in the contract and the same issues produce the need for forensic information in any state in order to determine matters of fact.

Insurance coverage for damages associated with chemical contamination in the environment may depend, among other things, on the imminence of off-site migration, whether coverage was triggered during a policy period, and whether the release was expected and intended, or sudden and accidental. When multiple parties have contaminated a site, equitable cost sharing may also be a coverage issue (Murphy, 1993).

Parties may agree on the facts but still produce different descriptions for the same facts in order to construe the policy language most effectively. Several examples of this are noted in the Trigger of Coverage and Sudden and Accidental sections. Although there may be no correct point of view, it may be useful to consider whether a particular point of view only arises in a litigation context and hence is not a customary point of view.

1.4.1 IMMINENCE OF OFF-SITE MIGRATION

Policies often apply only to third-party property. However, if there is an imminent threat to off-site locations, coverage may exist for on-site cleanup. In some states, groundwater under a site is off-site. To predict whether significant migration off-site is likely, soil leaching and soil erosion runoff models may be used. If the chemicals of concern are only slightly soluble in water and sorb appreciably to soils, then chemical transport through the vadose zone will be slow and concentrations reaching the water table may be below regulatory limits.

Groundwater transport models may be used to determine if a threat is imminent when groundwater contamination has not yet reached the property line and groundwater is considered off-site. Because of biodegradation in the plume as well as weathering and sequestration of mobile waste constituents in the source region, some plumes may reach a steady state before going off-site, or even recede over time. This is a common observation for BTEX (benzene, toluene, ethylbenzene, and xylene) plumes from gasoline spills (National Research Council, 2000).

1.4.2 TRIGGER OF COVERAGE

Some policies provide coverage only if in force when a claim is made. Coverage in other policies is triggered in environmental remediation cases by property damage. However, states differ in their determination of when damage actually occurs and if it can occur only once or can occur in a continuing fashion.

The possibility of triggering multiple policies in different time periods with multiple triggering events can lead to different interpretations of the same events. For example, a groundwater plume from a spill on one occasion may be stabilized and even shrinking, but since new water molecules are always entering the plume, some might argue that new damage is continually being done. Others, of course, would argue that the plume itself demarcates the extent of the damage.

Determining when policies are triggered often involves back-calculating a time of release or time to reach the water table as described in Chapter 8. In

some cases, structures such as cesspools, french drains, and leaching pits, which were specifically designed and installed to facilitate disposal of wastes to groundwater, negate the need for model calculations. A one-time liquid release, which is large enough to penetrate to groundwater, will generally do so over a period of hours or days. A cumulative or drip release will reach groundwater over a period determined by the drip rate. If the total quantity released is insufficient to reach the water table, the rate of contaminant travel will be controlled by the rate at which precipitation infiltrates the soil column and carries soluble waste components downward.

Reverse groundwater modeling, discussed in Chapter 11, can be used to determine the time when a property line was crossed or groundwater was first contaminated. However, there are always substantial uncertainties introduced by limited measurements in the subterranean environment. In addition, care must be used in defining the plume front; while the peak plume concentration may move with the retarded velocity, contamination in front of the peak moves more rapidly, up to and in theory even exceeding the groundwater velocity.

It may be possible to establish the time of release by linking the observed contamination to known process changes, such as a change in degreasing fluids from trichloroethylene to 1,1,1-trichloroethane (TCA). TCA releases can also be dated by the amount of the hydrolysis product, 1,1-dichloroethylene, present (Morrison and Murphy, 2006).

1.4.3 EXPECTED AND INTENDED

It generally will be important to determine if the damage was expected and intended. There can be issues that vary from state to state as to precisely what was expected and intended; that is, the release to the environment or the damage. Depending on the state, a "reasonable man" standard may apply or it may be necessary to produce evidence of actual knowledge by specific individuals. Of course what is reasonable for an individual to know depends on his or her background and role in an organization. Expectations are different for an accountant and an engineer, whose job might require him or her to read professional literature in that field.

Thus in some cases it will be useful to compare facility practices with historical waste disposal practices as evidenced by the engineering literature for the appropriate time period. The following illustrate the type of information that can be found.

- "The old fallacy of the speedy self-purification of streams was once pretty firmly fastened upon the engineering profession itself, and it is only in relatively recent

times that it has been wholly abandoned." Editor of Engineering News, Stream Pollution Fallacies, *Engineering News*, Vol. 42, No. 9, 1899.

- "The discharge of manufactural waste into streams without purification and treatment has frequently resulted in serious pollution. Manufacturers are coming to realize the seriousness of the conditions and consequently much study is being devoted to methods of rendering the wastes innocuous before their discharge into bodies of water." Disposal Methods for Manufactural Wastes, *Engineering Record*, August 27, 1910.
- "In the arid and semi-arid regions of the West, many large communities are virtually dependent upon groundwater supplies . . . Surveys show that refinery wastes in particular penetrate to considerable distances from sumps and stream beds." Burt Harmon, Contamination of Ground-Water Resources, *Civil Engineering*, June 1941.

As evidenced by these examples, the early pollution incident literature tends not to be chemical specific. It also is concerned with levels of contamination much higher than the levels that can be recorded with present measurement technology.

Pollution control legislation, practices at other companies in the same field, or trade association publications also may be introduced to illustrate the state of knowledge or practice at a given time. Generally, the engineering literature will provide a picture of a more advanced state of knowledge at an earlier time than these other references. Chapter 2 describes some sources of historical information.

If documentary information as to practices at a particular facility is lacking, it may still be possible to discern historical waste disposal practices from the spatial location or footprint of contamination at the site. For example, a groundwater plume emanating from a dry well could be linked to disposal of chemicals down a laboratory sink drain.

It may be important to distinguish contamination that arose from routine operational spills, which could be argued to be expected and intended, from such things as tank failures. Estimating the mass of contamination in soils and groundwater and characterizing the location relative to process areas can help in making such a distinction by determining the origin of contamination.

1.4.4 SUDDEN AND ACCIDENTAL

In the 1970s a clause was introduced to CGL policies that stated coverage for various kinds of releases would apply only if these were sudden and accidental. Some states consider sudden to have a temporal meaning and others consider

it to be more akin to unexpected. In the former case, it may be important to determine if a release was gradual or sudden in a temporal sense. However, even if the parties agree on the facts different interpretations can arise. For example, a leaking underground storage tank might be viewed as the result of years of electrochemical corrosion or it might be viewed in terms of a single instant when the tank is finally breached. Similarly, routine periodic degreaser cleaning and discharge to the environment might be characterized as a series of sudden releases or as a chronic operating condition.

1.4.5 EQUITABLE COST SHARING

Equitable cost sharing becomes an issue if there are multiple PRPs at a site. An unfavorable cost allocation scheme may be a basis to dispute full policy coverage. For example, as discussed in Section 1.2, when there are wastes that differ greatly in toxicity or mobility, a scheme based solely on the quantity of waste will be unfair to the disposer of large volumes of innocuous waste and its insurers.

Equitable cost sharing requires that waste streams be identified with specific PRPs. Methods are available for unmixing commingled waste streams. These include isotope techniques, discussed in Chapter 10, as well as principal components analysis (PCA) and polytopic vector analysis (PVA), discussed in Chapter 7. When indemnification costs are presented or settlements proposed, the question may arise, Should these techniques have been used?

1.5 TOXIC TORTS

In a toxic tort the issue is most often whether an injury was more likely than not caused by exposure to chemicals or other substances (e.g., radiological or biological). The causation requirement may also be phrased as "but for" the exposure the injury would not have occurred or that the exposure was a substantial contributing cause. Environmental forensics enters because historical chemical concentrations in air, water, soil, or foodstuffs are needed to estimate exposure and dose.

Dose is exposure times some uptake rate (e.g., cubic meters of air inhaled or an average number of grams of fish eaten per day). Exposure is determined by the concentration in environmental media (air or fish in the preceding examples), and by the time period over which uptake occurs. Since chemical concentrations may vary with time, exposure may be characterized over various time periods, acute or peak exposure, subchronic, or chronic (long-term) exposure. The averaging time of interest depends on the specific health effect being investigated.

Proof of causation for the injury may proceed in at least three ways: (1) if a sufficient number of people have been exposed epidemiological evidence may be offered, (2) a differential diagnosis may be performed for specific individuals, (3) although this is less frequent than the other two procedures, a risk assessment may be performed.

1.5.1 EPIDEMIOLOGY

Epidemiological information often is presented as an odds ratio for a specific type of injury. The odds ratio is the ratio of the number of observed cases in the putatively exposed community to the expected number of cases for a community of that size and demographic composition. Usually adjustments are made for age, smoking, ethnicity, and so on in determining the expected number of cases. It is often claimed that a probability of causation can be calculated directly from the odds ratio. If the odds ratio is OR, the probability of causation, P_c, is said to be

$$P_c = \frac{\mathrm{OR} - 1}{\mathrm{OR}} \tag{1.4}$$

P_c is thus just the fraction of total number of cases represented by the excess cases above background. When the OR > 2, then $P_c > 0.5$ and a cause other than background is sometimes said to be "more likely than not."

However, this equation is based on the assumption that background and source-specific causes act independently. In reality, disease manifestation may be a result of multiple causal factors, some of which are related to background and some of which are related to the specific source. The causal chain may be different for different individuals and the role played by background and the specific source may differ even for the same causal chain. In addition, as just presented there are three other things wrong with this argument and they show the role that historical exposure information, and hence forensic analysis, can play in assessing epidemiological evidence of causation: (1) association is being confused with causation; (2) no accounting is given of the number of different disease endpoints that were examined in order to find an OR > 2; and (3) in the preceding discussion the uncertainty associated with OR itself does not enter into determining P_c.

1.5.1.1 Association and Causation

Epidemiological evidence by itself describes association, not causation. To move from association to causation the Hill criteria formulated by Sir Austin Bradford Hill often are invoked (Hill, 1965).

These criteria are:

(1) *Strength of the statistical association.* As noted earlier, this is often measured by an odds ratio, comparing exposed and unexposed populations.
(2) *Consistency of the association.* Is the disease observed with similar exposures in other places and times? Do studies using a variety of techniques arrive at similar conclusions?
(3) *Specificity of the association.* Does the disease have many causes? Can the chemical in question cause many diseases?
(4) *Temporality.* Does exposure precede disease? Is the disease onset consistent with what is known concerning latency?
(5) *Biological gradient of the disease with exposure.* Are data for the population under study consistent with a dose–response relationship? Do the data show increasing rates of disease with increasing dose?
(6) *Plausibility.* Is a causal relationship between disease and exposure biologically plausible?
(7) *Coherence.* Is a causal interpretation consistent with other scientific understanding?
(8) *Experiment.* For example, if the suspected cause is removed, does the disease rate change?
(9) *Analogy.* Does experience with similar situations give any guidance?

The Hill criteria are intended as different viewpoints for examining causality rather than constituting a pass/fail exam. An additional criterion, which is sometimes added, is:

(10) *Elimination of confounders.*

Several of the Hill criteria are exposure related. The biological gradient criterion asks whether the number of cases increases with increasing exposure. The temporality criterion asks whether exposure preceded effect and if so whether it was by enough time to be consistent with what is known about disease latency. Consistency of the association is also exposure related. The injury may have been observed to occur elsewhere only when exposure was above some level. Both concentration in the exposure medium and averaging time enter into level of exposure.

1.5.1.2 Texas Sharpshooter Effect

The second thing that is wrong with the simple odds ratio/probability of causation argument is that how many end points were looked at is an important consideration in interpreting the results. If enough disease end points

are examined, an odds ratio greater than 2 may be found for some condition on a purely statistical basis. Restricting results to a 95% confidence level will not prevent this. At a 95% confidence level one test out of 20 will appear to be statistically significant, even if exposure conditions are actually identical in the two communities being compared. If a large number of disease end points are examined and only the high odds ratio and high confidence level cases are then presented, this constitutes what is sometimes called the Texas sharpshooter effect—where the bull's eye is drawn after the gun is fired!

1.5.1.3 Statistical Significance

The third thing that is wrong with an odds ratio greater than 2 simply equating to a causation probability greater than 50% has to do with the statistical uncertainty inherent in the odds ratio determination.

The basic concept is that even if we could do a series of identical studies with identical test populations and exposures, there would be a distribution of odds ratios because of statistical fluctuations. Let x be the odds ratio and $f(x)$ the distribution of odds ratios. Then the probability of causation is:

$$P_c = 1 - \int_0^\infty \frac{f(x)dx}{x} \tag{1.5}$$

Charrow and Bernstein (1994) show that Equation 1.5 implies:

$$P_c < 1 - \frac{1}{\int_0^\infty xf(x)dx} \tag{1.6}$$

for any distribution $f(x)$ subject to the normalization condition:

$$\int_0^\infty f(x)dx = 1 \tag{1.7}$$

If the odds ratio is identified with the expectation value of x:

$$\mathrm{OR} = \int_0^\infty xf(x)dx \tag{1.8}$$

then it follows that:

$$P_c < 1 - \frac{1}{\mathrm{OR}} \tag{1.9}$$

Thus just the fact that a distribution of odds ratios would occur if the epidemiological study could be repeated with different, but equivalent, populations causes the probability of causation to be overestimated by the simple expression $P_c = 1 - 1/\mathrm{OR}$.

1.5.2 DIFFERENTIAL DIAGNOSIS

The second way of determining causation is through a differential diagnosis. In a clinical setting this term means determining the underlying disease from among various possibilities through an analysis of symptoms. In a toxic tort context the term has taken on the meaning of determining the cause of a disease from among various possibilities. For example, if the claim is that a heart attack was chemically induced, among the factors that should be looked at as part of a differential diagnosis are the individual's weight, smoking habits, blood pressure, and age. This is in addition to whether the specific chemical is associated with heart disease and what the level of exposure was.

Causation is often considered in two parts. General causation addresses the question of whether the chemical in question is believed capable of causing the injury in question at any level of exposure. Specific causation addresses the question of whether the chemical caused the injury in the specific individual. The Hill criteria are used to support the general causation argument. Of course, for substances where the effect is well known, appeal to medical textbooks or government documents may be sufficient. Specific causation relies on the subset of the Hill criteria that may be applied to an individual rather than a population. These are the consistency of the association for the level of exposure, specificity of the individual's injury for the chemical, temporality (e.g., exposure preceding disease), experiment, or whether symptoms are alleviated when the supposed chemical cause is removed. In addition, as indicated earlier, confounding factors or other potential causes are eliminated as part of a differential diagnosis.

Historical exposure is thus an essential part of a correct differential diagnosis and can enter into a causation analysis in a quantitative way, most obviously, if we know the human exposure level at which disease is likely to occur. However, this condition is a rarity. More often we know the exposure level at which disease is not likely to occur. Thus one may compare the estimated historical exposure for an individual with chemical specific standards and criteria, both public health and occupational. Chemical exposure is unlikely to be a significant cause of disease if this exposure is less than these criteria. Of course, some caveats are necessary. Occupational criteria are generally less

stringent than public health criteria and there may be issues in comparing a healthy worker to the general population. Also criteria may not be based on carcinogenic effects, particularly for chemicals with limited evidence of carcinogenicity.

Similarly, one may investigate the significance of specific exposure levels by comparing the exposures or the concentrations involved with the exposures and concentrations of those same chemicals that people normally are exposed to through the natural or anthropogenic background, such as consumer products or urban air.

1.5.3 RISK ASSESSMENT

The basic risk assessment algorithm, risk = toxicity × dose, demonstrates the key role that dose and hence exposure plays. A risk assessment may be conducted at a hazardous waste site using EPA methods to calculate the lifetime risk of excess cancer; that is, of a cancer that would not have occurred otherwise. EPA's criterion for cleanup at a Superfund site is a computed risk, ΔR, larger than 10^{-4} to 10^{-6}. Computed risks rarely approach the more likely than not criterion of 0.50. This might be taken to imply that risk assessment is not a useful tool for demonstrating a causal toxic exposure. However, when there is a background rate for cancer of a certain type of R_0 and a computed chemical risk ΔR, the probability of causation for an individual who has cancer of that type can be written as:

$$P_c = \Delta R/(\Delta R + R_0) \tag{1.10}$$

Thus, even computed risks in the 10^{-4} to 10^{-6} range can be used to support a causation argument for an individual who already has cancer. Of course, as discussed in Section 1.5.1, this assumes a single cause for disease, and that background and other factors operate independently.

EPA risk assessment methods also can be used to calculate what is known as a *hazard index* for noncarcinogens. The hazard index is simply the computed dose over some averaging time, usually 24 hours, divided by a reference dose. The reference dose is a dose at which no adverse effects are believed to occur. Thus computing a hazard index <1 means that no adverse effect would have been expected from the exposure. Computing a hazard index >1 leaves the question of a chemically caused adverse effect open. In that case an informed judgment requires reviewing the primary medical literature including the occupational or animal studies upon which the reference dose is based.

1.6 NATURAL RESOURCE DAMAGE ASSESSMENT

In the United States, Natural Resource Damage claims can be brought by federal and state government agencies and Indian tribes (the "Trustees") under the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), the Oil Pollution Act (OPA), the Clean Water Act (CWA), as well as numerous state statutes. The Department of Interior and the National Oceanic and Atmospheric Administration have issued regulations indicating how assessments are to be done (http://www.epa.gov/superfund/programs/nrd/nrda2.htm).

Natural resource damage claims seek compensation to the public for damages resulting from a release of a hazardous substance or oil. The affected area can include virtually any aquatic or terrestrial environment and is not limited to living organisms or ecosystems. For example, many claims recently have been brought in the United States seeking damages to groundwater. Damages may be monetary or service-based, or a combination of the two, and are based on the value of "primary" restoration, compensatory restoration for lost natural resource services from the time of the release until such time as baseline services are restored, and reimbursement of assessment and response costs of the Trustees. For example, the compensatory damage component for the *Exxon Valdez* oil spill has cost Exxon close to $1 billion in addition to a cleanup bill ("primary" restoration cost) of nearly $2 billion.

Because of the varied nature of natural resource damages, the value of any Natural Resource Damage claim must be expressed in terms of some common currency. In some cases resources services can be monetized because there are markets for the services provided. Two such cases are: (1) diminished recreational fishing due to placement of fishing advisories, and (2) restriction of a groundwater source for drinking due exceedances of drinking water standards. The impact on fishing, for example, can be quantified by the travel cost to alternate fishing locations and/or lost fishing days. Similarly, the groundwater impacts can be quantified by the cost of providing an alternative water supply and/or the cost of treatment. More controversial methods include contingent valuation, a technique used to provide monetary values for goods, services, and public programs for which market data do not exist. The technique determines the value of goods and services based on the results of opinion surveys. In other cases, such as degradation of habitat and reduction of populations of biota, tools such as habitat equivalency analysis (Dunford *et al.*, 2004) can be used to determine the amount of restoration needed to offset lost services over time, thereby avoiding the need for a direct monetary metric.

Forensic issues, which frequently arise in Natural Resource Damage Assessments, whether conducted according to the preceding framework or just as a scientific matter, include: (1) establishing the baseline conditions that would exist absent the release, recognizing that the baseline itself may change over time; (2) determining the area of injury; (3) determining how much of the variation from the theoretical baseline may be attributable to natural variation; (4) determining causal relationships between injured resources and hazardous substances or petroleum compounds; and (5) sampling, laboratory, and statistical strategies to determine (1–4). An ecological issue that may arise is the significance of injury to individual organisms versus communities since a community of organisms may compensate for loss or injury to individual organisms so that the overall viability of the community is unaffected.

1.7 MARINE OIL POLLUTION

Marine petroleum pollution can result from tanker accidents; bilge water and ballast water discharges, or disposal of tank wash slops containing oil; runoff from land-based sources; and natural seeps. The petroleum can be in the form of crude oil or refined products. Generally the forensic task is to match the fingerprint of the suspected petroleum source to the petroleum found in the environment. Both the place of origin of the crude oil and the refining process contribute to a refined product fingerprint. The National Research Council estimated in 2002 that 1,300,000 tons of petroleum are spilled into the sea worldwide.

Prevention of operating and accidental discharges from vessels are controlled under the International Convention for the Prevention of Pollution from Ships, 1973, as modified by the Protocol of 1978 relating thereto commonly referred to as MARPOL 73/78 (http://www.imo.org/Conventions/contents.asp?doc_id=678&topic_id=258). In the United States the Oil Pollution Act of 1990 (OPA 90), administered by USEPA, provides resources to deal with oil spills including an Oil Spill Liability Trust Fund to address accidental spills. The Act also establishes requirements for contingency planning for both government and industry. Even when a source seems obvious because of a tanker accident or specific spill event, there is still a fingerprinting issue since in determining the extent of the impact, the background due to natural and other anthropogenic sources must be accounted for.

Petroleum biomarkers feature prominently in marine petroleum fingerprinting because of their strong dependence on the characteristics of petroleum formation or crude oil source(s) and because of their resistance to weathering. A review of this topic is provided by Wang *et al.* (2005a, 2005b).

ACKNOWLEDGMENTS

Some of this material appeared in different forms in *Environmental Claims Journal*, Vol. 5, No. 4, Summer 1993 and Vol. 8, No. 3, Spring 1996 © Aspen Law & Business, Inc. as well as in *Environmental Forensics*, Vol. 1, No. 3, September 2000 © Academic Press.

REFERENCES

Aronovsky, R.G. (2000) Liability theories in contaminated ground water litigation. University of Wisconsin, Environmental Litigation: Advanced Forensics and Legal Strategies, April 13–14.

Butler, J.C. III, Schneider, M.W., Hall, G.R., and Burton, M.E. (1993) Allocating superfund costs: Cleaning up the controversy. *Environmental Law Reporter* 23, 10133–10144.

Charrow, R. and Bernstein, D. (1994) *Scientific Evidence in the Courtroom: Admissibility and Statistical Significance after Daubert.* Washington Legal Foundation, Washington, DC.

Dunford, R.W., Ginn, T.C., and Desvousges, W.H. (2004) The use of habitat equivalency analysis in natural resource damage assessments. *Ecological Economics* 48(1), 49–70.

Ferrey, S. (1994) The new wave: Superfund allocation strategies and outcomes. BNA *Environmental Reporter* 25, 790–803.

Hall, R.M., Harris, R.H., and Reinsdorf, J.A. (1994) Superfund response cost allocations: The law, the science and the practice. *The Business Lawyer* 49, 1489–1540.

Hill, A.B. (1965) The environment and disease: Association or causation. *Proceedings of the Royal Society of Medicine* 58, 295–300.

Marryott, R.A., Sabadell, G.P., Ahlfeld, D.P., Harris, R.H., and Pinder, G.F. (2000) Allocating remedial costs at Superfund sites with commingled groundwater contaminant plumes. *Environmental Forensics* 1(1), 47–54.

Mink, F.L., Nash, D.E., and Coleman, J.C. II (1997) Superfund site contamination: Apportionment of liability. *Natural Resources and Environment* 12, 68–79.

Murphy, B. (1993) Technical issues in Superfund insurance litigation. *Environmental Claims Journal* 5, 573–592.

Murphy, B. (1996) Risk assessment as a liability allocation tool. *Environmental Claims Journal* 8, 129–144.

Murphy, B. (2000) Allocation by contribution to cost and risk at Superfund sites. *Journal of Environmental Forensics* 1(3), 117–120.

National Research Council. (2000) *Natural Attenuation Groundwater Remediation*, National Academy Press, Washington, DC.

National Research Council. (2002) *Oil in the sea III: Inputs, fates, and effects*, National Academy Press, Washington, DC.

Rockwood, L.L. and Harrison, J.L. (1993) The Alcan decisions: Causation through the back door. *Environmental Law Reporter* 25 ELR 10542.

Stewart, R.B., Clegg, R., and Goldsmith, B.J. (eds.). (1995) Natural resource damages: A legal, economic, and policy analysis. Nat. Legal Center for the Public Interest. Washington, DC, 208.

Wang, Z., Fingas, M., Yang, C., and Christensen, J.H. (2005a) Crude oil and refined product fingerprinting: Principles, Chapter 16 in *Environmental Forensics: A Contaminant Specific Guide*, Robert D. Morrison and Brian L. Murphy, eds. Academic Press, Burlington, MA.

Wang, Z. and Christensen, J.H. (2005b) Crude oil and refined product fingerprinting: Applications, Chapter 17 in *Environmental Forensics: A Contaminant Specific Guide*, Robert D. Morrison and Brian L. Murphy, eds. Academic Press, Burlington, MA.

Wise, K.T., Maniatis, M.A., and Koch, G.S. (1997) Allocating CERCLA liabilities: The applications and limitations of economics. *BNA Toxics Law Reporter* 11, 830–833.

CHAPTER 2

SITE HISTORY: THE FIRST TOOL OF THE ENVIRONMENTAL FORENSICS TEAM

Shelley Bookspan, A.J. Gravel, and Julie Corley

2.1 Introduction 24
2.2 Hypothetical Situation 1: Single-party sites with a history of preceding uses and users that may have contributed to the current problem 30
2.2.1 Task 1: Corporate Succession Research 30
2.2.2 Task 2: Facilities Research 32
2.2.3 Task 3: Ownership and Operations Research 32
2.2.3.1 Ownership History through Chain-of-Title and Tax Research 33
2.2.3.2 Operations Research 33
2.2.3.3 State Level Research 34
2.2.3.4 Local Level Research 35
2.2.4 Task 4: Work Product 35
2.3 Hypothetical Situation 2: A multiparty site, such as a regionally spread groundwater plume, where operations that resulted in chemical releases to the environment occurred *in situ* 36
2.3.1 Task 1: Past Occupants and Locations of Operations 37
2.3.2 Task 2: Target Entity List 38
2.3.3 Task 3: State and Local Agency Research 38
2.3.4 Task 4: Conduct Federal Archive Research 40
2.3.5 Task 5: Industry Standards Research 42
2.4 Hypothetical Situation 3: A multiparty site, such as a landfill, to which wastes generated off-site were transported and commingled over a period of time 43
2.4.1 Task 1: Landfill History 43
2.4.2 Task 2: Generator Identification 44
2.4.3 Task 3: Narrative Completion 45
2.5 Conclusion 45
References 46
Further Reading 46

2.1 INTRODUCTION

There is a polluted site; cleanup costs are in the millions of dollars. There have been many industrial occupants over a period of decades, but there are no obvious culpable ones. Under provisions of the federal Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA), and particularly as it was revised and reauthorized in the Superfund Amendments and Reauthorization Act of 1986 (SARA; 42 *US Code* Sections 9601, *et seq.*), current landowners who are not themselves accountable for all or any of the contamination found on their property can sue certain responsible parties for recovery of their cleanup costs. Such potentially responsible parties include those whose activities actually caused some or all of the contamination, no matter how long ago they occurred, or those parties' successors in interest (i.e., generators). Other potentially responsible parties are those who owned or operated the site when the contamination, or part of it, occurred (i.e., past owners or operators at the time of disposal). Others still are those who may have taken contaminating materials to the site for disposal, or who arranged for them to have been taken there (i.e., transporters or arrangers for disposal, respectively) (42 *US Code* Section 107).

CERCLA and analogous state level legislation, then, provide an array of prospective sources for financial assistance. Practically, however, how is it possible for the current landowners, faced with a cleanup bill, and their legal counsel to begin to unravel the necessary information about the site to ascertain whether there are other extant responsible parties? How is the context derived within which science-based environmental consultants can interpret the findings of field and laboratory studies to argue that a particular party's practices contributed to the environmental problem? A narration of the site's history may be what provides disputing parties, scientists, and adjudicators alike with clear and compelling information about the origin of the environmental problem. Certainly, it is not enough to find the pathways that chemicals have taken through the environmental media, nor enough to characterize the offending chemicals. In order to connect those findings to an existing party and to argue effectively that the party is responsible for paying some or all of the response costs incurred, the scientists and legal specialists also need to define the universe of past owners, occupants, generators, transporters, and/or arrangers. They need to be able to link the contaminants and the site conditions today somehow with the activities of the past. They need a narrative of the site's history to make sense of their findings.

Compared with the seemingly arcane formulae and models of science experts, the stuff of narrative history is ordinary documentation that most lay people can understand. The narrative history itself will become a source of

information for the entire environmental liability investigation. Because it is developed after the fact by a nonparticipant, this narrative is called a secondary source of historical information. A secondary source is a verbal reconstruction of events over a period of time that reflects the historian's best interpretation of the facts as discerned from the primary sources of information. The primary sources in the field of history are original documents, or documentary materials that were produced at the time of an occurrence. They may reside anywhere, in public agency files or in public or private collections. An archive is specifically a place housing original documentation, and there are archives around the country that maintain distinct types of collections. (For guidance to the country's archives see, for example, the National Historical Publications and Records Commission (1988) or try logging onto the Library of Congress's National Union Catalog of Manuscript Collections: http://www.loc.gov/coll/nucmc/. Or, for another site identifying and describing some 5000 plus repositories of primary documentation, see http://www.uidaho.edu/special-collections/Other.Repositories.html. In addition, a number of archives have begun to put portions of their collections online. See, for instance, the On-line Archive of California: http://www.oac.cdlib.org/ and the University of California History Digital Archives: http://sunsite.berkeley.edu/UCHistory/.)

Primary history sources can be verbal or graphic, and the ability to discern their meaning usually does not depend on a grasp of chemistry, engineering, hydrological, geological, or mathematical or other scientific jargon. This does not mean, however, that the work of history reconstruction is uncomplicated or lacking rigor. Ordinary documents—letters, testimony, photographs, maps, field reports—are the raw materials of a site history, indeed. Nonetheless, identifying them, locating them, and piecing them together in a way that coheres, assesses their relative merit, is true to their meaning, and suspends value judgment requires adherence to historical methodology. Like studies conducted using the scientific method, history conducted using the historical method must withstand tests that include peer scrutiny, reproducibility, and reliability (Bookspan and Corley, 2001).

Developing a work plan for a site history. How might a historian proceed to find and analyze documents relative to the history of a contaminated site for which responsible parties are being sought? Among historians who specialize in research and analysis for the purpose of clarifying the origin of specific circumstances of pollution, it is axiomatic that there is no one-size-fits-all formula for designing a site history work plan. A time line or chronology serves as the backbone of a narrative, and there are a handful of standard types of sources for which anyone compiling a site time line must look. Topographic maps and aerial photographs, for example, help define the period of development and thereby

help define the research periods. Sanborn fire insurance maps (block-by-block insurance maps of American cities published by the Sanborn Map Company), if they exist for the subject property, can supply a snapshot of the on-site land uses at a given point or points in time. Building permits, tax assessment records, city development maps, and city directory listings, if they exist for the subject site, may define some more of the bones of the land-use skeleton.

These are places to start, necessary but insufficient. A historian working a pollution liability claim or dispute needs to know much more about a site than simply what was on it, who was associated with it, and when, although these are vital data indeed. So vital, in fact, that such data can lead to the formation of a tailored work plan designed to unearth the idiosyncratic data on which a case may turn. For what gains the attention of the mediator, judge, or jury is not the listing in the city directory of a salvage yard. It is the 50-year-old transcribed testimony of a neighbor who told the Planning Department that he saw dozens of leaking drums buried on the site. It is not the aerial photograph that shows a sump, it is the letter of complaint about the rotten egg smell from that sump, found within the microfilmed files of the County Board of Supervisors. More than likely, historians will use disparate sources, some common, some obscure, to develop a sequential picture of what happened on the site. It is rare that a single source, or even single type of source, of historical information is complete, unambiguous, or unequivocal.

The methodology guiding historians through their planning for and conducting of site research, then, starts with the basic sources and proceeds through a warren of avenues. The work plan the historian develops at the outset of the research is effectively a road map identifying the main highways through a selected set of prospective sources, including the type of document sought, the information it is expected to provide, and the place or places where the document might reside and be found. The use of keywords to navigate through published indexes to libraries and to archival collections will aid in the construction of this initial work plan.

Internet searches designed to identify repositories of potentially relevant primary sources can also be of use, especially in this important first phase of the historical work. The scope and types of information and resources available on the Internet vary greatly. Some of the information is accurate, and a lot of it is not. The sheer amount of information has made it more difficult to determine the veracity of the data available. Nonetheless, certain searches can be particularly helpful when starting out on a project, for example:

- Manuscript collections pertaining to the companies and/or agencies of interest
- Environmental Protection Agency (EPA) databases for information regarding sites under agency scrutiny that are located in the vicinity of the target site

- Library and special collection databases
- Databases containing corporate information

A growing amount of information of the documentary type is available online. For instance, some city building departments have put their historical and current building permit collections online. However, in some instances only certain records have been cataloged for Internet use, and a trip to the agency will yield additional information. For thorough historical research the Internet is a vital tool and serves as a means to find places to look, but it is not the end point of the research project. It is more important than ever to reach back into the bowels of research facilities to obtain the documents necessary to reconstruct a thorough history. Many agencies, libraries, and archives simply do not have the resources necessary to develop and maintain extensive electronic catalogs. Even the most respected research facilities lack the funds and/or personnel to ensure that every collection or source is presented in electronic form either in its dedicated electronic catalog or on the Internet. The need to conduct in-person research in these facilities and examine all forms of indices has become more critical as large amounts of potentially relevant materials may reside in hard copy or on media such as microfiche and microfilm, which has fallen out of favor or been supplanted by newer imaging technology that may not be captured in online sources.

In addition, when relying on certain corporate or other collections, it is equally important to determine the types of records retained over time, the forms in which they were retained, and the most efficient way to extract potentially relevant data. For instance, it is not uncommon to require the use of sophisticated computer-based extraction tools to gain access to historical data. In some cases potentially relevant historical materials may have been stored on backup tapes or other computer-based media by companies attemting to preserve intellectual capital and storage space. Although the preservation methods may be considered rudimentary by today's standards, the various types of storage and preservation techniques employed can present unique challenges for data recovery and authentication. As a result, the historian must carefully word inquiries and be aware of a multitude of data recovery and extraction techniques and tools available.

The initial work plan, or research map, starts the historian on the treasure hunt, but is essentially an uncharted guide that requires revision as they actually explore those routes. They may find dead ends, or new or forgotten trails, back roads, or shortcuts. The value of the work plan at the outset will be commensurate with the historians' ability, through an understanding of change over time, to step away from current infrastructure, systems, and values of the prevailing society in which environmental discharges now occur. Files from

the 1920s, for example, will simply not contain the words "environment" or "ecology." That does not mean that there are no relevant files, however. They may be found perhaps by use of engineering nomenclature rather than "green" words, reflecting what were then the central values in that heyday of the civil and sanitary engineer.

Context, then, is key to a successful site history work plan. A work plan tailored to the site will reflect:

- The time periods of potential interest
- The jurisdictions in which it falls and into which it has fallen over time
- The types of chemicals currently found on or under the site
- The types of land uses or businesses occupying the site and when
- The remedial and legal issues at hand

Although there may be some similar elements, a work plan designed for each of, say, the following situations, will vary greatly from each other. Each of the situations contains clues for crafting the plan. Table 2.1 describes three distinct site history problems for pollution liability matters and extracts the clues to the explanatory context and consequently leads to useful primary source documents therefrom.

So it is that specific historical questions as well as the locations of the sites and the types of chemicals (and their containers, if any) causing the problems will suggest to the historian an era of origin and a range of sources. Review of the standard kinds of historical sources, such as the topographic or fire insurance maps, aerials, city directories, and building permits, will help confirm the initial impression and, perhaps, provide some of the other clues, such as the name of a former site occupant, on which to build that in-depth work plan. For more detailed listings of sources of site history information and repositories in which to find them, see Bookspan (1991).

Hypothetical cases and workplans. To provide further illustration of the interrelationship between research goals, historical context, and historical research methodology, provided next are three additional hypothetical situations in which the key question is who is responsible for a specific pollution problem. These hypotheticals represent each of the three basic situations in which potentially responsible parties (PRPs) may find themselves:

(1) They may be involved in a single-party site with a history of preceding uses and users that may have contributed to the current problem.

Table 2.1

Three examples of pollution liability site-history problems.

Research Objectives	Contextual Elements Suggested	Some of the Possible Primary Sources Suggested
To name contributors to an acid sludge pit in Louisiana	Refinery wastes reflective of Louisiana industry? Off-site generation? Multiple contributors? Possible association with aviation gasoline production and wartime era land disposal? Predating of specific state or federal environmental regulations?	Contemporaneous petroleum industry trade literature, such as *Oil and Gas Journal*. Parish courts for early pollution or nuisance complaints. Local history ephemera, as booster/Chamber of Commerce literature. World War II aviation gasoline program records in the National Archives.
To identify owners or the successors thereto of a defunct manufactured gas plant in Pennsylvania	Era of operation of the gaslight era, between 1880 and 1940? Utility selling gas service to the public in an era predating specific federal environmental regulations? Owners possibly benefiting from extensive acquisitions and mergers in the utility industry in early twentieth century? Possibly coming under state regulation for rates because of early Pennsylvania state involvement? Possibly making enough change through its operation in town to receive local attention?	County recorded grant deeds. State Public Utilities Commission ownership files and records. *Brown's* directories for utilities. *Moody's Industrial Securities* for acquisition and merger information. Securities and Exchange Commission filings for post-1930s, corporate tracings.
To identify parties contributing to an extensive trichloroethylene (TCE) plume in a California aquifer	Post-1930s and pre-1980s, reflecting the market for TCE versus carbon tetrachloride or trichloroethane? Multiple contributors? Metal working industries? Industries using heavy mechanical equipment? Cold war-related industries, as flourished in California? Local permitting for sewer use and special land uses beginning?	Various federal record groups within the National Archives for wartime, military-related industry records. Industrial waste disposal permits, conditional use permits, discharge regulations at the state and county level for 1950s onward.

(2) They may be involved in a multiparty site, such as a regionally spread groundwater plume, where operations that resulted in chemical releases to the environment occurred *in situ*.

(3) They may be involved in a multiparty site, such as a landfill, to which wastes generated off-site were transported and commingled over a period of time.

A detailed work plan follows each of these hypotheticals. Although neither the situations nor the work plans account for all the possible configurations of environmental events or available documentation, they are intended to

show how the case historian interweaves goals, context, and sources to reconstruct history.

2.2 HYPOTHETICAL SITUATION 1: Single-party sites with a history of preceding uses and users that may have contributed to the current problem

Problem. The client is a law firm representing a multinational chemical company (Chemicals, Inc.) that, among other things, owns several paint and dye manufacturing plants in the northeast of the United States. Ownership derived from the acquisition of two paint companies, one of which dates to the 1890s and the other to 1910. Some of the facilities remain in use, although, of course, they have been rebuilt and expanded many times over the years. Other facilities, however, are not functional, and others still have been sold to new operators.

As part of its program for minimizing its environmental exposure and for obtaining, where possible, third-party contributions for the costs of environmental remediation, the company has hired the client law firm to undertake a cost recovery program on its behalf. That law firm, in turn, has asked the historical consultant to assist in developing the requisite information by which to evaluate the cost recovery potential for the various sites.

The goals of the research are as follows.

(1) Insuring that all of the facilities that have been acquired over time are identified.
(2) Identifying and analyzing those facilities that have the potential for federal government cost recovery through wartime associated ownership and/or operatorship.
(3) Identifying and analyzing those facilities that have the potential for private third-party cost recovery through other theories, including prior or subsequent ownership and/or use; toll contracting; and off-site contributions.

Workplan. In the Chemicals, Inc. example it would be important to use a tasked approach that would allow the researchers to build on the research done in each task, ultimately to answer the questions posed in the research goals.

2.2.1 TASK 1: CORPORATE SUCCESSION RESEARCH

In this particular situation, prior to conducting what undoubtedly will be extensive research on ownership and operations issues, it is necessary to determine which company or companies to examine. To do this, it will be critical

to gain as much insight as possible from the attorney(s) representing Chemicals, Inc. and company personnel regarding existing common knowledge about the company's corporate lineage. In addition, the research will focus first on the main company and, subsequently, on the companies it acquired over time. A study of Chemicals, Inc. may well lead to research of Paint Company 1, Inc. (Paint 1), Paint Company 2, Inc. (Paint 2), and so on.

Primary sources. Because of its evidentiary value, the most desirable and useful information for the client will be obtained from primary, or documentary, sources. The three primary sources of corporate succession data for this case will include:

(1) Company-specific collections: Researchers should always begin by conducting research and speaking with company personnel to determine if company-specific collections exist either internally or in the public domain. Often, company archives can be a rich source of relevant information. In addition, many company-specific collections exist in the public domain. Reference material to such collections always should be consulted to identify whether such collections of relevant materials exist.

(2) Secretary of State filings: The next step in the corporate succession research is to conduct research in the appropriate Secretary of State office(s). Relevant information may be found in more than one state as companies sometimes reincorporate elsewhere over time for tax or other purposes. The types of information obtained should include:
   - Articles of incorporation
   - Amendments
   - Name changes
   - Mergers
   - Acquisitions
   - Dissolutions
   - Foreign corporation registrations
   - Certificates of good standing

(3) Security and Exchange Commission (SEC) filings: In addition to the Secretary of State, SEC filings can also be a rich source of corporate filings, if the company under scrutiny is or has been traded publicly. The purpose of federal securities laws, namely, The Securities Act of 1933, The Securities Exchange Act of 1934, and The Investment Act of 1940, was to require companies seeking to raise capital through the public offering of their securities, as well as companies whose securities are already publicly held, to disclose financial material and other information about their companies. Annual reports, 10K filings (forms filed annually by publicly traded companies with the SEC) and investment

prospects are the main vehicles for this communication and often contain very useful information for this type of research. This information can be obtained through a variety of sources including several service bureaus (e.g., Disclosure, Inc.) that make a business of providing this information.

Published sources. In addition to primary sources, there are a variety of published contemporaneous sources to be researched for the corporate succession phase of the project. Some of the most useful information can be found in business- and industry-specific directories, and often these can be found in libraries, particularly business libraries at universities or community research libraries. These sources tend to be published by research services, business services, or educational institutions and are compiled by researchers or through voluntary submission by the companies listed in the publications. In contrast, other secondary sources such as trade literature, business magazines, newspapers, and the like often contain articles or profiles of companies that are authored by a variety of individuals or groups. Published sources for the Chemicals, Inc. corporate succession research may include but not necessarily be limited to:

- *Moody's* Manual of Corporations
- *Standard & Poor's* Corporation Records
- *Ward's* Business Directory
- *Directory of Corporate Affiliations*
- *Million Dollar Directory*
- *Wall Street Journal* Index
- *New York Times* Index
- Chemical Manufacturers Association *Newsletter*
- *Chemical Week* Magazine

2.2.2 TASK 2: FACILITIES RESEARCH

Once the corporate succession work is done, the historians will compile two important pieces of information into one client work product: a detailed corporate history of Chemicals, Inc. and a list of facilities linked to Chemicals, Inc. through its predecessor companies. To compile the latter, many of the same sources, particularly company-specific records and directory information, will be used. This information will then be used to inform Task 3.

2.2.3 TASK 3: OWNERSHIP AND OPERATIONS RESEARCH

Taking the list of facilities compiled in Task 2, the historians can now focus on conducting ownership and operations history research.

2.2.3.1 Ownership History through Chain-of-Title and Tax Research

Chain-of-title research. In order to answer the question of ownership over time at each of the Chemicals, Inc. facilities, the researcher must determine the parcels of land related to each facility and conduct detailed chain-of-title research at the appropriate county, parish, city clerk's, or other office. In addition to researching the subject site, chain-of-title research to identify the owners of adjacent property may be conducted using the same time period criteria used in the Chemicals, Inc. research. This information may prove very useful in identifying off-site contributors to contamination of the subject property over time and in identifying parties from which Chemicals, Inc. can recover costs. In this phase of the work, it is critical to detail all the transactions related to the subject properties, to analyze any subdivisions of property, and to note any lease agreements and/or easements recorded in the deeds.

Tax research. In addition to deed research, it is often helpful to conduct research in the local tax collector's office. This office generally holds all information related to taxes paid, taxpayer identification, parcel information, and property liens and foreclosures. These data can be very useful in documenting the history of a property and confirming the identity of PRPs for cost recovery.

Once the ownership research is complete, the data should be compiled, analyzed, and recorded so they can be used to inform the operations history research.

2.2.3.2 Operations Research

Next, the researcher can focus on the operations portion of the project. To complete this phase of the work, the researcher or research team must now examine federal (if appropriate), state, and local and industry information sources to piece together a history of operations at the Chemicals, Inc. sites. Research sources may vary depending on the facility's location and other factors such as types of operations and period of operation. However, based on the Chemicals, Inc. example, the following types of sources should be helpful.

Federal data sources. Since there is a question about the federal government's involvement with the Chemicals, Inc. facilities, the research team will analyze the chain-of-title and other relevant documents to determine any government ownership or other types of control, particularly during times of war. If evidence of ownership or government involvement at an operational level were found or suggested, a more detailed review of federal records would be initiated.

Depending on the agency or branch of the military service involved, the researcher might examine records contained in:

(1) The National Archives and Records Administration (NARA) archival research facilities, which house records accessioned by the Archives, deemed to have potential historical value, and maintained within Record Groups signified by their federal agency or department of origin. (Online indices to the collections held within the regional branches of the National Archives and within the main research facility in College Park, Maryland, may be found, respectively, at the following sites: nara.gov/regional/findaids/findaids.html and nara.gov/nara/nail.html.)
(2) The Presidential Libraries.
(3) The Federal Records Centers, also operated by NARA, which house old files of federal agencies not accessioned to the Archives and still under the originating agency or department's control.
(4) Specific military collections such as the Air Force History Collection at Maxwell Air Force Base.

(See Hypothetical Situation 2, later, for further discussion of federal record sources.)

Site circumstances might suggest the utility of federal records other than ones relating to ownership or operations. For example, if issues of eminent domain because of the construction of an interstate were present, a review of Federal Highway Administration or U.S. Department of Transportation records could be appropriate. Likewise, if a navigable waterway or dredging were at issue, a review of the records of the U.S. Army Corps or Engineers or the U.S. Coast Guard could be of use. In the Chemicals, Inc. example it is likely that World War II records reviewed would include but not necessarily be limited to:

- Chemical Warfare Service
- War Production Board
- Defense Plant Corporation
- Office of Price Administration
- Ordnance Department

2.2.3.3 State Level Research

In the Chemicals, Inc. example, the state agencies targeted for research would be determined based on the specifics of the facility. However, generally the state level research would include but not necessarily be limited to agencies, whose exact names will vary based on the state, such as the:

- State Archives
- Attorney General's Office
- State Department of Health

- State Water Resources Board
- State Library
- State Department of Transportation
- State Lands Bureau
- State Historical Society
- State Department of Environmental Conservation
- State Department of Public Works

2.2.3.4 Local Level Research

In this area the same criteria for selection would apply, but it is likely that local level research would include but not necessarily be limited to agencies such as the:

- Department of Public Works
- Clerks Office
- Fire Department
- Health Department
- Library System
- County Offices
- County Court Records
- Historical Society
- Local specialty collections

2.2.4 TASK 4: WORK PRODUCT

Once all the information is collected, it is analyzed and compiled in a user-friendly format. At this point, the researcher, who has at least intellectually been sorting information along the way, must take the disparate pieces and weave together a coherent story about the company. In the Chemicals, Inc. example, the work product will have several distinct components including:

(1) Detailed ownership history of the real properties/facilities to be acquired: In the Chemicals, Inc. example it is likely that this portion of the work product would include a detailed ownership history for all of the properties associated with the acquisition. The ownership history would include a description of the lands over time and a detailed chronology of ownership past and present.

(2) Detailed information on operations at the various sites: Based on the information gathered at the federal, state, and local levels along with industry and company-specific operations information, an operations history would be developed weaving all the disparate pieces of information to tell a coherent story about the site. Some of the components of the site history might include manufacturing, building improvement, infrastructure improvement, waste

generation, and handling and disposal activities that have taken place at the site over time.

(3) An evaluation of the federal government's involvement at Chemicals, Inc. facilities: In the Chemicals, Inc. example it would be critical to assess the federal government's involvement in site ownership and operations as would be done for any other PRP. For instance, if records showed that the government owned the property and controlled operations during World War II, by virtue of its exclusion from sovereign immunity under the Superfund law, the federal government would be liable for contamination that occurred during its ownership period and subject to cost recovery efforts.

(4) An evaluation of the involvement of prior site owners and operators: Like the earlier example of the federal government, the researcher would also conduct a liability analysis of owners and operators of both the subject properties and adjacent lands to make a potential cost recovery assessment.

In sum, the historical investigation will result in a useful reference guide, complete with documentary evidence, for the future assessment of any of the facilities should they present with environmental liability problems. The information garnered will provide background regarding at least two of the major categories of parties liable under CERCLA: past owners who may have owned the property or facility when the release(s) occurred, and past operators whose activities may have resulted in hazardous materials being released to the environment.

2.3 HYPOTHETICAL SITUATION 2: A multiparty site, such as a regionally spread groundwater plume, where operations that resulted in chemical releases to the environment occurred *in situ*

Problem. In this case, the client is a law firm representing a small city in the southwest that owns a municipal airport that sits atop an extensive groundwater plume contaminated with residual volatile chlorinated hydrocarbon solvents (VOCs), and, in particular, trichloroethylene (TCE). The extent of the plume is not completely defined, but through a program of monitoring wells and sampling, it appears to cover an underground area at least six miles (9.5 km) long and two miles (3.2 km) wide. The airport, which succeeded the Army Air Force's World War II-era use of the facility, has been in operation since about 1954, and there are two repair hangars on site. Even so, there is no record of any use of TCE even in the degreasers there. Instead, the records, which date from about 1970, show that the solvent used in the degreasers has been 1,1,1-trichloroethane (TCA). There are some small industries, particu-

larly automotive related, that reside above the plume as well. Nonetheless, the EPA has assigned responsibility for the cleanup of the plume to the airport, as the current owner of what the agency believes to be the site of origin.

The potential costs for the cleanup of this plume are in the millions of dollars, and there is the added prospect of providing water to, and facing toxic tort liability from, users of a municipal well that has had to be taken out of service. The city has, therefore, hired a law firm to assist in its defense. The law firm, in turn, has hired a consultant specializing in historical research to help clarify the factual issues involving responsible, and potentially responsible, parties. Those issues in particular are:

- Ascertaining whether the airport or any of its lessees used and disposed of TCE at any time in the past
- Assessing the likelihood of a prior site occupant, such as the Army Air Force, having used and disposed of TCE while on the site

Workplan. The following discussion illustrates how the historical consultant would approach each of these problem areas to develop the informational basis analyzing the defense potential for the client.

2.3.1 TASK 1: PAST OCCUPANTS AND LOCATIONS OF OPERATIONS

A good place to begin is to determine who the past occupants were, where they operated on the site, and to identify the types of operations they conducted. These entities most likely will be the backbone of the target entity research list. One of the first steps taken to identify the target entities is to perform a chain-of-title search, documenting owners and recorded lessees of the property. Deeds and leases normally are filed at the county recorder's office and are searchable through grantee and grantor indices. For the purpose of this example, the results of the search showed the following information:

Owners	*Lessees*
Agricultural Company A, 1924–1941	
War Department, 1941–1946	
J. Smith, 1946–1953	Company B—auto shop, 1948–1953
	Company C—print shop, 1952–1955
	Company D—warehouse, 1952–1965
Municipal Airport Authority,	Company E—storage, 1956–1970
1953–present	Company F—sheet metal, 1956–1960
	Company G—auto parts, 1956–1995
	Company H—frozen pizza, 1965–1980
	Company I—dog groomer, 1972–1990

The legal descriptions found in the leases will help to identify where each entity operated at the site, which in turn will help determine their proximity to the plume. In some cases, the legal description of the area may not clearly encompass all the entities' activities, as operations may have occurred outside the area designated in the lease. The key is to remember that the chain-of-title search is only a starting point.

Not all leases are recorded, so it would be prudent to check other sources that identify site occupants. For this project, two good sources would be city directories and the airport annual reports. If available, both of these sources may provide some additional names and some context about the types of operations. For instance, a city directory listing from 1952 might include a brief description or an advertisement about the company's services. Airport annual reports may provide an indication about the magnitude of the operations or even discuss site conditions. Other sources that would be of particular use at this point in the research are any historic site tenant maps that the airport authority may have as well as fire insurance maps, which typically identify structures, major equipment, site occupants, and other useful information.

2.3.2 TASK 2: TARGET ENTITY LIST

The title, directory, annual reports, and map research will provide a target entity list of companies, with a documented association or connection to the site, to be used for further research. Review of the list may indicate a two-pronged research approach. On one hand, the target entities are companies that operated at the site in a relatively recent time period. On the other hand, there are industrial activities dating from the 1940s, when the military owned and operated the site. Sources of relevant information for the two categories in some cases may be the same, but those instances will be relatively few.

2.3.3 TASK 3: STATE AND LOCAL AGENCY RESEARCH

The next step is to implement a scope of work in which information about the chemicals used by site owners and occupants is assembled, as well as to document waste handling and disposal practices at the site. Certain federal, state, and local sources that would be consulted and the types of information that they typically contain include:

- Building permits: facility construction and major equipment installation.
- Fire department: inspections, notices of violations, and incident reports.

- Aerial photographs: site features and changes over time.
- Court records: nuisance/pollution complaints.
- Local historical society: contextual site and/or area history, and local business histories.
- Local public library: newspaper clipping collections, local business histories, contextual site and/or area history.
- County Department of Health, Environmental Section: inspections, investigations, notices of violation, releases and/or incidents.
- State Environmental Agency: inspections, investigations, notices of violation, enforcement activities, releases and/or incidents.
- Regional Air and Water Boards or Districts: inspections, investigations, notices of violation, enforcement activities, releases and/or incidents.
- State Archives: contextual site and/or area history, local business histories, state agency investigations.
- EPA: remediation files, investigations, inspections, permit records, compliance documentation, enforcement activities, and the like.

At this point the research results may seem to be disparate pieces of information. For instance, the researchers may have learned that Company B filed an application to install a paint spray booth with the building department in 1952. It is not clear if the permit was issued or spray booth installed. Court records informed the researchers that an employee sued the company in early 1953 for health problems associated with painting cars, but the case was apparently settled out of court and details about the painting activities and any solvents that may have been used in those activities were not in the file. Company B vacated the portion of the site that it occupied in late 1953 and the airport subsequently leased it to Company G, which began operating there in 1956.

Likewise, records from the fire department and the county environmental health department provide limited information about poor housekeeping practices that the two agencies observed at the operations of Companies C, D, F, and H between 1954 through 1969, but do not indicate the names of chemicals used at the site. The records found at the state archives indicate that the State Department of Health initiated a large study of a nearby river in 1947 that reported numerous pollution problems in the vicinity, but the report identified only operations that were active at the time of the study and did not examine past operations for their potential contribution.

The list of potentially relevant findings could easily continue. Researchers often prefer using different approaches to the questions at hand. Some prefer to identify the exact type of document they want to locate and then set out to

find it, whereas others prefer to see what they can find and then pull the various pieces of data together to tell a story about the operations at the site. Regardless of the method selected, it is essential to periodically examine the information obtained to identify gaps and leads to other relevant information.

2.3.4 TASK 4: CONDUCT FEDERAL ARCHIVE RESEARCH

At this point the most glaring gap is a lack of information about the activities of the military during the World War II era. The necessary research for the purpose of finding out about the daily operations of the Army Air Force is quite different than that described earlier. The records of the federal government are stored in various places, including the National Archives, Federal Record Centers, and with individual agencies.

The Federal Records Act of 1950 established a federal record management staff and a series of regional Federal Record Centers (FRCs) throughout the country. The FRCs work as a temporary holding center for the records of various federal agencies. Typically, an agency retains its records for a relatively short period of time (e.g., two to seven years), after which it transfers items with potential lasting value to the FRC. Items without lasting value usually are discarded by the agency at that time. Those materials transferred to the FRC are supposed to contain a notation regarding a schedule for their destruction or an indication that they are to be retained. Ownership of the records remains with the originating agency, and the FRC on average holds records for about 20 years. Prior to carrying out the ultimate disposition of the records—which can include return to the agency, transmittal to the National Archives, or destruction—the FRC usually contacts the agency to verify the action, or to request direction about schedules that were not correctly completed. Both the FRC and the originating agency normally hold documentation of the transfer of records in all aspects of the process. Records from the agencies are assigned a record group number and their materials, for the most part, retain that number at both the FRC and the National Archives.

For this project, record groups (RGs) of interest in this matter might include:

- RG 18, Army Air Forces
- RG 51, Office of Management and Budget
- RG 160, Records of the Air Service Forces
- RG 179, Records of the War Production Board

- RG 219, Records of the Office of Defense Transportation
- RG 270, War Assets Administration
- RG 340, Air Force, Office of the Secretary
- RG 341, Air Force, Headquarters
- RG 342, Air Force, Commands

Another key repository for the task at hand is the Air Force Historical Research Agency (AFHRA), which holds, among other things, the records of Army Air Force facilities. Military units were required to prepare reports detailing their unit or section's activities on a regular basis. The reports usually were submitted to a person designated as the base historian, who in turn compiled the information from the various units into a comprehensive history of the base, camp, post, depot, or whatever. The comprehensive histories usually were compiled semiannually, quarterly, or even monthly.

The level of detail found in the reports varies tremendously and reflects information that the soldiers who prepared the reports found of interest. Although such reports may not contain information explicitly about TCE use, they may provide the most detailed information available about the military operations at the site. In addition, the unit histories also will contain the names of personnel stationed at the base. The individuals who worked at the base often can provide first-hand recollections that could provide pivotal or persuasive information about chemical use, waste handling operations, disposal practices, and the like. In the absence of direct documentary evidence of TCE use at the site, the information supplied by the soldiers can provide convincing testimony. World War II veterans, however, are by now quite elderly, and more are lost to us daily.

Once the research at the National Archives, Federal Record Centers, and AFHRA is complete, the historians most likely will be faced with additional disparate information. The research into the military's operations will have yielded, at a minimum, the types of activities that occurred at the base, such as aircraft maintenance, motor vehicle maintenance, gun maintenance, equipment renovation, railroad spur activity, dry cleaning and laundry facilities at the base, and the like. In an ideal world, they would have found explicit documentation that several of the site occupants used TCE and that they dumped, discarded, and disposed of TCE-containing wastes in an unacceptable manner (e.g., poured down drains, used it for weed or dust control, dumped it in a pit that was located in the middle of what is now the highest hot spot at the site, etc.). If these were, in fact, the findings, and if the entities or their successors were active and financially viable, the historians could compile the

information, prepare a report, and submit it to the EPA, which would most likely name the additional entities as PRPs. However, it is not the end of the trail if no direct documentary evidence of TCE use is found.

2.3.5 TASK 5: INDUSTRY STANDARDS RESEARCH

The next step would be to identify standard industry practices regarding chemical use and waste handling and disposal practices. Trade literature as well as military technical orders, manuals, directives, and the like can be used to help make connections—albeit inferred—between some of the activities at the site and the TCE contamination. The additional effort of locating and interviewing people who worked at the site as to their first-hand knowledge of solvent use there could provide further persuasive information. For instance, a soldier might recall that he used a special solvent to clean engines right before inspections so that the equipment would pass a particularly grueling white glove test. Another might recall using a vapor degreaser at the site in which only Tri or Triad could be used. The former soldier might not know that Tri and Triad were brand names for TCE, but that particular information can be documented elsewhere. Another worker might remember mixing a solvent concoction to produce a cleaner that removed carbon from engine parts more effectively.

If direct evidence of TCE use is not found, it would be worthwhile to define further the entities for which additional research is required. To do this evaluation, the historians should consider the types of operations and next conduct a preliminary corporate succession tracing to locate likely viable parties. For instance, it is unlikely that the dog groomer used TCE and also unlikely that the company—if it is still in business—has substantial assets. The auto-related lessees at the site, however, may be worth consideration because TCE was a solvent of choice for metal cleaning during the years in which the companies operated at the site. The agricultural company also may be worth further examination because TCE and carbon tetrachloride were used as grain fumigants prior to World War II. The U.S. Government is definitely worth inclusion in any further research because the military was the largest user of TCE during World War II.

In sum, in order to address both of the research goals, the historians would develop a twofold plan: one in which they investigate possible TCE users among the private industry occupants of the area above the plume, and one to investigate the military's possible use and disposal of TCE during its wartime tenure. They would use the information gleaned from one source to enhance their ability to interpret another, rather than discard as useless documents that are not stand-alone or smoking guns. In the end, the historians' ability

to construct a convincing narrative from the unearthed information will depend on the interplay among secondary, primary, and oral sources.

2.4 HYPOTHETICAL SITUATION 3: A multiparty site, such as a landfill, to which wastes generated off-site were transported and commingled over a period of time

Problem. In this situation, the client is a municipality that owned and operated a landfill, and only one landfill, for its citizens for a period of more than 40 years, but that landfill ceased operations in 1978, a time period effectively predating waste manifest regulations. Now, chemicals have been found leaching from the landfill and migrating through the geological substrata to threaten the aquifers from which the municipal water district draws its supplies. The client city is facing liability not only for cleaning up the landfill and the aquifer, but also for providing alternate water sources to the affected water district clients, and, possibly, for toxic tort claims based on the citizens' suits alleging harm from consuming contaminated drinking water.

The city's attorneys have hired historical investigators to assist in developing evidence relating to two categories of potentially responsible parties under CERCLA: the generators of the contaminating wastes brought to the site over time and the haulers of those wastes.

Workplan. As the city's own consultants, the historians have unusually direct access to the files and records of the city, and need not await the results of a clerk's file search initiated on receipt of a formal letter request.

2.4.1 TASK 1: LANDFILL HISTORY

In order to narrow the list of potentially responsible parties in such a complex site as a landfill serving innumerable parties over a long period of time, the first step is to set the parameters of the research. Examination of the city's Public Works Department, accounting, and City Council records, among others, should provide answers to some basic historical questions set out here.

When did the landfill open and was it run continuously until it closed? This will help define the period(s) of time relevant to investigate further and it will reveal the regulatory era(s) in which the landfill operated, so as to suggest which current or historic agencies' records, if found, may prove valuable.

Did the municipality own or lease the property? This will identify any co-owners who may be joined into the landfill action, and, if there was a change

in ownership or a renewal of leases or the like along the way, it may signal a time when discussion about the landfill occurred in city administration or council. Successful research often depends on finding occasional, rather than regular documentary evidence, and pinpointing decision times can ease the search for such records.

How did the landfill operate? Did the municipality operate it through its own staff, or contract its operations to a private management company? How was income generated? Were customers billed, and, if so, what municipal agency did the billing? Were receipts retained? Answers to these questions will refine the researchers' search for specific types of prospectively helpful documents.

In this case, the landfill history reveals that this single landfill, when in operation, served all of the municipality. Moreover, the city operated its own refuse pickup and disposal service, paid for by citizen taxes, and disposed of all the wastes picked up in the subject landfill. This seems to exclude finding anyone other than the city to qualify as a PRP by virtue of having been a hauler. It also suggests that anyone who ever lived in or operated a business in the city is a prospective PRP by virtue of having been a generator. Nonetheless, certain generators are sure to have contributed more chemical wastes to the landfill than others, so the next step is to identify who they may have been.

2.4.2 TASK 2: GENERATOR IDENTIFICATION

Primary source research. Because all the occupants of the city during the time the landfill operated were generators, that is, their wastes all went there, the researchers will inventory the city through screening types of sources (such as city directories or Chamber of Commerce directories and Sanborn fire insurance maps) and, next, set criteria for priority investigation targets. So, perhaps, the researchers will use a series of city maps covering every five years of the landfill's life. Then, they may create a database of businesses listed in the city directories by their street addresses. Next, they may select some of those businesses by their business type, such as dry cleaner or metal shop or automotive repair, if such type generally has been associated historically with chemical usage and waste discharge, and chart those businesses by their locations on the series of maps. These businesses will represent the prospective PRPs, and there may be among them some entity, an aircraft parts manufacturer, for example, that not only particularly suggests chemical waste generation, but also suggests other document leads, such as regulatory or federal agency files. Even so, the nexus between the existence of the PRP in the city and its wastes being in the landfill has yet to be made directly.

Witness interviews. Barring the finding of some highly suggestive PRPs, the utility of documents to identify generators may diminish until there is another breakthrough to follow. It may be useful, then, to locate individuals with first-hand knowledge about the wastes generated by the businesses identified and/or first-hand knowledge about the disposal of those wastes into the landfill. Given that in this case, the city hired its own employees to conduct the business of refuse pickup and disposal, and given that the historians on the matter are consultants to the city, the personnel files may be available for access and review. In them are identified individuals who worked for the Refuse Division of the Public Works Department at the time the landfill operated, and, if appropriate truck drivers or other refuse workers, such as *swampers*, those who picked up the cans and rode the back of the truck, can be found, they may provide testimonial evidence about the wastes they picked up from the target entities. Sometimes, a ride with such individuals through the city, following the route of their work, regardless of how altered the streetscape now is, can prompt memories that can flesh out the survey documentation in a convincing way.

2.4.3 TASK 3: NARRATIVE COMPLETION

Through the landfill history and the documentary and interview work undertaken to identify generators, the researchers will have compiled disparate data that will need to be organized and accounted for in order to provide the clients with the PRP evidence they requested at the outset. A narrative that explains the work done, the sources consulted, the rationale for the PRP selection, a summary of the evidence regarding each PRP, and appends copies of all cited documentation or interviews, is the final work product of this level of investigation.

2.5 CONCLUSION

One of the operative terms in the conduct of site histories for environmental forensics purposes is *disparate.* The types of sources available will depend on the city, county, and state in which the site is located, the years of concern, the environmental issues of concern, the liability issues, and, of course, the researchers' ability to navigate among these issues in the sea of documentation. The examples here are intended to illustrate these interactions. Equally, they are intended to dispel the notion that a smoking gun document is needed to provide good evidence in a PRP case. After all, in environmental cases especially, the historical information will provide only an important piece, even if it is sometimes the cornerstone, of the causation

and resposibility puzzle. Although the historian may find that the subject property was home to a dry cleaning operation during the 1920s, and that the now-defunct operation used carbon tetrachloride in its process, it will remain for members of the scientific team to assess the relevance of that information in relation to the overall contamination problem. If the site is contaminated with chromium, then the information about carbon tetrachloride, a chlorinated hydrocarbon, may be interesting, but not related to cleanup costs. On the other hand, the historical information may lead the scientists to conduct additional tests of, say, the place on the site that formerly held the dry cleaning equipment and, perhaps, add tests for chlorinated solvent analytes to the tests for metals currently being taken from the existing sampling locations. This interaction and the iterative nature of the historical and scientific investigation processes is in many cases key to the success of both efforts. The historical and scientifc investigations should never be conducted "in a vacuum" as the ongoing evaluation of evidence may lead to additional questions that need to be answered. Overall, the historical story will provide the scientists with a way to interpret their findings relative to a past user's on-site activities and to build the core of the environmental forensics argument.

REFERENCES

Bookspan, S. (1991) History requited: Historians and toxic waste. In *History and Public Policy*, D. Mock, ed. Krieger Publishing, Malabar, FL.

Bookspan, S. and Corley, J. (2001) The paper trail: How and why to follow it. In *Practical Environmental Forensics: Process and Case Histories*, P.J. Sullivan, F.J. Agardy, and R.K. Traub, eds. John Wiley & Sons, New York.

National Historical Publications and Records Commission (1988) *Directory of Archives and Manuscript Repositories.* Oryx Press, Phoenix, AZ.

FURTHER READING

Barnett, H.C. (1994) *Toxic Debts and the Superfund Dilemma*. University of North Carolina Press, Chapel Hill, NC. [A study of the Superfund program and its legacy elements.]

Brown, M.H. (1979) *Laying Waste: The Poisoning of America by Toxic Chemicals*. Pantheon Books, New York. [This book, together with those by Tarr (1996) and Colten and Skinner (1996), provides background regarding the history of chemical waste disposal and American industry.]

Colten, C.E. and Skinner, P.N. (1996) *The Road to Love Canal: Managing Industrial Waste before EPA*. University of Texas Press, Austin, TX.

Tarr, J.A. (1996) *The Search for the Ultimate Sink: Urban Pollution in Historical Perspective*. University of Akron Press, Akron, OH.

Winks, R.W. (1969) *The Historian As Detective: Essays on Evidence*. Harper & Row, New York. [A classic discussion of methods of historical investigation and, particularly, forensic use of historical documentation.]

CHAPTER 3

PHOTOGRAMMETRY, PHOTOINTERPRETATION, AND DIGITAL IMAGING AND MAPPING IN ENVIRONMENTAL FORENSICS

James I. Ebert

3.1 The Aerial Photographic Record 49
3.2 Principles of Photogrammetry 51
3.2.1 Photographic Scale 51
3.2.2 Radial Displacement 52
3.2.3 Stereo Photographs and Stereo Viewing 52
3.2.4 Resolution and Photographic Media 53
3.2.5 Digital Image Processing 54
3.3 Photointerpretation 58
3.3.1 Photointerpretation versus Photo Reading 58
3.3.2 Photointerpretation in Environmental Forensics 60
3.4 Analytical Products and Preparation of Exhibits 64
3.4.1 Maps 64
3.4.2 Geographic Information Systems 66
3.4.3 Terrain Visualization 67
3.5 Case Studies 67
3.5.1 Case Study: Informal Dumping Area Covered by Plant Addition 67
3.5.2 Case Study: Product Transfer and Storage Spillage at Circuit Board Fabrication Plants 69
3.5.3 Case Study: Chemical Drainage from Treated Wood at a Creosote Plant 72
3.5.4 Case Study: Dark Staining at a Maintenance Yard 72
3.6 Conclusions 78
References 79

3.1 THE AERIAL PHOTOGRAPHIC RECORD

Aerial photographs comprise a huge international archive of hundreds of millions of images, taken over the period of slightly more than a century, that document the earth's natural landscape and humankind's interaction with it.

As soon as the process of chemical photography was developed, photographers began experimenting with taking aerial photos, first from balloons and then from aircraft. The collection of aerial photographs systematically covering large contiguous areas began in the United States in the early 1930s by government agencies, largely for purposes of making topographic, soil, agricultural, intelligence, and other sorts of maps and photointerpretations.

The maps made by the United States Geological Survey (USGS), Department of Agriculture, National Aeronautics and Space Administration (NASA), the U.S. Army, and other agencies from systematic aerial photographs through time are problem oriented, and each map-making or interpretive effort focused on only a narrow range of the vast amount of information that can be extracted from those aerial images. Although they are conserved and made available in many ways, the aerial photographs remain in government and private archives for others to use in the context of their own problem orientations, background, and expertise. Regardless of colloquialisms that suggest that aerial or any other sort of photographs speak for themselves, they do not. In order to get answers from aerial photographs, questions must be asked of them. And there are as wide a range of kinds of questions to be asked, and ways of asking them, as there are scientific, technical, and other disciplines.

Some of the earliest applications of aerial photographs were by the military, with balloons being used to photograph enemy lines by Union forces during the Civil War. Another early use of balloon platform photography was the imaging of Stonehenge in the United Kingdom in 1906 by archaeologists, who have been pioneers in aerial photographic methods. Geographers base much of their science on information—much of it represented in maps—derived from aerial photos. Aerial photographs also serve as basic data sources for geologists, foresters, range managers, agricultural scientists, hydrologists, biologists, and environmental scientists. Engineers and architects use aerial photographs to interpret, map, and measure information important in their work such as topography, soils, and the locations of the built environment. Many other creators and users of aerial photographs do so for historical or artistic purposes, and a search of the aerial photographic literature turns up as many hobbyist or artistic titles as scientific or engineering ones.

Photogrammetrists are scientists and experts specializing in the development and application of methods specifically focused on using aerial photographs and other sorts of imaged data, as well as many associated digital imaging and mapping methods. Photogrammetry is defined by the American Society for Photogrammetry and Remote Sensing (ASPRS) as "the art, science and technology of obtaining reliable information about physical objects and the environment, through processes of recording, measuring, and interpret-

ing imagery and digital representations of energy patterns derived from noncontact sensor systems" (Colwell, 1997).

Photogrammetry as a practice originated before the advent of aerial photographs, but as a profession emerged by about 1930 when hundreds of photogrammetric engineering firms became established in the United States and around the world, taking photos with sophisticated aerial cameras and using optical-mechanical plotters to make contour maps. Over time these firms and government agencies have blanketed the landscape with aerial photographs. Since the beginning of systematic coverage of the United States, in most places from the early to late 1930s, it is not uncommon to find up to 30 to 40 different dates of aerial photographs from then to the present, and in urban areas there can be hundreds.

The shift to digital data and techniques, beginning in the mid-1970s and accelerating through the next two decades, profoundly affected the profession and practice of photogrammetry. In November 1985 the American Society of Photogrammetry changed its name to the American Society for Photogrammetry and Remote Sensing in recognition of the fact that the profession encompasses not only interpreting and measuring from aerial photos, but using space pictures, other digital imagery, digital imaging techniques, and computer mapping technologies including geographic information systems (GIS), computer aided drafting (CAD), and global positioning systems (GPS), which are parts of the field of imaging and mapping today.

3.2 PRINCIPLES OF PHOTOGRAMMETRY

An understanding of the principles of photogrammetry can assist any scientist, expert, or other aerial photo user to better understand and interpret the information contained in aerial photographs and other sorts of images. The most basic properties of photographic images derive from the fact that photographs are the accommodation of three-dimensional, real-world scenes to a two-dimensional medium.

3.2.1 PHOTOGRAPHIC SCALE

The scale of objects in an image (i.e., the ratio of their size in the image to their actual size) is a function of camera geometry and the distance of objects from the camera. For aerial photographs, this relationship is:

$$\text{scale} = \text{focal length}/\text{flight height} \tag{3.1}$$

If, for instance, an aerial photo is taken at an altitude of 12,000 feet with a six-inch lens (a common focal length for aerial cameras), then the scale of

the photo is 1:24,000. Engineers sometimes express this in terms of inches to feet; in this example 1″ = 2000′.

Scales of about 1:20,000 to 1:30,000 are commonly found in historic aerial photographs taken for mapping purposes until about the 1980s, when improvements in films and plotting devices made it practical to use somewhat smaller scale aerial photos and the USGS began using scales between 1:40,000 and 1:80,000 for producing topographic maps. The scale of aerial photos has a profound impact on what can be seen in them; at 1:24,000 scale, for instance, a 50-gallon drum measures only about a thousandth of an inch in diameter on film or contact prints made from it. Quite clearly, optical or digital magnification of an image is required to recognize such objects in an aerial photo of this scale.

The scales given for aerial photographs by those who took them are nominal scales calculated by using the average height of the aircraft above the ground. Actually, the exact scales of objects with an aerial photo are a bit more complicated than that, varying with the distance of each specific object from the camera. For instance, an object on top of a hill is closer to the camera than the same-sized object in a valley, so their scales vary. This is not much of a problem in relatively flat terrain, or with small-scale aerial photos, but can result in significant variation in the sizes of objects in large scale (low altitude) photos taken over rugged terrain.

3.2.2 RADIAL DISPLACEMENT

The heights of objects has another effect on how they are represented in an aerial photo, which photogrammetrists refer to as radial displacement. Every object in a vertical aerial photograph, except at its exact center, appears to lean outward from the photo's principal point, to a degree that increases with its height and the distance it falls from the photo center. This is why one sees the sides of some objects, for instance buildings, in an aerial photograph and why the exact placement of objects with height differs in two overlapping photos. Radial displacement can be thought of as a photographic distortion, but actually is a property of photographic geometry, and in fact provides the information that allows one to measure from, and see, a three-dimensional view of the landscape, in stereo photographs.

3.2.3 STEREO PHOTOGRAPHS AND STEREO VIEWING

Systematic aerial photo coverage typically is flown in straight flight lines with 60% overlap between consecutive frames, ensuring that everything on the ground can be seen in at least two of the photos. When adjacent photos—a

stereo pair—are viewed with a stereoscope, a three-dimensional view of the ground and objects in the photo is seen. Stereoscopes range in sophistication from pocket models costing $20 to precision instruments offering high-power zoom magnification costing tens of thousands of dollars. Perhaps the most efficient type of stereoscope for serious use is the mirror stereoscope, which can be fitted with a parallel-tracking mount, allowing the photointerpreter to scan across the stereo image without having to constantly reposition the prints. Binocular eyepieces available for mirror stereoscopes typically offer about 2× to 6× magnification. If higher magnifications are necessary to see very small details in small-scale photos, more expensive instruments are required; some zoom stereoscopes can magnify images as much as 200×. The light-efficiency of such optics quickly decreases as magnification is raised, however, and extremely intense illumination is required. This and the fact that it is hard to find tiny features on photos at high magnification can make using such stereoscopes difficult.

Stereoscopic viewing is always very helpful, and in most cases absolutely necessary, in aerial photointerpretation. Identifying many sorts of features at environmentally relevant sites—for instance drums, structures, piles, pits and trenches, which can be nearly impossible with a monoscopic photo—becomes much easier with stereo. Topographic details, which are of vital importance in many environmental situations, cannot be appreciated without stereoscopic viewing.

3.2.4 RESOLUTION AND PHOTOGRAPHIC MEDIA

There is some debate in the professional photointerpretive community regarding the best photographic medium to obtain from those firms or agencies holding the original photo negatives. Aerial photos are exposed on film, and the original film is referred to as the first generation by photogrammetrists. The original film is archived at firms and agencies that make second-generation copies of aerial photographs for users. Black-and-white negatives are the first generation for almost all black-and-white photo products; color aerial photos can be taken with color negative or color positive transparency film. Second-generation products can be made from negatives as either transparencies or paper prints; second-generation positive transparencies or prints can be made from positive, first-generation film by using reversal film or paper (such as Cibachrome). Positive paper or film products are sometimes also made from first-generation, positive film by making an internegative, in which case the positive copies are third-generation products.

Some photogrammetrists assert that one should obtain positive film transparencies as the basic data for environmental photointerpretation, stating

that what one sees in a film transparency is far superior in resolution than a paper print. Although it may be true that paper prints sometimes provide less spatial resolution than film duplicates, it is often difficult to tell the difference even when one is inspecting extremely small objects or features in the film. The resolution and contrast performance of duplicating films and papers vary widely, as well (Teng, 1997). The main advantage of a photographic film product over a paper print for most uses is the wider exposure latitude and spectral resolution of film, especially black-and-white panchromatic films. In first-generation panchromatic film negatives, there are rarely any portions of an exposed frame that are totally black or totally white; some level of gray-scale detail appears almost everywhere. Each successive generation of photo products—film or paper—introduces increased contrast and less spectral resolution. If details in areas that appear very dark or bright are important in a specific photo-interpretive effort, details in those areas can be maximized by under- and overexposure. A disadvantage of using film products for photointerpretation is that illumination must be provided by a light table.

A better strategy for evaluating potential differences in resolution between film and paper second-generation aerial photo products is to use enlargements made from original negatives for interpreting details that are small in small-scale aerial photographs. If identical areas are enlarged from the two frames making up a stereo pair, these can be used with a stereoscope just like contact prints. Enlargements can be produced digitally, but even with the high-resolution scanners used by photogrammetric engineering firms (typically about 12.5 microns, or 2032 dpi), tiny details scanned from 10″ aerial film—for instance drums that measure 1/1000″ on 1:24,000 scale photos—are practically unresolvable. If such details are important, obtaining photographic enlargements and then scanning those to produce digital image files allows comparable or better results to be obtained with much more affordable scanners.

3.2.5 DIGITAL IMAGE PROCESSING

Digital imaging and image processing can be advantageous in a number of ways in environmental forensic photointerpretation. When these methods were first becoming widely available in the mid to late 1980s, there was considerable discussion concerning whether they could be used to alter or fabricate information, thereby making them problematic in legal applications. Such concerns have become less frantically voiced as scientists and the public in general become more familiar with digital imaging. Nonetheless, it is true that digital imaging can be used to alter the appearance of images. Therefore,

it is necessary when using digital imaging techniques in a legal context to fully understand digital imaging processes and their effects on images.

Of course, traditional photographic techniques alter images as well—two photo prints from the same negative can look quite different. It has always been incumbent upon those using photographs for forensic purposes to exercise ethical behavior and thought processes when interpreting the meaning of images. An example of a common photointerpretive pitfall is the attribution of meaning to dark staining of surfaces, the appearance of which can be changed or in fact created by many factors, for instance varying contrast caused by photographic or digital processing, or variation among lighting, vegetation, and other factors in aerial photos taken on different dates. This theme will feature prominently in a case study to be presented later in this chapter.

One of the most straightforward and immediately valuable ways digital imaging can aid in environmental photointerpretation is as a means of enlarging small details from aerial photos. As discussed earlier, obtaining photographic enlargements from the original data (i.e., the film negatives) is often a useful first stage in examining important details. The enlargements can then be scanned using relatively affordable flatbed scanners to create digital images.

One way to ensure that interpretations are not biased by differences introduced by digital imaging processes is to work interactively with both photographic products and the digital images made from them. A useful practice is to view the stereo photographs with an optical stereoscope, while zooming in and out to see specific details on a computer screen. Although optical stereoscopes are still the best way to view stereo, they have some disadvantages, including that only one person can view the stereo image at a time. Many people can view an image together on a computer screen, and this can often be quite advantageous in cooperative photointerpretation among the photogrammetrist and other experts in a case, and of course for showing others what the photogrammetrist is seeing.

Stereo can be viewed on a computer screen, and although shuttered or polarized glasses can be used to do so with special software, a quite workable and affordable alternative is anaglyphic viewing, in which one frame of a stereo pair is assigned to the red band, and the other frame to the blue band, of a color (RGB: red-green-blue) image. The image is then viewed with red/blue glasses, which are really just colored filters in a frame and can be purchased from a number of sources for from less than a dollar for cardboard models, to $20 for comfortable plastic frames. Anaglyphic stereo viewing for extended periods takes a little getting used to, and what one sees is qualitatively somewhat different than viewing through a stereoscope, but this

technique can be particularly valuable for seeing small details of an image in stereo, and for showing them to others. Anaglyphs can be made manually in digital imaging software programs such as Adobe Photoshop, or by a number of special software packages.

Once one has a digital image file, image processing techniques can be used to increase the visibility of certain sorts of details in a fairly simple and understandable way. The ideal black-and-white, second-generation aerial photo product is of relatively low contrast to preserve details at both ends of the gray-scale (light-to-dark) range, but digitally changing the brightness and contrast of parts of an image can help clarify details. One way to vary these image properties without losing the dark or light ends of the spectrum is to adjust an image's histogram by moving the gamma point (i.e., the center of the dark-to-light range) without moving either end of the histogram (see Figure 3.1). Certain sorts of digital filtering, such as unsharp masking, can increase the contrast of the places in the image where the dark-to-light changes occur most abruptly. This is often called *edge enhancement*, since in fact such changes often do define the edges of objects or patterns in a photo.

There are many types of image processing that can be applied to digital images made from aerial photographs. For instance, it is sometimes possible to remove periodic noise or image blur using certain procedures. An almost unlimited number of digital spatial filters are available or can be designed that have widely varying effects on images. There are many software packages that do image processing at various levels, each doing slightly different things, ranging from freeware through packages costing tens of thousands of dollars. New versions constantly become available. Each original aerial photograph shows things differently, so it is impossible to specify any sort of formula for how to do image processing in any specific instance. Image processing is best viewed as a tool to be used interactively, along with optical stereoscopy, to help make different sorts of details in aerial photos easier to see and interpret.

The spatial arrangements of the picture elements or pixels that make up digital images can be altered in controllable ways to achieve image registration or rectification, or in fact to produce planimetric orthophotos. One of the most efficient and flexible ways to rotate and scale a photo to fit map coordinates or another photo is to use software packages that work with computer aided design (CAD) programs to display and manipulate images within a spatial coordinate system. Many of these programs also perform *rubber sheeting*, which can be used to match multiple points across an image to map- or field-derived control point coordinates. These methods can be used to register images in relatively flat portions of the landscape—where many sorts of industrial sites in fact are located—quite acceptably for many mapping purposes. The software creates a world file (*.res or *.wrl) that translates the image to the

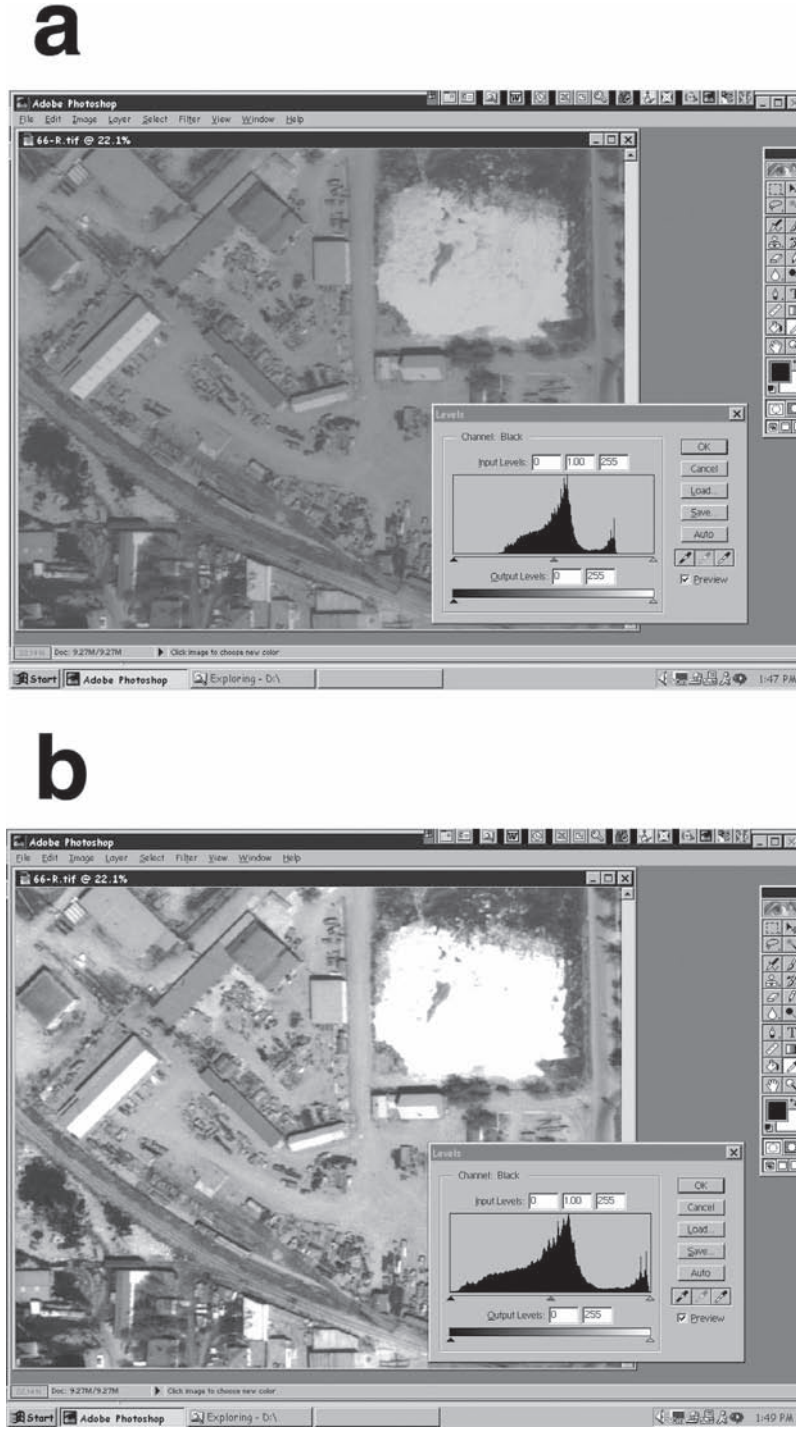


Figure 3.1

Histogram equalization redistributes the range of dark-to-light values in a digital image to increase contrast and thus the visibility of details. In Figure 3.1a, a scanned image exhibits a narrowed range of dark-to-light values (shown in the little graph in the lower right corner). By moving the light and dark ends of the graph to the edges of the histogram, and adjusting the middle of it (the 'gamma') the image is optimized. The new image and its histogram are shown in Figure 3.1b.

coordinate system again when it is brought into a CAD drawing, or geographic information system (GIS) such as ArcView. Lines and polygons showing photointerpreted objects or areas can then be drawn precisely on top of the image in CAD or GIS, and are then also registered to real-world coordinates.

3.3 PHOTOINTERPRETATION

Aerial photographs reveal information about natural and cultural features, processes, and interactions through time that cannot be derived from any other source, and for this reason they are important data sources in virtually every science and profession that focuses on spatially distributed phenomena. Aerial photos are used by biologists to map distributions of plants and animals, and by geologists, soil scientists, and engineers to characterize sediments and mineral deposits. Some of the earliest uses of aerial photographs, taken from kites and balloons before the advent of airplanes, were for military purposes, and military funding usually has been responsible for cutting-edge advances in imaging and mapping technology.

Interestingly, another profession that has pioneered many advances in photointerpretive and remote sensing methods is archaeology. Aerial photointerpretation has been an integral part of both archaeological fieldwork and the archaeological literature since the early 1900s in both the United States and Europe. The first archaeological aerial photographs were taken in 1906 over Stonehenge from a balloon. A number of early military pilots flying over Europe and northern Africa in World War I noted subtle patterning in vegetation, particularly in agricultural fields, that could not be seen from the ground and that revealed the locations of buried Roman forts and other ancient sites. Following the war, some of these aviators became some of the first "aerial archaeologists" at British and European universities. At about the same time, American archaeologists began using aerial photos, photographing sites such as the Cahokia Mounds and Southwestern pueblo ruins (Avery and Lyons, 1981). Some of the first and still most useful expositions of principles of photointerpretation were written by archaeologists such as O.G.S. Crawford (1923, 1924) and Poidebard (1929, 1930) (Ebert, 1984, 1997).

3.3.1 PHOTOINTERPRETATION VERSUS PHOTO READING

There are interesting parallels between the uses made of aerial photographs and remote sensor data used by archaeologists, and the ways environmental forensic scientists should approach the use of aerial photos as an adjunct to their work. Archaeologists' basic data consist of objects such as flaked stone tools or bits of pottery discarded or lost by past people in the course of their

lives and interaction with the environment to reconstruct not only how they were using a scraper or a pot immediately before they left it behind, but to explain things that can never be seen: mobility, settlement systems, religion—the sum of past human behavior. A fragment of pottery may be mildly interesting, and in some very rare cases one might even be able to derive a story from it, for instance, if a number of shards with the same pattern are found adjacent to one another, someone dropped the pot and it broke. But anecdotes such as this are of passing interest, illuminating only an instant in time. It is the intangible components of continuous and ongoing behavior and culture that are the goal of archaeological analysis. In fact, archaeologists do this same thing with another data source, aerial photographs, looking for indications of past sites and structures.

In much the same way, one can view objects relevant to environmental forensic studies in aerial photographs and assign names, but the ultimate goal is to elucidate the systemic effects of ongoing human behavior on the environment through time from aerial photos that show only instants. This fact is the basis of a distinction made by Thomas Eugene Avery, one of the United State's foremost teachers of photointerpretation, in the fifth edition of his book, *Fundamentals of Remote Sensing and Airphoto Interpretation*:

> *Photo interpretation* is defined as the process of identifying objects or conditions in aerial photographs and determining their meaning or significance. This process should not be confused with *photo reading*, which is concerned with only identifications. As such, photo interpretation is both *reading the lines* and *reading between the lines*.
>
> (Avery and Berlin, 1992: 51)

Utilization of aerial photographs in environmental investigations and litigation frequently follow the model of photo reading, rather than photointerpretation. Experts involved in environmental photointerpretation insist that film transparencies must be used rather than paper prints, or that one kind of scanner is better than another, because it lets you see certain small objects better. Reports contain elaborate overlays with arrows pointing to "drums," "fences," "lagoons," and "trenches." Part of the reason for this may be, in a sense, historical. Some of the first photointerpretive products most environmental scientists are exposed to are the reports of the Environmental Protection Agency's Environmental Monitoring Systems Laboratory. The U.S. Environmental Protection Agency (EPA) and photointerpreters under contract with them prepare *Waste Site Discovery Inventories* and *Detailed Waste Site Analyses* following very specific procedures and formats (Mace *et al.*, 1997). The purpose of the photointerpretive efforts is to identify environmental problem areas in a very general way (waste site inventories), and then to

identify more specifically features of properties that give reason to suspect there may be environmental problems at individual locations (detailed waste site analyses). The intent is also clearly *not to* propose activities or events in the past, *not to* read between the lines. The EPA's reports provide an excellent starting place for those involved in environmental forensic studies and environmental litigation—they supply identifications of objects and features, to be interpreted by others, and are *very purposely* the products of photo reading.

3.3.2 PHOTOINTERPRETATION IN ENVIRONMENTAL FORENSICS

What environmental forensic scientists and environmental litigators need from historic aerial photos is information that goes beyond simple identifications of features, into the realm of understanding the past *processes* by which environmental damages occurred. This is similar to what archaeologists do with aerial photos. What the archaeologist is looking for is quite clearly not what was going on when the aerial photo was taken, but rather indications of things that happened long ago. What is more, it is indications of things that often are buried beneath the surface of the ground, contained in subsurface sediments, or indications of landscape modification that subtly mirror past interactions between people and their environment. Of course, just about everything the archaeologist is interested in happened long before any aerial photos, historical or not, were taken. Archaeologists use all the aerial photos available because weather conditions, sun angle, season of the year, and recent landscape modifications change from photo date to photo date, making indications of past landscape use such as the condition of vegetation or ground compaction more or less visible.

Environmental forensic scientists have an easier task in many ways, since activities important at environmental sites really *were* going on when the aerial photo was exposed. Sometimes one can actually see a truck pouring dark liquid onto the ground, or a leaking drum or a burst storage tank. Much more often, however, what is environmentally and legally important is what was happening between the aerial photo exposures. After all, photographic exposures take place in only hundredths of a second. The task of the forensic environmental photointerpreter is to extract information from aerial photographs to elucidate the activities and processes that occurred that are the cause of environmental damages. Although this information may be obvious in aerial photos, some of that which is most interesting in environmental litigation is much less so.

It is interesting that a survey of 25 years of literature concerning the use of aerial photographs and other remote sensor data in environmental investiga-

tions shows an overwhelming focus on purposeful waste disposal, and dedicated waste disposal sites (see, for instance, Garofalo and Wobber, 1974; Philipson and Sangrey, 1977; Erb *et al.*, 1981; Getz *et al.*, 1983; Nelson *et al.*, 1983; Evans and Mata, 1984; Lyon, 1987; Bagheri and Hordon, 1988; Barnaba *et al.*, 1991; Pope *et al.*, 1996). Waste disposal is an activity that is directly responsible for much environmental damage, and locating waste disposal sites, determining what waste materials were disposed of through time and exactly where, determining ownership and whether disposal methods were proper or improper, are important issues in environmental litigation. But *deliberate and organized* waste disposal at sites designated in the past for such purposes is only one sort of activity responsible for environmental damage. Often, waste disposal of a less formal nature occurred at industrial sites, and identifying these locations using aerial photographs requires deductive photointerpretation.

Another contribution of archaeology to a better understanding of what it is environmental forensic scientists and environmental litigators might be most interested in learning from historic aerial photographs and other spatial data sources such as historic maps is the realization that our environment today is the result of a complex overlay of human activities through time. This is not something that archaeologists initially realized; in the mid-1980s, a series of theoretical changes within the field questioned whether one could assume that clusters of artifacts were the result of single episodes of human activity. Archaeology looked to anthropology for the answer. Anthropologists study present-day human behavior, and a century of anthropological observation focusing on hunter-gatherer societies showed that human activities of various sorts often occur repeatedly over long time periods at a place. During such activities, tools and debris are discarded or lost and form a complicated overlay that must be understood with reference to the ways people do things rather than relying on anecdotal evidence (Ebert, 1992).

Human *industrial behavior*, which is responsible for the environmental damages and other characteristics important in environmental forensics and environmental litigation, is a subset of general human behavior. Over approximately the last 200 years in the United States, and longer in some European countries, industrial and manufacturing activities have affected the environment in locations that have shifted through time, in a wide variety of ways as industrial processes and practices evolved. Though most environmental engineers and attorneys are quite aware that industrial perceptions and operations have changed radically since the advent of the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), far less attention is paid to the fact that only a very small percentage of possibly as many as half a million hazardous waste sites in the United States are inventoried by the

EPA (Colten, 1990). And most of those sites are not disposal sites per se, but rather the places where industrial activities took place.

At industrial sites, activities are, of course, organized to get the job done. Raw materials as well as chemicals used in processes arrive at the plant by water, rail, road, or pipeline transport (or sometimes by other means such as aerial tramways). When materials arrive at the plant, they are transferred from whatever means of transport by which they arrive into some form of storage. For instance, solvents are brought to the plant in tank trucks, and are transferred by means of a hose and pump to storage tanks. From the storage tanks, they are carried in glass carboys into the plant to be used to clean circuit boards in open baths. Once the solvents are contaminated, they are transferred by hand into wheeled, open-topped metal carts, taken out of the plant and into a storage area where they are poured into 55-gallon drums. The drums sit in the storage area until they are taken away by trucks to a recycling or disposal facility.

At each juncture when products are transferred from one means of containment, transport, or storage to another, a small amount is spilled. In storage, containers leak or chemicals that have been spilled onto containers are washed off by rain. Empty containers typically are rinsed out, often with a hose on a loading dock, before being reused. Although the amounts of chemicals that enter the environment through such *transfer and storage spillage* are small per event, hundreds or even thousands of such events can take place daily, involving dozens of sorts of materials, metals, chemicals, and the like at large industrial facilities.

At many industrial sites, a majority of environmental contamination has occurred relatively constantly through time, and instead of being accidental by any definition, was instead part of the normal operation of the plant. When true accidents occur—for instance, a drum falls from a truck and bursts open—the event is obvious, notice is taken, and cleanup efforts ensue. But small amounts of chemicals released as the consequence of routine product transfer and storage spillage or leakage, in legal parlance, are *expected and intended*. Distinguishing among accidental and expected and intended environmental damages is an especially crucial point in insurance defense litigation in the United States, for insurers cover their clients against accidental occurrences, and not routine, intended ones.

Although routine and intended occurrences are individually minor and go largely unnoticed at industrial plants, indications of their locations and extent often are visible in aerial photographs. Of course, the spillage of a few ounces of liquid chemicals that took place during transfer from a drum to a carboy in the yard of a plant cannot be seen on an aerial photo. The most important indicators are contextual—that is, the identification of facilities or features

that indicate areas of activities such as product transfer or spillage, or other means by which contaminants are released. Docks or racks where tank trucks or rail cars are unloaded can be the sites of extensive spillage, often indicated by dark staining, and even tracks of spilled material made by vehicles. Storage areas are another contextual indicator of places where spillage, leakage, and other transfers of contaminants to the environment can take place, and can be of several sorts. Bulk materials such as ore, fuel, and other process materials, and slag at metal refineries, are stored in obvious piles that usually dominate the landscape. Other liquids or solid materials often are stored in drums or other uniform containers, which are distinguishable even in small-scale aerial photographs by their regular patterning. Particularly important are places where materials are poured from, or into, drums or other containers. This generally takes place at docks or near entrances to plants. It is worthwhile noting that, due to radial displacement of objects, the sides of buildings and entrances on them can sometimes be seen in aerial photographs if the site falls away from the photo's center.

The products of industrial operations are often themselves contaminants, or the source of them; for instance, metal is often heavily oiled before being placed in outdoor storage for transport from a site. Treated wood is another notorious example of products from which contaminants can be washed into the site drainage by rainfall. Waste materials, such as used solvents, often are removed from operational areas and poured into storage containers such as drums outside of a plant. Any information garnered from depositions or knowledge of operating procedures about how and what is transferred can be vitally important in guiding the photointerpretation of such activities.

In some rare instances, when one fortuitously has access to extremely large-scale aerial photographs, it may be possible to actually see spills or liquids leaking from a drum or a tank. It is really the cumulative effects of such occurrences that are important, and for which photographic evidence usually can be seen. *Staining*, that is, dark patterning noted on the surface of the ground or pavement, often is cited as evidence of spillage or leakage of contaminants. There are in fact many things that can cause dark patterning in an aerial photograph. Most environmental photointerpreters recognize that water, for instance puddles or damp ground from recent rainfall, can cause surfaces to appear dark in places in both color and black-and-white film, and check local precipitation records before pronouncing dark places to be chemical spills or stains. Puddles appearing in places where chemicals would not be expected to be spilled are another indication to look for (Morrison, 2000). Water is copiously used in many industrial operations, too. When staining seems to be caused by water, some care should be taken to attempt to determine whether

it is "just water," or might contain contaminants, for example water drawn from tank bottoms at petroleum refineries.

Shadows also look dark in aerial photos, and distinguishing shadows from other dark patterning is one very important reason for interpreting aerial photographs stereoscopically so that objects that would project shadows can be detected. This works the other way, too, with shadows sometimes providing important cues in identifying objects that cast them. Shadows can be modeled and illuminated three-dimensionally using rendering software in CAD programs, to determine whether ambiguously appearing shadows are cast by known objects (Ebert, 2000).

The overall tone of aerial photographs also is affected by photographic properties such as exposure settings and the way film and subsequent generations of photo products made from it were exposed and processed. Vignetting at the edges of photo prints, particularly those made in black-and-white with early cameras, can cause some portions of a photo to look darker than others. Differences in sunlight intensity and angle can cause such differences from date to date when comparing photos of a site through time.

3.4 ANALYTICAL PRODUCTS AND PREPARATION OF EXHIBITS

3.4.1 MAPS

Ultimately, images have to be registered to maps. Photointerpretations made from multiple years of aerial photos are not of much use unless they can be considered in spatial relation to one another and to the landscape of today. In some cases, for instance at some industrial sites, large-scale surveyed plant maps are available and should be used to scale and register aerial photos. In many situations, however, engineering maps are not forthcoming and often the only maps that can be found are U.S. Geological Survey (USGS) $7\frac{1}{2}$-minute topographic sheets, which are fortunately available for the entire country. Compiled at 1:24,000 scale, they usually contain enough cultural detail to allow the registration of at least relatively recent aerial photos. The USGS calls the digital versions of their topographic sheets, which have in the last few years become available, digital raster graphics (DRGs). These scanned, georeferenced TIFF files are available for purchase in groups of 36 on CD-ROM; they can also be downloaded without charge from a number of web sites as well. DRGs are digitized or rasterized at a resolution of 250 dpi, and since they come with associated world files they can be brought directly into CAD or GIS in their respective UTM (metric) map coordinates, where digital images can be registered to them. They can also be used as a base coverage

for display purposes, although at 240 dpi (about 2.4 meters per pixel) they can be a bit difficult to look at closely, and one might want to scan a paper USGS map at higher resolution, and register it to the DRG, to be used as a base map for display purposes.

An orthophoto is a digital image of an aerial photograph that has been systematically corrected so that objects and features across the entire scene appear in their correct planimetric positions; orthophotos fit exactly onto a planimetric map of an area (although some sorts of features with abrupt topography, such as buildings, are difficult to correct in this manner). Orthophotos are created by using a digital elevation model of the area covered by the image to "rubber sheet" the image to a regular network of closely spaced control points. Orthophotos (called digital ortho quarter-quads or DOQQs and covering 3.75×3.75 minutes of latitude and longitude, one-quarter of a $7\frac{1}{2}$-minute topo quad) are available from the U.S. Geological Survey. They often provide better registration control source than DRGs since many features absent on maps, such as noncultural landscape details like trees, drainages, and the like, are visible for use as registration points in the orthophotos. USGS DOQQs are based on recent USGS aerial photographs, and conform to National Map Accuracy Standards, which require that objects on the ground be faithful to reality within 1/50″ on the map, or 40 feet on the ground as in a 1:24,000 scale map.

Contemporary aerial photographs can frequently be easily registered to either recent maps or orthophotos. When cultural features such as buildings and road intersections are visible on the maps or orthophotos, enough common points can be defined to allow registration as well as the calculation of the closeness of fit arrived at in the registration process. Historic aerial photos taken decades ago, however, often contain little detail that is the same as that on recent maps or even orthophotos. In many instances, it is necessary to work back through time, registering relatively recent photos to the map, and then using details on those photos that can also be seen in aerial photos a few years older to register *those* photos. The serial registration of historic aerial photos back through time to real-world coordinates can be a difficult process requiring much concentration.

Registration of aerial photographs to maps, or to other photographs, using simple scaling and rotation, or projective transforms, can have varying results in terms of closeness of fit depending on the nature of the terrain and cultural features in the area covered by the photos. Scaling and rotation work well when the area of interest is level. Projective transformation of an image can remove perspective effects caused by a sloping surface, but that surface must be relatively planar to produce good results. If there is considerable nonplanarity in the topography of the study area, there will be

mismatches of points across the terrain except for those used as control in the matching process.

True parallel axis photogrammetric plotting can produce highly accurate maps from historic aerial photographs. Photogrammetric engineering firms that have such plotting systems can produce planimetric and contour maps from historic aerial photos, but there are often limitations imposed by the scale of the aerial photos and the control points that can be found that mitigate against such maps being much more accurate than USGS $7^1/_2$-minute quads.

3.4.2 GEOGRAPHIC INFORMATION SYSTEMS

The reason to register aerial photographs in a real-world coordinate system is so that photointerpretations made from successive historic aerial photo dates can be compared spatially through time. Some site features such as property boundaries, roads, and railroads can be defined as lines, but most objects or areas relevant to past activities and present-day problems at a site are more appropriately polygons so that their correspondence or lack thereof through time can be analyzed. Buildings, tanks, pits, trenches, pools, stains, storage areas, loading and unloading areas, and many other sorts of features can and often do change locations through time at sites, and one of the best ways to keep track of such changes and correlate them with other site data *not* derived directly from aerial photos (for instance, sample locations and depictions of chemical plumes derived from them) is by incorporating them all in a geographic information system (GIS) database. Coordinating the efforts and data of the diverse experts involved in any large-scale, serious environmental litigation virtually requires a unified database that can be used by all parties concerned.

Only a few years ago, compiling and using a GIS database required hardware, software, and technical expertise and training that was beyond even many large environmental firms. Fortunately, more affordable and user-friendly software coupled with the geometric increase in the capabilities of computers has resulted in GIS being a tool that belongs on everyone's desk. There are dozens of GIS software packages with varying levels of capability and sophistication available today, most of which are affordable enough that multiple licenses can be purchased within the scope of any significant litigation. In a GIS database, information is spatially organized in multiple layers by data type, time, and/or other appropriate classification, and can then be compared, combined, and manipulated in a number of ways. Overlapping areas of the same type can be combined, differenced, or viewed additively through time, something that will be discussed in the context of a case study

later in this chapter. Phenomena that vary across space, for instance chemical concentrations, can serve as the basis for producing contour maps, or sample locations compared in terms of their proximity to buildings, roads, or other spatially organized features. Some GIS software packages have sophisticated graphic and cartographic capabilities that can be used to produce exhibits illustrating locations of features and areas within a plant or other study site, and the results of analyses. A properly compiled GIS database can organize all the relevant spatial data in a case in one place so it can be used by everyone in many places.

3.4.3 TERRAIN VISUALIZATION

Nonanalytical digital visualization techniques are available for the exposition of information from aerial photographs and other data sources in environmental forensic cases. Terrain visualization is one of these, and can be used to help orient judges, juries, deponents, and other participants in litigation to the topography and the arrangement of site features. Terrain visualization consists of draping a map or aerial photograph over a surface constructed with a digital elevation model (DEM). The resultant rendered surface can then be viewed from any angle, elevation, or distance, or multiple views rendered and chained together to produce an animation that is essentially a "walk-through" or "fly-by." A number of GIS, image processing, and other software packages are available, which will create three-dimensional surface/subsurface visualizations.

3.5 CASE STUDIES

Four case studies illustrated by air photo images are included to illustrate some of the principles of environmental forensic photointerpretation and analysis. Those who would be interested in viewing the aerial images from the case studies in the form of stereo anaglyphs can do so at www.ebert.com.

3.5.1 CASE STUDY: INFORMAL DUMPING AREA COVERED BY PLANT ADDITION

In older urban areas, the environmental state of industrial sites and the soils beneath them is the result of a composite overlay of past industrial activities and changing industrial behavior. At a bearing manufacturing plant site in the northeastern United States, subsurface contamination was concentrated beneath the southern part of the factory, and there were questions as to its possible source. Leakage of oil and solvents from process areas within the

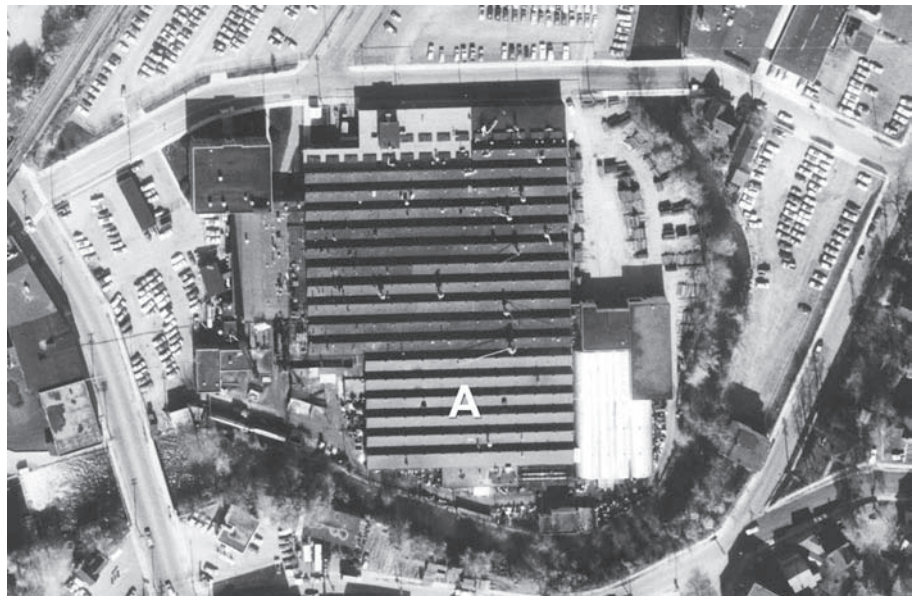


Figure 3.2
A bearing manufacturing plant in 1975 in the northeastern United States had unexplainably high concentrations of solvents and other chemicals under its slab at A.

plant, and through the concrete floor, was suggested but seemed implausible. Figure 3.2 is a high quality image of the plant in 1975, and around the southern portion of the plant under which contaminant concentrations were highest (marked A), large numbers of drums and containers can be seen to be stored in a haphazard manner in outside corners and along the sides of a service road. Such containers, whether filled with chemicals going into the plant, waste materials, or even empty, are potential sources of leakage and product transfer spillage, but if they were the major source of contamination then concentrations around the perimeter of the buildings would be expected to be as high or higher than under the slab.

The deposition testimony of a machinist who worked at the plant in the 1930s mentioned informal but continuous dumping and pouring of wastes "on the island," which was described as a place you would not want to go if you did not have to because it could ruin your shoes. Figure 3.3 is a stereo anaglyphic image derived digitally from aerial photographs flown by the Soil Conservation Service in 1938. Anaglyphs must be viewed with the axis of the flight line from side to side; the 1975 aerial photos were flown east-to-west, so in them north is up, whereas the 1938 photos were flown north-to-south, so north is toward the left. Although the resolution of the 1938 photos is low and the contrast very high, it can be seen that in 1938, in the area where the southern portion of the plant had yet to be constructed (A), an abandoned river channel is visible by virtue of reflected sunlight, and between it and the river lies the island. Dumping there was at least part of the source of hydro-

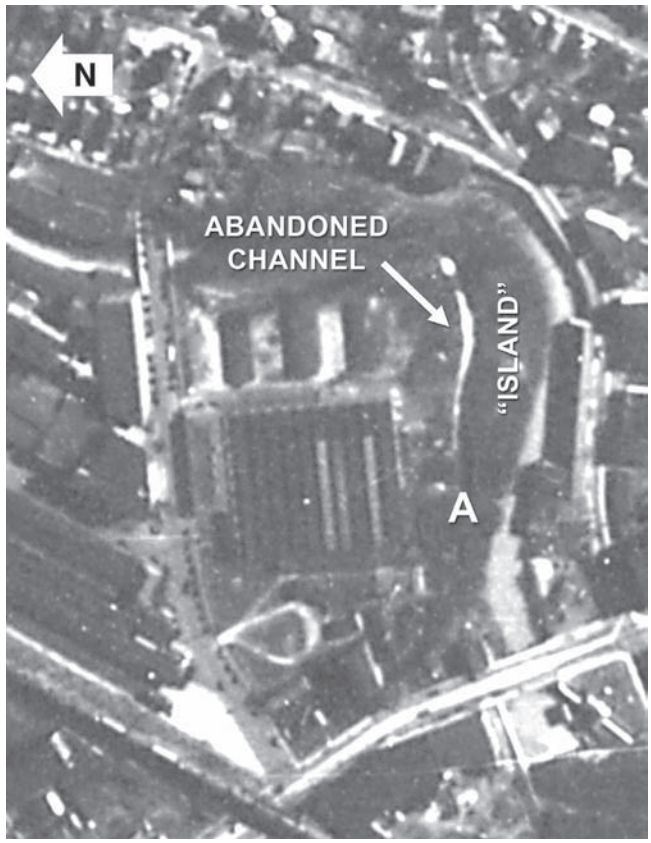


Figure 3.3
A 1938 aerial photograph of the bearing manufacturing plant in Figure 3.2 shows an abandoned river channel and an area called 'the island' by a deponent, where large amounts of dumping occurred prior to the building's annex at A being built.

carbon contamination beneath the southern annex to the factory in later times.

3.5.2 CASE STUDY: PRODUCT TRANSFER AND STORAGE SPILLAGE AT CIRCUIT BOARD FABRICATION PLANTS

Circuit board manufacture involves the use of etching acids, fluxes, and solvents for cleaning parts. These chemicals are brought to these industrial sites in bulk, usually in carboys and drums, which are unloaded from trucks and often stored outside the plant on or near loading docks. Since they are used inside the plant in relatively small amounts, for instance in tubs or machines in which circuit boards are cleaned, the liquids are transferred from drums or carboys into other containers—in the case of the two plants discussed here, according to deposition testimony, in specially made carts, consisting of a vertical tank attached to a two-wheeled hand dolly—using centrifugal pumps (for drums) or by hand pouring (from carboys). Centrifugal pumps operate by turning a crank affixed to a pump with a pipe that screws into the open drum bung; when pumping stops, and the hose is withdrawn from

the receptacle, whatever liquid remains in the hose drains out suddenly, and spillage onto the ground and drum is inevitable. Other methods of transferring liquids from drums include rocking cradles and funnels. Liquids in carboys, typically containing 5 or 10 gallons, are often simply poured, using funnels, into other containers. Filled carboys may weigh 100 pounds (45 kg) or more, so pouring carefully is difficult. Because spillage in such transfer operations is expected, these activities often occur outside the plant buildings, in storage areas.

This was the case in 1963 at the circuit board fabrication plant shown in Figure 3.4, where drums and other containers are stored around the loading docks at A, and in a long row along the side of the parking lot at B. At C in this stereo anaglyphic image are a number of aboveground storage tanks into which other liquid chemicals were transferred from tank trucks as well as drums to be piped into the plant. The parking lot at D shows mottled staining (staining can be light-colored, as well as dark), which may be the result of transfer spillage. The objects stored in the parking lot below D appear to be largely equipment, although some of them may be containers as well.

Some of the containers visible in Plate 3 are undoubtedly filled with used chemicals, to be removed from the plant for disposal or recycling. Spent liquids are poured into drums with more pumps and funnels, with additional spillage—often outside the plant, in the storage yard, because of the mess it would create inside.

Figure 3.5 shows another circuit board fabrication facility in the western United States. The most obvious black patterning is water flowing from the cooling tower, which probably should not be happening in such volume, but which is also probably environmentally benign. That water is flowing into a drain consisting of a small ditch in the parking lot with a perforated cover.

It is the product transfer area at this plant which is far more interesting in terms of environmental forensics, an area where, according to deposition testimony, you "wouldn't want to drive." In the enlargement of this area in Figure 3.6, a hydrant and hose near the spill-stained loading dock, and a huge plume of water flowing hundreds of feet down the drive to the drain, are apparent. The deponents said the dock manager liked to rinse out carboys and drums in the driveway from which chemical liquids were transferred into the plant so that they would be clean for reuse. Bulk liquid chemicals were also brought in drums and tank trucks to the covered loading dock, and drums unloaded there were stored outside at its right end, as can be seen from the stereo anaglyph. Also visible in the anaglyph is staining on the asphalt from liquid running from this outdoor container storage area into the nearby drain ditch.

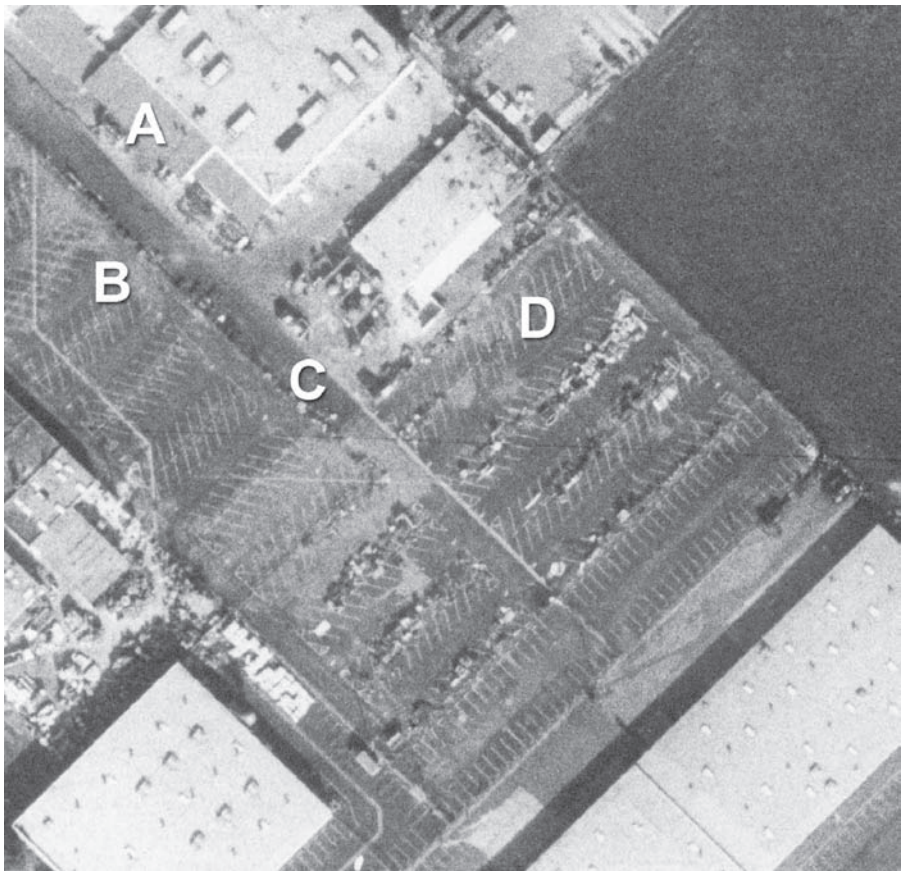


Figure 3.4

At a semiconductor board fabrication plant in 1963, liquid chemicals were being transferred from drums and carboys at the loading dock (A) and in the container storage area at B. Chemicals were being transferred from tank trucks to aboveground storage tanks at C. Mottled patterning on the asphalt at D may be the result of product transfer spillage as well.

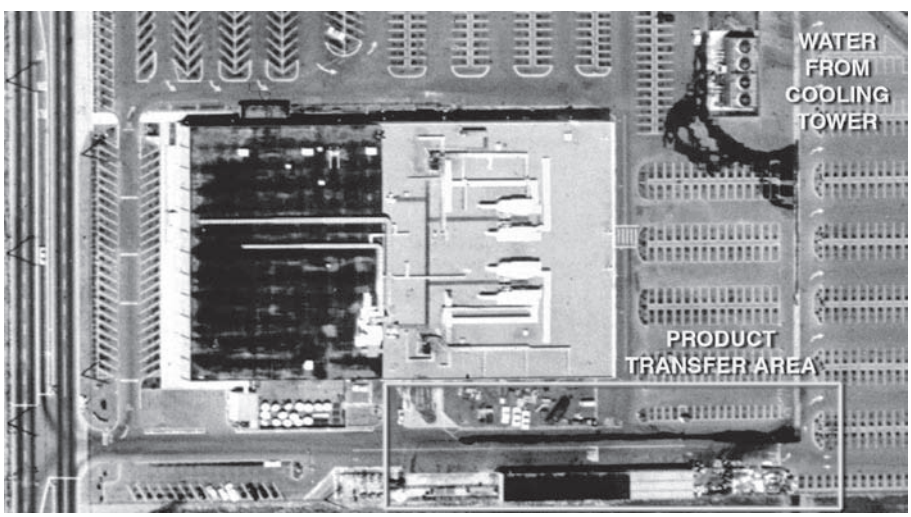


Figure 3.5

While water from the cooling tower is the most obvious 'dark stain' in this image, the product transfer area is of greater interest from an environmental forensic standpoint. A close-up of the product transfer area is shown in Figure 3.6.

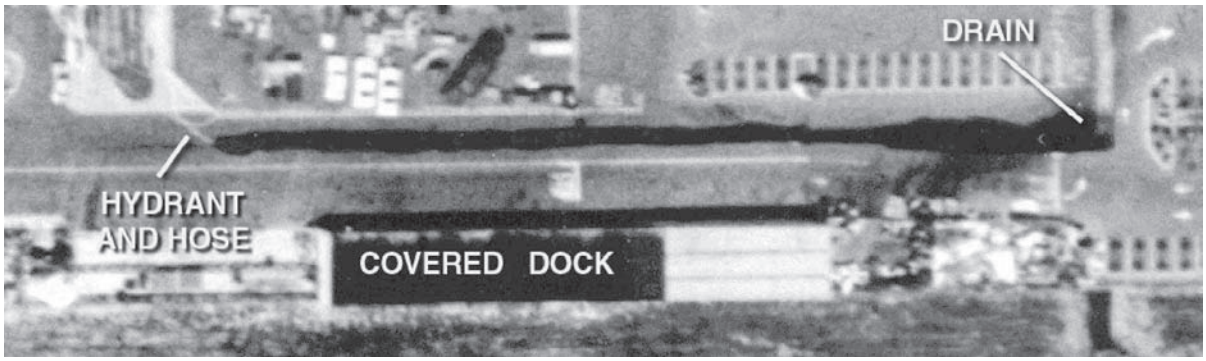


Figure 3.6

This close-up of the circuit board fabrication plant product transfer area outlined in Figure 3.5 shows that a hydrant and hose used to wash out drums and carboys. The effluent ran the length of the drive, and emptied into a drain at its end. The covered dock and the adjacent uncovered container storage area to its east were used to store chemical containers, and a dark stain resulting from product transfer spillage is visible issuing from the open storage area into the drain.

3.5.3 CASE STUDY: CHEMICAL DRAINAGE FROM TREATED WOOD AT A CREOSOTE PLANT

Although creosote used to treat wood is a natural product and relativly involatile, only transmitted in sediments slowly, it is susceptible to transport in water. This can occur when treated wood is stored outdoors (which of course it must be) and subjected to rainfall. Figures 3.6 and 3.7 show a wood treatment plant in the southeastern United States in 1959 and 1965. At issue in this case was the question of when wood treatment chemicals from the plant could have entered the new drainage ditch, designated as such on the 1959 aerial photo.

Untreated wood piles are light colored, almost white, in the aerial photos, and treated wood is dark. In 1959 (Figure 3.7) there is little treated wood near the new drainage ditch, but in 1965 (Figure 3.8) the practice seems to be to have stored more treated wood along the ditch, from which rainfall-washed chemicals could have drained into it.

3.5.4 CASE STUDY: DARK STAINING AT A MAINTENANCE YARD

An illustration of why it is necessary to exercise caution when citing staining or dark patterning as evidence of site contamination is provided by a case study involving an urban property in the southwestern United States damaged largely by petroleum hydrocarbons and associated chemicals. At issue was the question of the differential responsibility of two consecutive owners of the facility, a construction company who used the property for maintenance and storage of construction equipment for approximately 40 years, versus the

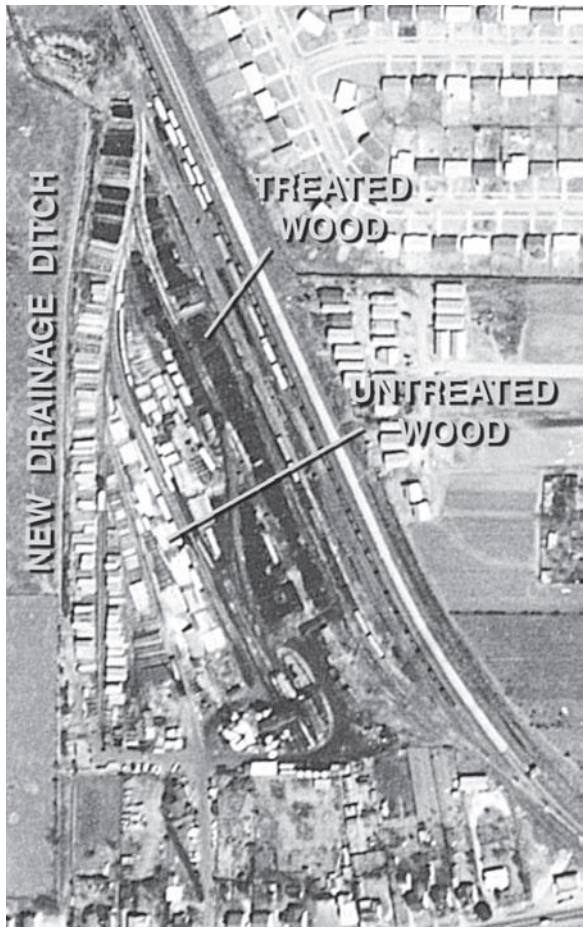


Figure 3.7
In 1959, the treated (dark colored) wood at this wood treatment plant in the southeastern United States was mostly stored on the east side of the plant. A new drainage ditch had just been dug on the west edge of the facility.

subsequent owner and plaintiff, who conducted an oil recycling business there.

The plaintiff engaged the services of an experienced and qualified photogrammetrist who located all the available aerial photos of the facility, photointerpreted them using mirror stereoscopes, and used stereo on-screen computer viewing. The photointerpreter counted construction vehicles on-site at various times of the year in a series of 20 sets of aerial photos spanning a period of 26 years, illustrating that over this time many vehicles potentially had been on the site. For each of the aerial photo dates, the photo-interpreter also delineated areas of what was defined as dark staining on the ground surface. This being a vehicle maintenance site, it was further reasoned, the dark staining was due to the spillage of lubricating oils and greases.

The historical aerial photos of the plant were registered to a common coordinate system, and the areas of dark surface staining were delineated and digitized and exported to ArcView GIS for analysis. Coverages within a GIS

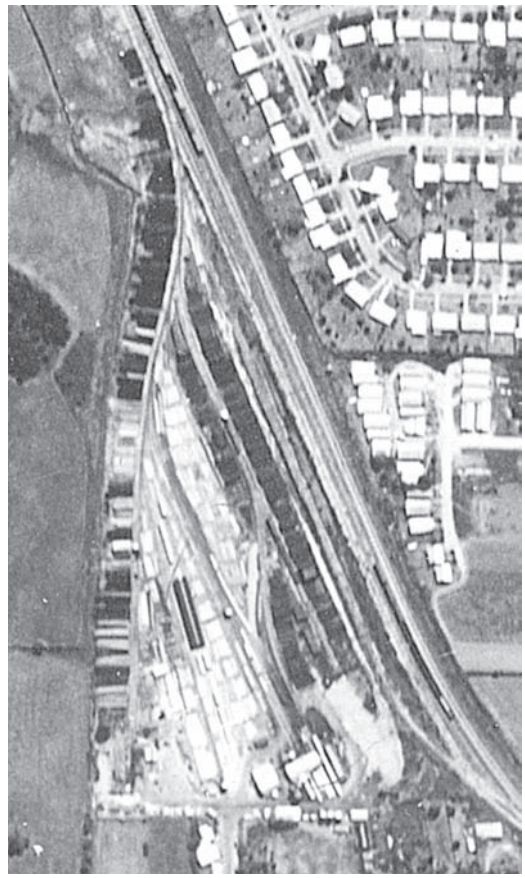


Figure 3.8
Storage of large piles of treated wood adjacent to the western edge of the wood treatment plant by 1965 allowed creosote laden runoff to enter the drainage ditch there and flow into nearby neighborhoods.

database, for instance the dark surface stains digitized from the air photos for each year, can be compared, combined, measured, and analyzed in a number of ways. One of these is a union of all the coverages through time. Essentially, the polygons or bounded areas in all the coverages are overlain upon one another, and their internal boundaries are dissolved, resulting in combining overlapping areas into one. This is what the plaintiff's expert did to show the extent of what had been interpreted as dark stained areas at the property that was the subject of litigation, resulting in a map like that shown in Figure 3.9, left. The implication, and the impression given, by this exhibit is that somewhat more than 40% of the surface of the plant site is covered with lubricating oils and greases.

At least two levels of criticism might be leveled at this analysis and its resulting implications. First, even if all those dark stains *were* indications of spills of lubricating oils and greases, each ArcView stain coverage records only those areas interpreted as dark when the aerial photos were taken. Representing their union as a single stain type does not take into account that some areas

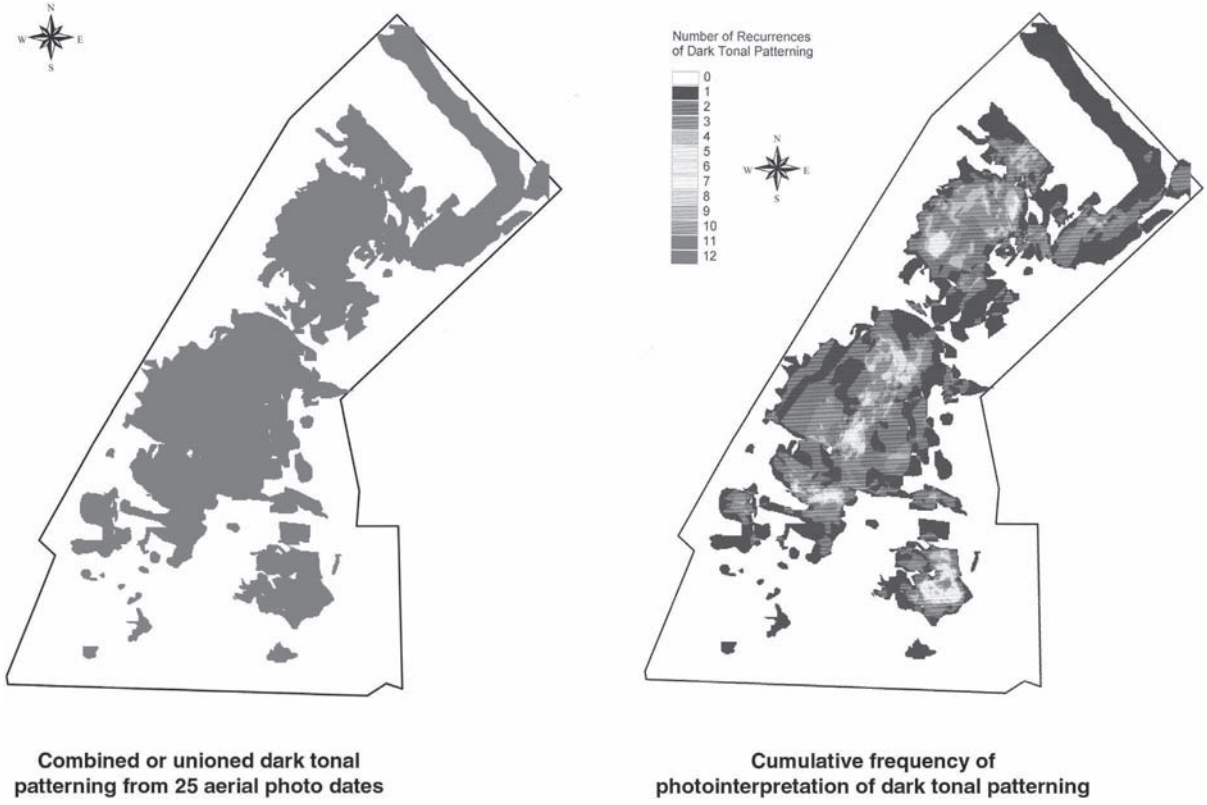


Figure 3.9

Two maps show alternate analyses of photointerpretation of 'dark staining' in the yard of a construction maintenance plant over 20 sets of aerial photographs spanning a period of 26 years. The map on the left is all of the interpreted dark staining unioned or overlain by ArcInfo, a geographic information system (GIS) software package. In the map on the right, the cumulative frequency of interpreted dark staining is shown in colors ranging from dark blue through dark red, showing that only a very small portion of the facility exhibited repetitive dark tones through time, a pattern consistent with human industrial behavior. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

must have been stained only once, whereas other areas would have been stained on more than one aerial photo date. Determining which areas were interpreted as stained more than once might be very informative in terms of determining where the environmental damage would be expected to be most severe.

A related but more severe critique, however, of what the unioned GIS coverages imply is more clearly behavioral, and that it is probably very unlikely, if not impossible, that so much of the facility's surface could be stained with oils and greases over even the 40 years that construction equipment maintenance took place there. The configuration of the facility (building locations, etc.) did not appreciably change during that period. In order for oil and grease to

have stained this much of the yard, they would have to have been spilled, poured, drained, and dumped virtually *everywhere*, constantly, over thousands and thousands of square feet.

This just is not the way any sort of human activities take place. Deposition testimony in the case revealed that, occasionally and unfortunately, oil from heavy construction vehicles was drained onto the ground at ongoing construction sites, but never at a maintenance yard.

A consideration of some of general properties of human behavior is helpful in understanding how activities—whether in hunter-gatherer camps or vehicle maintenance yards—are patterned in space. Anthropological observation illustrates that at large or small sites, specific sorts of human activities are spatially segregated, partly because the necessary tools and facilities (i.e., nonportable articles) are more or less anchored in space. For instance, in Bushman sites in the Kalahari Desert of southern Africa, *mongongo* nuts, a staple at certain times of the year, are cracked in circumscribed areas where a large, heavy stone anvil, cracking stones, and a hearth for roasting the nuts in their pods are juxtaposed. The pods and shells are disposed of in a pit or open area close to the processing locus. Another critical component is often a shade tree under which those doing the processing can sit.

In the same way, changing oil and greasing heavy equipment requires spatial correlations among implements and facilities. The necessary tools, receptacles for receiving waste oil and grease, and supplies of new oil and grease, must be located together—and not all over the yard. Construction vehicles are brought to the place where oil changes and greasing occur, and then when that is over they are driven to where they are stored, short or long term. Interim storage, from which waste oils periodically would have been removed from the site, would be expected near the locations where oil was changed. In order to attempt to discover where the majority of oil changing and greasing occurred at the site, the defense undertook an alternative spatial analysis strategy, using the plaintiff's expert's own GIS data, which was obtained upon discovery. Instead of unioning the coverages for the 20 aerial photo dates interpreted by the plaintiff's expert, an analysis was performed using the capabilities of the GIS that quantified and represented the number of times his dark staining overlapped—that is, *intersected*—throughout the entire aerial photo series. Figure 3.9 (right) is a contour map that illustrates how often the dark staining occurred repetitively, photo date after photo date.

The results of the occurrences of dark staining as a percentage of the plant site (see Table 3.1) were striking. Although the plaintiff's expert showed that about 40% of the property was covered by dark staining in his GIS analysis, about half that area was stained only once, that is, in areas from just one aerial photo date. Only 10% of the total plant site was found to be stained twice,

Occurrences of Dark Staining	Percentage of Total Plant Site
0	59.90
1	20.12
2	10.09
3	5.84
4	2.51
5	0.86
6	0.39
7	0.16
8	0.04
9	0.03
10	0.03
11	0.02
12	0.01

Table 3.1

Area of the construction equipment maintenance plant showing no dark staining and dark staining occurring from 1 to 12 times, over the span of 20 aerial photo dates interpreted by the plaintiff's expert.

and for large numbers of recurrence of staining, the percentage of the site area decreased asymptotically. The plaintiff's expert illustrated that in 40% of the site, dark staining was interpreted at least once. Areas of the plant in which the plaintiff's expert interpreted dark staining even seven times is only a bit more than 0.1%. The highest frequency of repetitive dark surface staining interpreted by the plaintiff's expert was 12 times in 20 photo dates, covering 0.01%, or one ten-thousandth, of the surface area of the property. Interestingly, that area was divided between two places at the site: the location of the filler tube, adjacent to the main workshop, for an underground tank where waste oil was stored to be taken away for recycling; and around a steam cleaning unit.

At the waste oil storage tank, it is quite likely that the dark staining *is*, in fact, the result of repeated oil spillage. Deposition testimony showed that waste oil was carried to the tank in a container made from an oil drum cut in half, and poured into a funnel-like receptacle into the tank. In the vicinity of the steam cleaning pad, the dark patterning noted by the plaintiff's photointerpreter was water. In some of the aerial photographs, looping patterns caused by spray from the steam cleaners' wand were visible. The steam cleaning occurred over a concrete pad in the gravel yard, in the center of which, in one of the higher resolution aerial photos, the opening of a drain was visible. Water sprayed on vehicles, which may have contained some residual grease and certainly dirt, went down the drain.

Just what all the less-recurrent dark patterning interpreted as staining at the site might have been is a matter for logical evaluation. Many things that are not staining in any sense cause dark patterns in aerial photos, as was discussed earlier. At the construction maintenance site, vegetation may have been



Figure 3.10
In this photograph of the construction vehicle maintenance yard in 1966 there is little contrast difference in the yard indicative of 'dark staining'. The steam cleaning pad is located at A and a very reflective area of bare earth or possibly stockpiled light colored material is at B.

responsible for much of the dark patterning. Figures 3.10 and 3.11 show enlarged portions of stereo aerial photographs taken of the construction equipment maintenance site in 1966 and 1969, respectively. The steam cleaning unit appears next to the A in the 1966 photo, which has a ring of dark patterning—water—around its pad. In the 1969 photo, steam cleaning is still occurring at the same location. The dark-and-light patterning of the gravel yard and at other features such as the roofs of buildings are different in the two aerial photographs. In 1966, the yard and the areas around the equipment and stored materials are all relatively dark in tone; in 1969, the roadways between the equipment and material are much lighter in tone, probably because they were recently graded. Other parts of the plant are dark in relation to the light-toned roadways. At B in the 1966 photo, a large area covered by light-toned, dumped material appears. By 1969 the area is covered by vegetation, and appears dark in tone.

3.6 CONCLUSIONS

Photointerpretation requires the use of the proper equipment, the correct types of aerial photographs, and sophisticated instruments, methods, and techniques that have been developed by the photogrammetric profession over the course of more than 100 years. In the course of about 15 years, however,

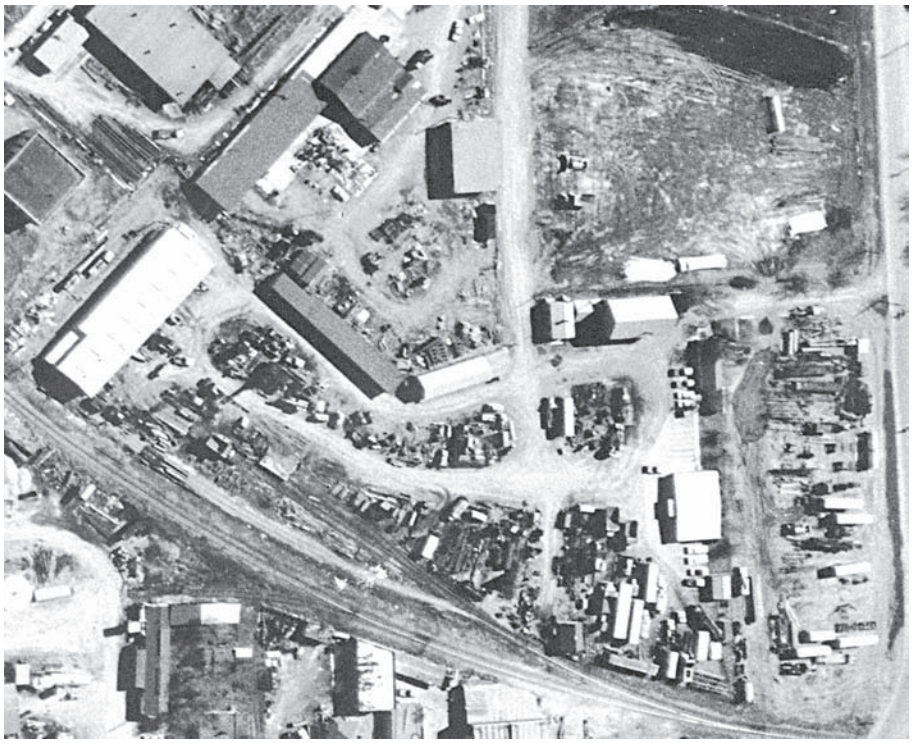


Figure 3.11
The construction vehicle maintenance yard as it appears in a 1969 aerial photograph shows considerably more contrast than it does in Figure 3.10, probably due to the printing of the aerial photograph coupled with grading of the open ways between stored equipment. Differences in photo exposure and sunlight often cause one aerial photo to look dark while another looks much lighter.

photogrammetry has been transformed by the advent of computer technologies that have added digital imaging and image processing and digital mapping methods such as CAD and GIS to its repertoire. The acquisition and careful examination of aerial photographs, using all the proper instruments, methods, hardware, and software is only part of using photogrammetry to its fullest advantage in environmental forensics and environmental litigation. In the final analysis, the most important part of photointerpretation—as opposed to photo reading—is to use the information presented by historical aerial photos to support logical and consistent reasoning about the processes and human industrial behavior that took place at sites through time, and shaped the nature of the environment there.

REFERENCES

Avery, T.E. and Berlin, G.L. (1992) *Fundamentals of Remote Sensing and Airphoto Interpretation*, 5th ed. Macmillan, New York.

Avery, T.E. and Lyons, T.R. (1981) *Remote Sensing: Aerial and Terrestrial Photography for Archeologists*. National Park Service, Cultural Resources Management Division, Washington, DC.

Bagheri, S. and Hordon, R.M. (1988) Hazardous waste site identification using aerial photography: A study in Burlington County, New Jersey, USA. *Environmental Management* 12(1), 119–125.

Barnaba, E.M., Philipson, W.R., Ingram, A.W. and Pim, J. (1991) The use of aerial photographs in county inventories of waste-disposal sites. *Photogrammetric Engineering and Remote Sensing* 57, 1289–1296.

Colten, C.E. (1990) Historical hazards: The geography of relict industrial wastes. *Professional Geographer* 42(2), 143–156.

Colwell, R.N. (1997) History and place of photographic interpretation. In *Manual of Photographic Interpretation*, 2nd ed., W.R. Philipson, ed. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 3–47.

Crawford, O.G.S. (1923) Air survey and archaeology. *Geographical Journal* 61, 342–366.

Crawford, O.G.S. (1924) Archaeology from the air. *Nature* 114, 580–582.

Ebert, J.I. (1984) Remote sensing applications in archaeology. *Advances in Archaeological Method and Theory*, vol. 7, M.B. Schiffer, ed. Academic Press, New York, 293–362.

Ebert, J.I. (1992) *Distributional Archaeology*. University of New Mexico Press, Albuquerque, NM.

Ebert, J.I., ed. (1997) Archaeology and cultural resource management. In *Manual of Photographic Interpretation*, 2nd ed., W.R. Philipson, ed. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 555–589.

Ebert, J. (2000) Shadowy defense: 3D modeling at the scene of the crime. *GeoSpatial Solutions* 10(7), 28–33.

Erb, T.L., Philipson, W.R., Teng, W.L., and Liang, T. (1981) Analysis of landfills with historic airphotos. *Photogrammetric Engineering and Remote Sensing* 47(9), 1363–1369.

Evans, B.M. and Mata, L. (1984) Aerial photographic analysis of hazardous waste disposal sites. *Hazardous Wastes and Environmental Emergencies National Conference*. Hazardous Material Control Research Institute, Houston, TX, 97–103.

Garofalo, D. and Wobber, F.J. (1974) Solid waste and remote sensing. *Photogrammetric Engineering and Remote Sensing* 15(2), 45–59.

Getz, T.J., Randolph, J.C., and Echelberger, W.F. Jr. (1983) Environmental application of aerial reconnaissance to search for open dumps. *Environmental Management* 7(6), 553–562.

Lyon, J.G. (1987) Use of maps, aerial photographs and other remote sensor data for practical evaluations of hazardous waste sites. *Photogrammetric Engineering and Remote Sensing* 53(5), 515–519.

Mace, T.H., Williams, D.R., Duggan, J.R., Norton, D.J., and Muchoney, D.M. (1997) Environmental monitoring. In *Manual of Photographic Interpretation*, 2nd ed. W.R. Philipson, ed. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 591–612.

Morrison, R. (2000) *Environmental Forensics Principles and Applications*. CRC Press, Boca Raton, FL.

Nelson, A.B., Hartshorn, L.A., and Young, R.A. (1983) A methodology to inventory, classify and prioritize uncontrolled waste disposal sites. US EPA, No. 68-03-3049, Las Vegas, NV.

Philipson, W.R. and Sangrey, D.A. (1977) Aerial detection techniques for landfill pollutants. In *Proceedings of the 3rd Annual EPA Research Symposium on Management of Gas & Leachate in Landfills*. St. Louis. EPA, Washington, DC, 11.

Poidebard, A. (1929) Les révélations archéologiques de la photographie aérienne: Une nouvelle méthode de recherches d'observations en région de steppe. *L'Illustration*, May 25, 600–662.

Poidebard, A. (1930) Sur les traces de Rome: Exploration archéologique aérienne en Syrie. *L'Illustration*, December 19, 560–563.

Pope, P., Van Eeckhout, E., and Rofer, C. (1996) Waste site characterization through digital analysis of historical aerial photographs. *Photogrammetric Engineering and Remote Sensing* 62(12), 1387–1394.

Teng, W.L. (1997) Fundamental of photographic interpretation. In *Manual of Photographic Interpretation*, 2nd ed. W.R. Philipson, ed. American Society for Photogrammetry and Remote Sensing, Bethesda, MD, 49–113.

CHAPTER 4

THE MEASUREMENT PROCESS

Dallas Wait and Charles Ramsey

4.1 Introduction 84
4.2 Measurement Process 85
4.3 Planning 90
4.3.1 Questions 91
4.3.2 Decision Unit 91
4.3.3 Confidence 93
4.3.4 Sample Plan Design 94
4.3.4.1 Sampling for Identification 95
4.3.4.2 Sampling to Determine Concentration 96
4.3.4.3 Quality Control Measurements 97
4.4 Implementation 99
4.4.1 Plan Implementation 99
4.4.1.1 Soil Gas 99
4.4.1.2 Soil 100
4.4.1.3 Sediment 102
4.4.1.4 Groundwater 103
4.4.1.5 Surface Water 103
4.4.1.6 Surfaces 104
4.4.1.7 Ambient Air 104
4.4.2 Sample Handling 105
4.4.3 Subsampling 106
4.4.4 Analysis 107
4.5 Assessment 111
4.5.1 Documentation 111
4.5.2 Quality Control Interpretation 112
4.5.3 Statistical Analysis 116
4.5.4 Data Usability Assessment 116
4.6 Summary 118
References 118

4.1 INTRODUCTION

The need for measurements that are reliable, defensible, and of known quality is paramount to a successful forensic investigation. Therefore, it is important that forensic investigators craft and implement sampling and analysis plans that properly address the objectives of their study (Wait, 2002). The threshold for what constitutes reliable and relevant data will vary from study to study depending on the needs of the study but these requirements have been legally defined, when applicable, in the 1993 Supreme Court Daubert ruling (USSC, 1993). The factors bearing on reliability and relevance include:

- Whether the method employed is generally accepted in the scientific community
- Whether the method has been subject to peer review
- Whether the method can be or has been tested
- Whether there are standards controlling the method's performance
- Whether the known or potential rate of error is acceptable

From a regulatory standpoint, after the 1979 Love Canal data quality debacle (OTA, 1983; Maney, 2002), Doug Costle, U.S. Environmental Protection Agency (US EPA) Administrator, threw down the data quality gauntlet by stating in part, "Reliable data must be available to answer questions concerning environmental quality and pollution abatement and control measures. This can be done only through rigorous adherence to established quality assurance techniques and practices." Further, "these data must be scientifically valid, defensible, and of known precision and accuracy" (US EPA, 1979). Costle's mandate is reflected in subsequent US EPA test method and quality system documents, guidance and regulations, and has helped shape current environmental forensic practices (Wait, 2000, 2002). These types of data quality initiatives are more recently reflected in international measurement programs such as the test method guidelines and good laboratory practices promulgated by the Organization for Economic Co-operation and Development (OECD, 1993; Olson, 1999) and the international laboratory competence requirements (ISO/IEC 17025) promoted by the International Organization for Standardization (ISO) and the International Electrochemical Commission (IEC) (ISO/IEC, 1999).

This chapter initially discusses the data life cycle and what constitutes a measurement. Guidance is then provided on how to design a defensible sampling and analysis plan. To support this discussion, sampling techniques, tips, and pitfalls are provided for numerous matrices important to forensic investigators. These matrices include soil gas, soil, sediment, groundwater, surface water, surfaces, and ambient air. Appropriate test methods and necessary

supporting documentation is then considered. The chapter concludes with recommendations on how to assess the usability, reliability, and admissibility of data.

4.2 MEASUREMENT PROCESS

The environmental measurement process consists of three phases: planning, implementation, and assessment, and is also known as the project life cycle (Maney and Wait, 2005). Measurement of an environmental process requires that numerous levels of detail be properly addressed. Some examples are provided in Table 4.1 (Maney and Wait, 1991). Many of the details listed in Table 4.1 may not be applicable to every forensic study, but should be actively considered prior to exclusion. Some of these details, such as site history, contaminant transport, and testing methodologies are discussed in greater detail in other chapters in this book.

Understanding the intricacies of the measurement process and associated errors are key to the collection of accurate measurements (Allchin, 2002). Errors relate to three general categories: blunders, systematic errors (bias), and random errors (imprecision), whose combined impact is referred to as integrated error (Maney, 2002). Blunders are mistakes that typically occur only occasionally, such as misweighings or mislabelings. Systematic errors are measurements that are consistently different from their true values, such as underreporting a contaminant concentration, that results in a bias. Random errors arise from the difficulty of repeating sampling and analytical processes, as well as the heterogeneity of the population being studied.

Excessive measurement errors jeopardize the integrity and reliability of forensic studies and occur more frequently than some may realize. For example, one investigation found that 11% of 2000 Food and Drug Administration (FDA) studies had "serious deficiencies" in measurements (Shapiro and Charrow, 1989). In another example, United States Geological Survey (USGS) researchers questioned the fundamental reliability of physical constant data used by environmental and forensic scientists. Specifically, Pontolillo and Eganhouse (2001) reviewed approximately 700 reports between 1944 and 2001 that provided data on the aqueous solubilities and the octanol-water partition coefficients for dichlorodiphenyl trichlor ethane (DDT) and dichlorodiphenyl dichloroethylene (DDE) and found "egregious errors in reporting data and references, and poor data quality and/or inadequate documentation of procedures" with no convergence over time. Reliable chemical and physical properties are key to producing valid modeling results that support forensic determinations (Wait and Maney, 2006). For instnace, Linkov *et al.* (2005) examined how different K_{ow} values for polychlorinated biphenyls (PCBs) may

Table 4.1
Considerations in planning a measurement program.

Analytical	Subsampling
	Preparatory method
	Analytical method
	Matrix/interferences
	Detection limits
	Holding/turnaround times
	Contamination
	QC samples
	Reagents/supplies
	Reporting requirements
Project Details	History
	Waste generation
	Waste handling
	Contaminants
	Fate & transport
	Sources of contamination
	Areas to study
	Adjacent properties
	Exposure pathways
Sampling	Representativeness
	Health and safety
	Logistics
	Sampling approach
	Sampling locations
	Number of samples (field/QA)
	Sample volume
	Compositing
	Containers/equipment
Objectives	Need for program
	Regulations
	Thresholds or standards
	Protection of human health
	Environment protection
	Liability
	Data quality objectives
	Company/agency directives
	Public relations
	End-use of data
Validation & Assessment	Data quality objectives
	Documentation of quality
	Documentation of activities
	Completeness/representativeness
	Bias and precision
	Audits
	Performance evaluation samples
	Chain of custody

have impacted decision alternatives at the Hylebos Waterways Superfund Site in Puget Sound near Tacoma, Washington. Achieving health-based remediation standards ranged in cost from approximately $7.5 million using a PCB log K_{ow} value of 6.0 to $55 million using a log K_{ow} value of 7.2.

Managing errors during the measurement process is key to the generation of accurate and reliable results. When managing errors it is important to realize that there are two general types of errors. Measurable errors are those

factors whose impact on the accuracy, sensitivity, and representativeness of a measurement process can be detected, monitored, and quantified by quality control samples (Taylor, 1987). The proper identification of the analyte being quantified is also essential, and should not be assumed. Nonmeasurable errors are those whose impact cannot be detected by quality control (QC) samples, but can be controlled through quality assurance (QA) programs, standard operating procedures (SOPs), documentation procedures, and training. Unlike measurable errors, which can be detected by QC samples, a nonquantitative and somewhat subjective evaluation by an experienced forensic chemist is necessary to determine if nonmeasurable errors have effected the accuracy and representativeness of a measurement. A more in-depth discussion of quality assurance and quality control practices is provided in Section 4.3.4.

There are eight areas within planning, implementation, and data handling activities where errors occur in the measurement process.

(1) Errors in planning

- **Errors in determining the correct question**
  Forensic investigations begin with a question. Two common analytical chemistry questions involve analyte identification and analyte concentration. Formulating the right question is critical to the success of the measurement process. Many times each participant in a project believes they know what the question is, but often their questions are different. The time to resolve what is the right question is at the beginning of the measurement process. The chosen measurement process may answer only one question at a time. The use of data from one question to answer another typically leads to errors in decision making.
- **Errors in selecting the decision unit**
  A decision unit (population) is the media or portion of the media to which the results of the investigation apply. For the question of identification, judgment may be used to determine the portion of the media likely to provide a clear chemical signature, and this becomes the decision unit. For the question of concentration, the decision unit is the entire volume of material where the analyte concentration is of interest, for example, a pile, drum, or lagoon. In both instances it is likely that there are multiple decision units for a given site.
- **Errors in selecting the correct confidence**
  The confidence with which the resulting data will correctly answer the question must be specified. There are two aspects to confidence: one, the type of confidence and two, the level of confidence. The two types of confidence are statistical confidence and professional confidence. For statistical confidence it is critical that the sampling design incorporate some

type of randomness within the decision unit. For professional confidence that requirement does not exist. The level of confidence drives the level of effort; that is, more confidence requires more effort. For statistical confidence the level is stated in numeric fashion. Some typical levels are 80%, 90%, 95%, and 99%. For professional confidence there are no numeric values, only a subjective opinion based on the experience of the person interpreting the data.

(2) Errors in implementation

- **Errors in sampling the decision unit**

In some instances the entire decision unit fits into the sample container. Sample collection errors occur when the entire decision unit does not fit into the sample container. When the decision unit is larger than the sample container it is imperative that the collected portions (samples) represent the entire decision unit. The common practice of arbitrarily selecting some material from a discrete location within a decision unit most likely will not represent the entire decision unit. To properly represent the entire decision unit, material from many locations within the decision unit should be sampled.

A tenet of sampling theory important to forensic investigations is that two types of heterogeneity should be considered. One is constitutional heterogeneity and the other is distributional heterogeneity (Mishalanie and Ramsey, 1999). These heterogeneities manifest themselves as sampling errors that can be measured and controlled.

Constitutional heterogeneity occurs when the particles in the population to be sampled have different concentrations of the analyte of interest. Constitutional heterogeneity always exists even in the purest of materials. If constitutional heterogeneity could be eliminated, there would be no sampling error since any and every particle that makes up the material to be sampled would have the identical concentration. Since constitutional heterogeneity cannot be eliminated, it must be controlled. Constitutional heterogeneity is controlled with adequate sample mass collected in the field. In the laboratory, constitutional heterogeneity is controlled with either increased subsample mass or grinding of the field sample to reduce particle size (Walsh *et al.*, 2002).

Distributional heterogeneity occurs when the particles are not randomly distributed within the decision unit. Distributional heterogeneity always exists due to the presence of gravity. In addition, distributional heterogeneity occurs whenever the analyte of interest is not uniformly spread across the entire decision unit. If distributional heterogeneity could be eliminated there still remains the error associated with constitutional heterogeneity. Distributional heterogeneity is controlled by collecting adequate sample mass

(calculated to control constitutional heterogeneity) with multiple increments throughout the entire decision unit (known as multi-increment sampling). These concepts are utilized in developing sampling plans for the two common environmental forensic sampling questions: concentration and identification.

- **Errors in preserving the integrity of the sample prior to analysis**
  Once a sample is collected, the integrity of the sample must be maintained. Maintaining integrity includes using inert containers, minimizing reactions, and preventing contamination and losses. Errors from sample handling have many causes and are also functions of an analyte's chemistry and its concentration. These errors typically are controlled using preservatives, appropriate holding times, specified sampling techniques, and specialized clean containers. It is important to know the class of analytes of interest, as well as potential interferents, prior to sampling to maintain sample integrity.
- **Errors in collecting the analytical subsample from the field sample (Subsampling)**
  Subsampling errors are similar to those encountered in the field. The portion of a field sample that the laboratory analyzes is typically smaller than the field sample (Ramsey and Suggs, 2001). When the laboratory selects a portion of the sample for analysis, care must be taken to ensure representiveness of the field sample. Proper communication between the data user and the laboratory is important to minimize representativeness errors. In some circumstances extraneous material may be collected in the field and should be avoided for analysis at the laboratory. However, laboratory personnel should not exclude materials without specific direction from the forensic investigator.
- **Errors in sample analysis**
  Testing includes sample preparation and instrumental analysis, and errors can affect both identification and concentration of an analyte. In addition to commonly identified errors in precision and bias, laboratory errors comprise: (1) data interpretation error, (2) sample management error, (3) laboratory procedure error, (4) methodology error, and (5) data management error (Popek, 2003). Understanding total error associated with all steps in the analysis process is key to producing reliable analytical measurements (e.g., Bonomo *et al.*, 2002; Bruya and Costales, 2005).

(3) Errors in data handling

One of the primary and most critical aspects of data handling is review and interpretation of the quality control data. Often this step is ignored, but a thorough evaluation is integral to data defensibility. If data fail the quality control measures, they may not be suitable for defensible decision making. In some instances failure of certain prespecified quality control parameters does

not affect the quality of data for decision making purposes. Only after careful scrutiny by someone competent in quality control data review can the results be used confidently for decision making purposes. In some cases, statistical analysis may be performed on the data. If statistical analysis is performed, it should also be consistent with the objectives of the study.

4.3 PLANNING

The level of data quality can vary to some extent depending on the objectives of the investigation. Data Quality Objectives (DQOs) include statements about the level of uncertainty that a decision maker is willing to accept. For example, a data quality objective for a site remediation could be "after cleanup, the average concentration of ethylbenzene for an entire hazardous waste site will be less than 100 parts per million (ppm) with a 90 percent level of confidence." In other words, after considering all aspects of the program, the data quality objective is the acceptable probability of making a wrong decision.

Formal procedures for developing data quality objectives, which are summarized in Table 4.2, have been established by both US EPA (2002a) and American Society for Testing Material (ASTM) (1996). Although it is not imperative that a forensic scientist conduct a formal data quality objective evaluation prior to the start of a testing program, it is important that the logic of the data quality objective process be adapted prior to initiating any forensic study if the goal of the investigation is an outcome that is reliable and usable (Maney, 2001; Popek, 2003).

Data quality objectives are often confused with acceptable levels of analytical bias and precision. However, analytical uncertainty is only a portion of an environmental measurement and only one element of an environmental measurement decision. Data quality objectives also should consider the uncertainty in health-based standards, exposure pathways, and sample collection, since they all contribute to the overall uncertainty of a decision (Bayne *et al.*, 2001). Unfortunately, uncertainty can be difficult to measure. Moreover, undefined uncertainty associated with sample collection activities, a prominent element in the measurement process, can be particularly significant

Table 4.2

The seven steps of the data quality objective process.

Step 1	State the problem.
Step 2	Identify the decision.
Step 3	Identify the inputs to the decision.
Step 4	Define the boundaries of the study.
Step 5	Develop tolerable limits on decision error.
Step 6	Specify tolerable limits on decision error.
Step 7	Optimize the design for obtaining data.

(Jenkins *et al.*, 1996; Crumbling *et al.*, 2001). Many researchers suggest that the major contribution to uncertainty of a sample measurement, possibly up to 90%, is that a sample is not representative of study conditions (Jenkins *et al.*, 1996; Crumbling *et al.*, 2001). For example, in one study of aflatoxin measurements in peanuts, error associated with sampling was 67% of the total variance, whereas error associated with the analyst preparing a subsample was 20%, and analytical error was 13% (Whitaker *et al.*, 1974).

The three most important inputs to develop a properly designed study are determining the question, population, and confidence (Mishalanie and Ramsey, 1999; Gilbert and Pulsipher, 2005). Although there are other information requirements associated with the data quality objective process, these three inputs are the technical basis for a defensible plan design. Some of the terms and concepts used in the planning process are discussed in Table 4.3.

4.3.1 QUESTIONS

The purpose of forensic sampling should be clearly defined prior to sample collection. One question, the identification of specific contaminants or ratios of contaminants, can be used to determine its source or the age of the released material. Another question, the concentration of a contaminant within a decision unit, may be used to determine the total amount of material used/discharged, or to pinpoint a source by discerning concentration gradients. Determining the identification and concentration of contaminants is integral to determining the who, what, when, and where questions of an environmental forensic investigation (Stout *et al.*, 1998).

When the objective is to determine the total contaminant mass, the average concentration is the main focus of the sampling effort. Since the mass of a substance cannot be measured directly, the average concentration along with the mass/volume of an area of interest is used to determine the total mass. If determining contaminant mass is the objective, it is important that the average concentration of the material is known within acceptable levels of uncertainty.

4.3.2 DECISION UNIT

Once the forensic question is determined, a decision unit should be identified. The decision unit is the total material to which the analytical results will apply. This could be as small as a fraction of a gram for identification purposes or may be as large as thousands of kilograms. This is a difficult and important step in the planning process. Typically some judgment or knowledge is used to select the decision unit. It could be a stained area, the top phase of a liquid, an individual particle, an entire lagoon, or a waste pile.

Table 4.3

Sampling terms.

Accuracy	This term is defined differently by various users and therefore always should be used with a definition. For some it is used interchangeably with bias as in the phrase "accuracy and precision." Others use it to mean both unbiased and precise. Since nothing is without any bias nor perfectly precise, accuracy generally means minimal bias and imprecision. ISO defines accuracy as the relative agreement between the results of a measurement and a true value of the measured.
Bias	A systematic error that causes the analytical results to be consistently high or consistently low compared to the true value. The true value is never known so great care must be exercised to prevent biases from occurring. Biases occur in all steps of the measurement process. Bias is a component of measurement error.
Blunder	Mistakes made in the measurement process that are not detectable by ordinary quality control events. As a result blunders can be very difficult to find. Some examples include accidentally switching samples in the field, transposing results in a final report, and incorrect documentation.
Colocated Sample	A sample collected next to another sample. This type of sample is sometimes used as a quality control sample for discrete sampling to demonstrate sample representativeness of the decision unit.
Constitutional Heterogeneity	(Also known as compositional heterogeneity) Occurs when different particles within the decision unit have different concentrations of the analyte of interest. As a consequence some of each type of particle (or molecule) must be collected to represent the material in the decision unit. Compositional heterogeneity always exists, even in the purest of materials. Constitutional heterogeneity expresses itself as fundamental error (an error of precision), which must be controlled as part of the sample plan design. The two methods of controlling fundamental error are increased sample/subsample mass (support) and comminution (grinding).
Confidence	Confidence can be either statistical or professional. Statistical confidence is the probability that your results did not occur from pure chance. If the probability that a result could be from chance is small (e.g., less than five percent) then we make the claim that the conclusion is true (e.g., 95% confident that the true mean is less than ten). Professional confidence is not based on statistics and therefore cannot be assigned numerical values. Instead it is based solely on the experience and the knowledge of the person making the claim.
Decision Unit	Decision unit is used interchangeably with population. It is the total material that is to be characterized and about which a decision is being made. It may be a drum, an exposure area, or an individual particle. For most investigations there is more than one decision unit. For instance, a project may involve 100 drums. If the concern is with what is in each of these drums, there would be 100 decision units. Every decision unit may not be sampled, but there are still 100 decision units. If the concentration of lead in a yard is the question, then the entire yard becomes the decision unit. If the objective is identification, the decision unit may be much smaller than for characterization. For instance, if the problem is to determine if a certain particle fell into a yard, it would be the particle that becomes the decision unit, not the yard.
Distributional Heterogeneity	Occurs when the particles are not randomly distributed within the decision unit. This can be both spatially and temporally. In addition segregation is a form of distributional heterogeneity. A common example is segregation of the field sample during the trip to the laboratory. The dense fines end up on the bottom (that is why the common lab subsampling practice of scooping some material off the top of the container for analysis is prone to large errors). In order to mitigate the effects of distributional heterogeneity multiple increments are collected throughout the entire decision unit to make up the sample. This is done both in the field and in the laboratory.
Increment	The portion of material from a population that forms a representative sample of the population.

Table 4.3
Continued

Laboratory Subsample	The portion of material the laboratory selects from the field sample for analysis. The process of removing the material is called laboratory subsampling. If this process is not done with great care, large subsampling errors will result.
Multi-Increment Sample	A sample made from increments from more than one location within the decision unit. Instead of the common practice of filling the container with material from just one location, the container is filled with material from multiple locations. This is done to mitigate the effects of the distributional heterogeneity within the decision unit and provide a more representative sample without the cost of additional samples.
Precision	The degree to which analytical results agree.
Representativeness	A very difficult term that typically relates to the quality of the field sample. A typical definition is "the degree to which the sample(s) reflect the population." The problem with this definition is the manner of determining representativeness if you don't know the true value. A more adequate definition would be a sample(s) that can answer a question about a decision unit with a specified confidence. In other words, if Data Quality Objectives are met, the samples were representative.

If the entire decision unit does not fit into the container, a sample should be collected that represents the entire decision unit. It is the responsibility of the forensic investigator to develop a sampling plan that results in a sampling error small enough to meet the project objectives.

When the objective is to determine the identity of the analyte, the main goal is to collect samples in the decision unit that provide a clear pattern of the material in question. One should not convolute the chemical signature with material that is not representative of the decision unit. In some cases a sample that represents the average concentration of the analyte within the decision unit also provides the best signature. Judgment, based on defensible science or adequate field screening methods, may be used to determine the portion of the population that provides the best signature. However, it should be noted that for field screening, the rules for defensibility of field measurements are no different than for laboratory measurements.

When the objective is to determine the average concentration of the analyte, the main goal is to collect a sample that represents the entire decision unit. If the sample does not represent the entire decision unit, data assessment is confounded, and incorrect or indefensible decisions may result. Determination of the area/mass/volume of the decision unit is not an easy task. The boundaries of the decision unit should be well defined and agreed upon by all relevant parties prior to sample collection.

4.3.3 CONFIDENCE

Once the question and decision unit have been determined, the desired type and level of confidence should be specified. The confidence is a measure of

how accurately the analytical results represent the population (Ramsey and Hewitt, 2005). Confidence can be statistical confidence (e.g., 95% confidence) or can be based on professional judgment. If confidence is based on statistics, there are certain sampling requirements that must be followed. If confidence is based on judgment, then the expertise of the person developing the plan is key. For legal situations, judgmental sampling may be counter to standard performance criteria stipulated in the Daubert ruling (USSC, 1993) since there is no method to measure sampling. With judgmental sampling, confidence is not independent of the sampler and determination of sample representativeness can become a battle of experts. Although judgment samples may be appropriate for a specific situation, its limitations should be considered before implementation.

Confidence in a decision is a function of all the errors in the entire measurement process. Larger measurement errors produce less reliable decisions, or in the case of statistical analysis, larger confidence intervals. If measurement error can be minimized, more confident decisions can be made (lower probability of mistakes) and fewer samples need to be collected. The amount of error that can be tolerated is a function of how close the average is to a limit or threshold. The closer the average is to the limit, the less error can be tolerated. For example, if the average is 8 and the limit is 10, the total error must be less than 25% to reliably conclude that the average is less than 10. If the limit is 16, an error of 100% would be acceptable to conclude that the average is less than 16.

4.3.4 SAMPLE PLAN DESIGN

Once data quality objectives are determined, the sampling plan can be designed and optimized (ASTM, 1998). If the US EPA DQO process is followed exactly, this would be step 7 of the process (see Table 4.2).

When dealing with particulate material, a scientifically based, defensible sampling plan using field sampling theory is important. Field sampling theory was first developed in the mining industry, but its application has become more widespread in the environmental community (Gerlach *et al.*, 2002). Sampling theory is critical to the determination of sampling error, which often is the major component of measurement error.

If sampling error is not controlled, the resulting data are questionable. Some of the consequences of uncontrolled sampling error are:

- Increased variability (imprecision) of analytical results
- Underestimation (bias) of the mean
- Higher failure rate of quality control

Variability confounds decision making. The higher the variability, the more difficult it is to make reliable decisions. Although some variability is inevitable, it must be controlled at a level that allows effective decision making. Concerns about variability are convincingly demonstrated for compliance testing of solid waste materials regulated in the Netherlands, where different types of frequency distributions and outliers occur regularly (Stelling and Sjerps, 2005). Bias is an even greater concern since sampling bias cannot be measured with traditional quality control techniques. If the magnitude of the subsampling variability and bias is not controlled, the resulting conclusions may be questionable even if they are correct.

4.3.4.1 Sampling for Identification

There are five potential areas where sampling errors can occur. In addition, there are also errors of analysis and data interpretation. Some of these errors are the same for both identification and quantification.

(1) Selecting the decision unit

Judgment is used to determine where the population of concern is likely to provide a clear chemical signature. For instance, a decision unit may be an office that possibly has been contaminated with anthrax. If the objective is to find some anthrax for identification purposes only, judgment can be applied to the investigation (e.g., anthrax may be on the computer screen due to the presence of static electricity). The size of the decision unit can be as small or as large as necessary to provide a clear chemical signature.

(2) Selecting the confidence

The type and level of confidence must also be selected before the sampling plan can be designed and implemented. If statistical confidence is chosen, the sample design must incorporate some type of random selection. Statistical confidence is probably used less often for identification than for determining concentration.

(3) Collecting the sample from the decision unit

It is difficult to fully describe sampling for identification since the issue is site specific and a function of known and unknown inputs, chemistry, and fate and transport, all of which may be represented by a conceptual model. The goal is to identify an area within the area of concern based on judgment and science, and then collect a sample without affecting the integrity of the contaminant. The size of the decision unit for identification can be much smaller than the size of the decision unit for concentration since the presence of any signature is typically adequate. Sampling designed for identification purposes typically cannot be used to establish concentration within the decision unit. In cases where the decision unit is difficult to locate (i.e., hot spot sampling), field screening analysis may be helpful.

(4) Preserving the integrity of the sample prior to analysis

Sample preservation has been previously discussed. However, it is important to note with identification that other analytes critical to the investigation and their integrity must also be preserved (i.e., co-contaminants, tracer analytes). In some instances these other analytes may not be known at the time of collection.

(5) Collecting the analytical subsample from the field sample

Laboratory subsampling should be designed to collect the portion of material in the field sample that provides the most reliable signature (Ramsey and Suggs, 2001). It is possible that extraneous material may have inadvertently been collected that should be avoided during subsampling. It is also possible that phases of the sample may have to be separated or mixed, or reactions may have occurred during transport to the laboratory. Communication between field investigators and laboratory personnel is important to obtain subsampled material that provides a reliable signature.

4.3.4.2 Sampling to Determine Concentration

(1) Selecting the decision unit

For many regulatory programs decision units are defined (e.g., a management unit under US EPA's Resource Conservation and Recovery Act (RCRA)), but for forensic sampling the investigator will need to define the units. Factors that determine a decision unit may include regulatory requirements and health concerns. For determining total mass, the decision unit is ideally the entire area where the analyte exists. Location of the decision unit may not be completely known and professional judgment or models may be needed. If modeling is used, similar levels of scientific rigor as with measurement processes are needed to be reliable (Denton *et al.*, 2006). There may be several decision units needed to properly ascertain concentration.

(2) Selecting the confidence

Before a sampling plan can be designed and implemented, the type and level of confidence must be selected. If statistical confidence is determined, the sampling plan design must incorporate some type of random selection. Statistical confidence is used more often for determining concentration rather than for identification. For concentration, the estimate of the error in the final result is important. The level of confidence chosen is related to the consequences of an incorrect decision. More significant consequences should be associated with higher confidence levels.

(3) Collecting the sample from the decision unit

Once the decision unit is defined, it is critical that a representative sample capture the true mean concentration. If the decision unit does not entirely fit into the sample container (a typical case), then the sample must be collected in a manner that minimizes sampling error. The commonly employed method of

taking a few discrete samples typically does not work well when trying to estimate average concentrations. Unfortunately there is very little guidance on reducing sampling error during environmental sampling (Gerlach *et al.*, 2002). There is a field of sampling theory, cited earlier, that is very helpful for designing scientifically valid sampling plans for particulate matter. This theory is slowly being integrated into the development of environmental studies (Gerlach *et al.*, 2002). Defensible sampling necessitates that sample mass and the number of increments be considered. The details are beyond the scope of this chapter, but can be found in Gy (1992, 1998), Pitard (1993), US EPA (1992a), and other sources (e.g., Maney, 2001; Jenkins *et al.*, 2005).

(4) Preserving the integrity of the sample prior to analysis
Error associated with sample integrity is independent of the question of the investigation. Care should always be taken to maintain the integrity of the sample. Ample guidance is available about preservation methods that maintain sample integrity (e.g., Smith, 2003; US EPA, 2004, 2005).

(5) Collecting the analytical subsample from the field sample
Subsampling in the laboratory should be performed in a manner that provides the best estimate of the average concentration of the field sample (Ramsey and Suggs, 2001). A common laboratory practice is to scoop a few grams of material from the top of the sample jar for analysis. This is not a good method for subsampling. In most cases the concentration of the analyte in the top and bottom portions of the field sample are different. This is not only a function of how the container was filled, but also settling and phase separation. It is also possible that extraneous material may have inadvertently been collected that should be avoided in the subsampling portion. If extraneous material is not considered in the subsampling protocols, the resulting concentration data may be suspect.

4.3.4.3 Quality Control Measurements

Quality control samples provide a mechanism to identify, measure, and monitor errors. Table 4.4 lists the types of information that can be gleaned from different quality control samples, and some of these samples are described in more detail here.

Trip Blanks—Sample containers prepared at a laboratory or other off-site area that are not opened, but travel with the samples that are collected. Trip blanks are analyzed in the laboratory with the field samples and provide information to determine whether the sample containers were clean and if cross contamination from outside sources exist. The typical rate for using trip blanks is one per 10 or 20 field samples. These samples are most beneficial for low-level analytes or situations where cross contamination is a concern.

Table 4.4

Information derived from quality control samples.

Quality Control Sample Type	Types of Error										
	Contamination					Precision				Bias	
	Containers	Field	Equipment	Cross Contamin.	Laboratory	Sampling	Splitting	Subsampling	Analysis	Laboratory-preparation	Laboratory-analysis
Blanks											
Trip	X			X	X						
Field	X	X		X	X						
Equipment	X	X	X	X	X						
Method					X						
Replicates/Splits											
Splits (Field)							X	X	X		
Replicate (Field)						X		X	X		
Splits (Lab)							X	X	X		
Replicate (Lab)						X		X	X		
Spikes											
Field										X	X
Matrix (Lab)										X	X
Performance Sample											
Performance Evaluation Standards										X	X
Reference Material										X	X

Field Blanks—Empty sample containers prepared like trip blanks are opened in the field at one of the sample locations, and then analyzed with the field samples. Field blanks provide information regarding contamination of field samples that include the time period when the container was open. An example where field blanks would be useful would be collecting low-level organics in an area where heavy equipment is operating. Field blanks do not provide information related to the representativeness of the samples collected.

Decontamination Check Blanks—Samples of water collected in the field from the rinsing of sampling equipment. In some cases equipment is used several times and must be cleaned between uses. After the sample tool is cleaned it is rinsed with clean water, and this sample is analyzed to determine if the sampling tool was clean. The typical rate for using decontamination check samples is one per 10 or 20 field samples. The need for decontamination check samples can be avoided with the use of disposable/dedicated samplers.

Splits—Field samples that are split into two or more samples to allow for separate analyses. This typically is done to check or verify one laboratory's results against another. Splits are used to check laboratory performance, not field performance.

When collecting split samples it is imperative that the split samples are identical in composition. If not, disagreement of results may be due to splitting error, not laboratory error. There is no practical way to determine these discrepancies afterward. Proper splitting of particulate samples, multiphasic liquids, and liquid samples with suspended material is difficult (e.g., Pitard, 1993). Liquids with suspended materials can be separated with a churn splitter (USGS, 1997).

Replicates—Extra field samples (e.g., duplicate, triplicates) collected under the exact same conditions as the primary samples. These samples are used to determine sampling error, which in turn indicates sample representativeness. In some cases a colocated sample is collected for the replicate. However, colocated soil samples provide little, if any information related to the representativeness of the samples. If statistical analyses are to be performed, triplicate samples are preferred over duplicates.

Individual quality control events often measure more than one type of error. For instance, field replicates measure precision in the entire measurement process, including both field precision and laboratory precision. This is beneficial because the precision of the entire measurement process is contained in one quality control event. Therefore, the field replicate sample is the most important quality control sample and, if possible, should be included in every sampling campaign. To estimate the precision of the sampling process alone, the laboratory precision has to be subtracted from the field replicate precision value.

4.4 IMPLEMENTATION

4.4.1 PLAN IMPLEMENTATION

Forensic investigators need to be adept at examining various materials indicative of contaminant sources. Sampling plans must consider the media that are to be sampled, as well as the types of contamination that may be present. Advice for sampling matrices pertinent to forensic investigations follows.

4.4.1.1 Soil Gas

Soil gas measurements are suited to identifying contaminants and sources. Soil gas surveys provide a screening method for detecting volatile compounds so that subsequent investigative activities can be focused with the highest probability of contaminant detection (Marrin and Kerfoot, 1988). Technology used for soil gas measurements is most effective in detecting compounds with low molecular weights, high vapor pressure, and low aqueous solubilities. Soil gas surveys often are used to locate nonaqueous phase liquids (NAPLs) (US

EPA, 1994). NAPLs are organic liquids or wastes that are sufficiently immiscible in water such that they may persist as a separate phase in the subsurface for many years.

There are many techniques available for soil gas analysis: (1) active, (2) passive, and (3) flux chamber (Hartman, 2002). An active soil gas survey consists of the withdrawal of a soil gas sample, typically from a perforated sampling probe, followed by analysis in a stationary or mobile laboratory (Hartman, 2002). Passive soil gas surveys provide for the burial of probes containing absorbent materials to identify the presence of a broad range of volatile and semivolatile organic compounds (US EPA, 1998a,b). Flux chamber measurements are used primarily in research applications, risk assessments, and toxicological studies when direct vapor flux values are desired (ASTM, 1997; Hartman, 2003). Soil gas testing technology is constantly evolving; hence a review of the current literature is important whenever technology of this type is employed. Since identification is the major objective of soil gas analysis, it is important that technology used to identify compounds is appropriate. Typical analytical issues include selectivity, detection limits, and background or cross contamination.

If soil gas is used to determine concentration or mass as is the case with most passive techniques, the ability of the technology to quantitatively remove the soil gas of interest from a specified volume must be well understood. This is difficult to determine due to the heterogeneity of the soil matrix. In addition, it is difficult to determine the volume of soil gas removed. The use of soil gas to determine the total mass of an analyte should be undertaken with extreme caution; however, Hewitt (1999a) has shown that using a robust and controlled vapor collection procedure may produce trichloroethene soil gas measurements that are quantitatively reliable.

4.4.1.2 Soil

The definition of soil is critical to an investigation. In fact, the definition and description of earth materials such as soil can itself be an important forensic tool (Murray and Tedrow, 1992). Scientific disciplines and agencies define soil differently (e.g., US EPA, 1991a). This is not helpful when measuring analyte concentrations in soil. The easiest method is to select a cutoff particle size and assume that everything under that size is soil. Whatever particle size is chosen, it should not be decided cavalierly since concentration is often related to particle size. Another concern is the presence of organic matter and how it is considered as part of the soil matrix. Many contaminants sorb onto organic matter, and to exclude the organic material for analytical convenience may omit a significant portion of the contamination. Total organic carbon (TOC) measurements can be a useful indicator of the relative organic content of the

soil being studied. Care should be exercised when sampling soil surfaces, especially if atmospheric deposition of the contaminant is a concern. The correct geometry of the soil increment collected should be cylindrical (Pitard, 1993). There are a variety of soil coring devices that collect cylindrical soil samples rather than pointed trowels (Walsh, 2004).

The soil definition should be decided as part of the data quality objective process. Once decided, sieving may be used if the target analytes are not volatile. Sieving can be conducted either in the field or at the laboratory depending, in part, on the ease of work, health and safety issues, and shipping.

Soil is frequenthy heterogeneous. Care must be exercised to mitigate the physical (compositional) and spatial (distributional) heterogeneities of the soil when determining the mass and the number of increments needed to represent the decision unit. Obviously, it is easier to sample surfaces than subsurface soils due to accessibility.

One challenge to collecting representative soil samples arises from the small-scale spatial variation seen in contaminant concentration resulting from NAPL sources (Kram *et al.*, 2001; Feenstra, 2005). NAPLs include, in part, petroleum products such as gasoline and fuel oils, creosote, coal tar, polychlorinated biphenyls (PCBs), and chlorinated solvents such as trichloroethylene (TCE). Petroleum products are less dense than water and are known as light nonaqueous phase liquids (LNAPLs), whereas creosote, coal tar, PCBs, and chlorinated solvents are denser than water and are referred to as dense nonaqueous phase liquids (DNAPLs) (Weiner, 2000). Lack of knowledge about the presence and precise location of NAPL in a subsurface zone can confound an investigation. Feenstra (2005) believes "this challenge can be overcome for characterization of small-scale properties of the source zone by careful selection of sampling locations based on soil stratigraphy and field screening, and for larger-scale properties of the source zone by the collection and analysis of large-quantity soil samples and composite samples."

Another challenge is collecting representative soil samples where volatile organic compounds, such as solvents, are target analytes. Due to losses by volatilization and biodegradation, laboratory data can significantly underestimate actual concentrations of volatile organic compounds if attention is not given to sampling and analysis techniques (Smith *et al.*, 1996; Hewitt, 1999b). In the field, losses may occur during sample transfer from the sampling device to a sample container, during sample transport and storage as a result of biodegradation, abiotic degradation, or nongas tight samplers and from laboratory handling prior to analysis (Siegrist and Jenssen, 1990; Hewitt, 1994). Numerous preservation techniques, such as methanol extraction, freezing, or chemical preservation have been developed to minimize volatile organic losses (e.g., Hewitt, 1997, 1999c).

Determining background concentrations of contaminants in soil, as well as other matrices, may be an important element in determining the source and impact of the contamination (Qian and Lyons, 2006). For instance, polycyclic aromatic hydrocarbons (PAHs) are ubiquitous in the environment, originating from various natural and anthropogenic sources. These sources include, in part, diagenesis of organic matter, atmospheric deposition from burning petroleum, vegetation and coal, surface runoff, urban and industrial outfalls, and petroleum products (Blumer, 1976; Laflamme and Hites, 1978; Lima *et al.*, 2005; Boehm, 2006). For example, the compositional similarity of degraded coal tar PAHs to combustion-related PAHs typical of urban background can make it difficult to determine local PAH background concentrations in soils and sediments near former manufactured gas plant (MGP) sites (e.g., Costa *et al.*, 2004; Stout and Wasielewski, 2004). Since representativeness is fundamental to a forensic investigation, background must be properly considered to ensure results used for decision making are reliable and admissible (Maney and Wait, 2005). For instance, background was a key issue in *Dodge v. Cotter* (328 F.3d 1212, 10thCir. [2003]), where a successful appeal of tort claims associated with contamination from a uranium mine occurred because the expert geologist's conclusion was based on a general text for establishing backgrounds for metals that was not reflective of the mineral's diverse and rich background in the study area.

4.4.1.3 Sediment

Sediments include depositional silts, clays, fine sands, and small particulate matter that are found below an aqueous layer, such as a pond, lake, river, stream, or wetland. Sediment samples include interstitial (pore) water, but not standing water collected with the sediment sample. Most US EPA test methods used for sediment analysis originally were designed for soil samples (US EPA, 1998c). Soil test methods analyze a known wet weight of sample, then back-calculate the results to a dry weight basis. Since soil samples typically contain less than 10% moisture, whereas surface sediment samples may contain up to 80% moisture, the amount of sample analyzed is key to realizing method sensitivity requirements (US EPA, 1998c). This issue affects detection limits and representativeness, and must be considered during project planning.

As with soil, it is important to classify sediment particle size. Sediment particle size is a major factor that determines the capacity of sediments to bind pollutants (Mudge *et al.*, 2003). In addition, with flowing streams, sediment characteristics can change with the velocity of the stream, which, in turn, can be affected by varying stream widths and seasonal variations of input. The physical, spatial, and temporal heterogeneity of the sediment particles should also be considered when developing a sampling plan. Particle size

normalization procedures may be needed to understand differences in contaminant concentrations. TOC is more routinely measured for sediments than for soils, and can be an important factor in biochemical and contaminant partitioning studies (Boehm et al., 2002; Ryba and Burgess, 2002).

4.4.1.4 Groundwater

The most important issue in groundwater sampling is sample representativeness or the accuracy with which a groundwater sample represents the formation water. A myriad of considerations arise when interpreting chemical results from groundwater samples for source identification and/or use in groundwater modeling (Morrison, 2000). Examination of chemical results without exploring other factors can result in the misinterpretation of the data relative to identifying the contaminant source(s). If groundwater samples are collected from a monitoring well, common issues to evaluate include well design, location, depth and screen interval, water level management, well purging, and sample collection procedures. Potential sources of bias include well construction materials, cement grout, improperly installed security covers, and contamination introduced during drilling activities.

Groundwater collection procedures should be designed and implemented to retrieve representative samples. Sufficient removal of stagnant water, followed by well stabilization are key procedures. Parameters used to monitor well stabilization include pH, temperature, conductivity, oxidation-reduction potential, dissolved oxygen, and turbidity. Selection of a sampling technique should be decided as part of the planning process. Techniques to be considered include (1) conventional purging and sampling using a pump or a bailer, (2) low flow minimal draw down or micro-purge technique, (3) direct push techniques, and (4) passive diffusion bag samplers (Popek, 2003). If understanding colloidal transport or dissolved metal content are objectives of the study, filtering procedures will need to be defined as part of the planning process.

It is generally assumed that groundwater is representative of the aquifer if the well is properly purged and the chemistry of the water does not change during storage and transport from the well to the laboratory. There are times, however, when well purging is not appropriate. For example, the proper approach for collecting nonaqueous phase liquids (NAPL) is to sample NAPL prior to purging the well (Mercer and Cohen, 1990; Johnson, 1992).

4.4.1.5 Surface Water

Surface water is water that flows or rests on land and is open to the atmosphere, such as lakes, ponds, lagoons, rivers, streams, ditches, and man-made impoundments. When examining a surface water column, there needs to be

a focus on assessing both physical and chemical characteristics of the water (i.e., suspended solids, dissolved contaminants), and then also relating these characteristics to hydrologic conditions, sources, and transport (Shelton, 1994). There are two complicating factors when sampling surface water. One factor is whether suspended material should be considered as part of the sample. For identification purposes, collection of suspended material may not be a problem. For determining concentration, collecting a representative sample containing suspended material is a difficult problem. The best method for representing suspended material in flowing surface water is using an integrating sampler (USGS, 2005).

The other complicating factor is the temporal nature of surface water. Analyte concentrations in surface water can change rapidly over short periods of time. Sampling plans should consider temporal heterogeneity.

4.4.1.6 Surfaces

Most studies evaluating surface contamination focus on identifying contaminants for source identification, risk assessment, and determining the extent of contamination for the purpose of remediation. Published wipe sampling procedures provide semiquantitative data for nonporous surfaces such as metal (e.g., Gaborek *et al.*, 2001), but are considered poor for porous surfaces such as concrete (Slayton *et al.*, 1998). For many contaminants, nonporous surface sampling can be confidently performed using wipe sampling procedures to identify the contaminants, but for quantification these types of procedures may be useful only for mass loading determinations (e.g., mass per unit area), but not to determine concentration (e.g., mass per unit mass) (Robbins, 1980; US EPA, 1989; ASTM, 2001). Many protocols and sampling plans do not properly capture spatial variability on surfaces. Sampling rough or porous surfaces such as upholstery, carpeting, concrete, and bare wood is a more challenging endeavor, requiring either specialized wipe sampling methods, micro-vacuum techniques, or bulk sampling (EPA, 1991b; Slayton *et al.*, 1998; ASTM, 2005).

4.4.1.7 Ambient Air

Outdoor ambient air presents difficult sampling problems. Spatial and temporal heterogeneity is potentially enormous. There are few air samplers or sample collection techniques that can address this type of variability, such as those defined for particle mass (PM) regulated under National Ambient Air Quality Standards (NAAQS) (US EPA, 1997). For identification purposes the goal is to place a sampler in the area where the contamination is most likely. If the analyte is collected correctly and identified properly, the objective of the sampling is achieved. For concentration, outdoor ambient air data may

not be representative of the place or time of interest. There is no efficient method to collect ambient air samples to determine the concentration of outdoor air beyond what was present at a specific point over a specified time period. Spatial and temporal heterogeneities may be too great to produce reliable measurements unless a long-term continuous sampler is used. Indoor air contaminants are easier to determine if the sampling device has access to the entire decision unit. Either moving the sampling device to a strategic location or moving the air within the decision unit can be effective. Air samplers are typically compound specific so the correct type of sampler must be chosen. In addition, continuous monitors often do not have the required sensitivity to measure toxic air pollutants at very low concentrations. The collection of a representative air sample also requires attention to other parameters, such as temperature, pressure, humidity, sampling rate, and volume collected (Wight, 1994).

4.4.2 SAMPLE HANDLING

In the United States, regulatory standards exist that specify the proper containers for soil, water, and air samples for a given test method. If the standards are not followed, the containers can introduce a potential bias, especially near detection limits. Samples with low concentrations of target analytes may require preservation by refrigeration or the addition of chemical reagents to slow physical, chemical, or biological changes. Chemical and biological activities altering sample chemistry include the formation of metal and organic complexes, adsorption/desorption reactions, acid-base reactions, precipitation/dissolution reactions, and microbiological activities affecting disposition of metals, anions, and organic molecules. Forensic investigators should conduct due diligence on sample handling requirements as part of their data quality objective effort that reflects both sound science and the vagaries of applicable regulations. For example, a survey of soil sample handling procedures of state pesticide regulatory agencies showed 10 states use some sort of bag for sample storage, 30 states use glass jars, two states use either bags or jars, and four states use other types of storage containers such as plastic tubes, aluminum foil, or boxes (Saxton and Engel, 2005).

Although sample handling may violate various regulatory sampling/handling requirements or guidance, sample results should not be necessarily disregarded (Simmons, 1999). In the United States, such data may be relied upon by an expert witness. In *People v. Hale* (59 Cal. App. 3d [1994]), for example, an element of the case involved the illegal disposal of 1,1,1-trichloroethane (TCA) into waste dumpsters. Although sampling procedures (no sampling plan was used), handling (samples were frozen rather than cooled to 4°C),

holding times (the specified 14-day holding time exceeded), and analytical methods (Method 8015 with a flame ionization detector was used instead of accepted EPA Methods 8010 or 8240 using a mass selective detector) deviated from EPA SW-846 standards (US EPA, 1986), the court found that these deviations did not preclude the introduction of this information as evidence.

Another example of the value of chemical testing information that violated recommended sampling and/or handling procedures is illustrated in the matter of *People v. Sangani* (94 C.D.O.S. 1273 [1994]), which involved the illegal disposal of hazardous waste into a sewer system. The defendant was convicted, but appealed because the laboratory performing the analysis was not certified. The court found that "Failure to follow precise regulatory or statutory requirements for laboratory tests generally does not render the test results inadmissible, provided the foundation requirements for establishing the reliability of the tests are met." The necessary foundation requirements are: (1) the testing apparatus is in proper working order, (2) the test was properly administered, and (3) the operator was competent and qualified. If questionable data can be properly recreated and shown to be reliable, the axiom that all forensic data may have value is important to consider since the assumption that a court will find the analytical data inadmissible may be misplaced.

4.4.3 SUBSAMPLING

Typically the mass of a field sample is much greater than the mass required for analysis. Therefore, as part of the analytical process, the laboratory removes the mass required for the analytical technique from the field sample. This process is called laboratory subsampling. If the subsampling is not performed properly the quality of the analytical results will be compromised. For example, US EPA contract researchers evaluated the effectiveness of five different soil subsampling practices (riffle splitting, paper cone riffling, fractional shoveling, coning and quartering, and grab sampling) and found that subsampling accuracy was at least two orders of magnitude worse than the accuracy of the analytical method (Gerlach *et al.*, 2002).

Historically, laboratory subsampling and the errors associated with subsampling have received little attention. For solid materials a common approach is to open the field sample container and scoop a few grams off the top for analysis. In some cases, the top portion of the material is stirred with a spatula prior to subsampling. It is questionable if either approach will produce data of sufficient quality for environmental decision making. In addition, improper or inadequate quality control will confound one's ability to determine the magnitude of subsampling errors.

Laboratory subsampling should be designed to provide the best estimate of the average concentration of the field sample. There are exceptions when only a portion of the field sample is represented in the laboratory. Some examples are the presence of extraneous material, isolating a single phase of a multiphase liquid, or filtering suspended material.

As with the collection of field samples, the heterogeneities inherent in the media must be controlled in the laboratory. The same sampling theory used to measure and control field sampling error should also be used in the laboratory (US EPA, 2000a). Techniques, such as grinding, can be effective in reducing sampling error (Walsh *et al.*, 2002).

4.4.4 ANALYSIS

Regulatory test methods typically are used to investigate the presence of contamination, to characterize human and ecological risks, to determine regulatory compliance, to aid in the design and evaluation of remediation alternatives, and to monitor cleanup and disposal activities. Although these types of target analyte methods may be useful to today's forensic chemist, they may be inadequate to determine the source, fate, and transport chemistry being investigated. In effect, regulatory target analytes, considered rather ubiquitous in the environment, often do not satisfy the unique distribution attributes sought by the forensic chemist. Forensic chemists often are required to devise an analytical chemistry program focused on unique marker compounds and/or a mixture's "chemical profile" to develop linkages between contaminants and contributing parties (Morrison and Murphy, 2006). As an example of a marker compound, the additive methyl tert-butyl ether (MTBE), an environmental contaminant of recent notoriety, has been used as an octane enhancer in gasoline since the late 1970s (Davidson and Creek, 2000). Forensic chemists have used this marker compound to differentiate modern gasoline contamination from older gasoline formulations (Petrisor, 2006; Stout *et al.*, 2006). Another example is biomarkers, stable complex molecules derived from formerly living organisms, that are often useful in determining the source of petroleum contamination (Peters *et al.*, 2005; Wang *et al.*, 2006).

There is an extensive literature available that will aid a forensic investigation in determining whether a standard test method exists for their purpose (e.g., Suggs *et al.*, 2002; Smith, 2003; US EPA, 2003a). Of particular note is the National Environmental Methods Index (NEMI), first available in 2002, which is an online database of water method summaries that provide information to compare one method with another for the purpose of method selection (Keith *et al.*, 2005; available at http://www.nemi.gov). When forensic chemists

are confronted with a situation for which a standard method is not available, it is often prudent to modify a proven or standard method rather than start anew (Wait and Douglas, 1995). Novel methods are often fraught with unforeseen problems that are not amenable to problem-solving research within the schedule and budget of a forensic study. The resultant method's performance should then be shown to be adequate for matrices and analytes of interest using, in part, method detection limit studies (Currie, 1999; Georgian and Osborn, 2001; Douglas *et al.*, 2004), and if possible, compared to other test methods (*e.g.*, ASTM, 1991; Swartz and Krull, 2006).

Once a method meets all analytical data quality objectives, a QA/QC system must be instituted that will define, demonstrate, and document method performance. Bias, precision, representativeness, and sensitivity are parameters that should be measured in evaluating the suitability of the method. The manner in which these parameters are measured and evaluated for nonstandard methods may vary from standard methods. However, it is advisable to start with measures and acceptable criteria of similar standard methods and then revise these criteria, if necessary, during the method optimization process (Wait and Douglas, 1995).

Two outputs of a test method are identification of an analyte and determining the concentration of that analyte. The reliability of both outputs should not be assumed. There are various types of instruments available to analyze for substances. Most environmental contaminants are organic compounds, and the predominant instruments of choice to analyze for organics use chromatography (Wait, 2000). A chromatography method may provide reliable identification of a compound if the method is chosen, designed, and implemented properly. For example, benzene, a contaminant commonly associated with gasoline spills, can be analyzed on a gas chromatograph (GC) coupled to either a photoionization detector (PID) or a mass spectrometer (MS). The PID is only a quasi-selective detector, hence its output can be prone to interferences when a GC/PID method is used. Some examples of benzene interferants encountered using a GC/PID method are shown in Table 4.5 (Smith, 2000).

The same type of coelution problem exists for MTBE. Lawrence Livermore National Laboratory conducted an extensive method evaluation study of commonly used GC methods for oxygenates, such as MTBE, and found the following:

> "EPA Method 8020A/21B (photoionization detection) was unfit for monitoring of tert-butyl alcohol and frequently yielded false-positives (12–15% of samples) and inaccurate results when ether oxygenates **[including MTBE]** were monitored in aqueous samples containing high concentrations of TPH **[total petroleum hydrocarbons]** (<1000 µg/l). Thus, care

Potential Benzene GC/PID Interferences	
Methyl acrylate	Dichlorohexadione
Cyclohexene	Hexyne
Dichlorocyclohexene	Ethylisobutylether
Dimethylisobutylamine	Hexadiene
Cyclohexadiene	Pyran
Bromocyclohexene	Ethylchloromethyl ether
Diisopropylamine	Bromochloroethene
2,3-Dichloroethane	Hexatriene
Ethyl acetate	3-Methyl-1-hexene
Methylhexylether	Dichloropropionaldehyde
Isocyanoethane	1,2-Dimethoxyethane
Nitroethane	2-Methyltetrahydrofuran

Table 4.5
Partial listing of compounds that can coelute with Benzene using a GC/PID.

1. Request that the laboratory confirm all chromatography results with at least a second GC column or a second detector.
2. Whenever appropriate with compound chemistry and range of concentrations, request confirmation be performed by GC/MS.
3. Give preference to GC/MS methods for forensic investigations since unknown sites will be more reliably characterized.
4. For long-term monitoring studies, occasionally confirm GC analyses with GC/MS.
5. Give preference to chromatography methods that use selective detectors. For example, for PAHs choose EPA Methods 8310 and 8270 over EPA Method 8100 (US EPA, 1986).
6. For PAHs requiring low detection limits, select EPA Method 8270 with selected ion monitoring (SIM) rather than EPA Method 8310 (US EPA, 1986).

Table 4.6
Chromatography tips.

> should be taken in the analysis of LUST **[leaking underground storage tank]** databases populated with EPA Method 8020/21 data because results reported for MTBE in samples containing high samples of TPH have a likelihood of being inaccurate or false-positive."
>
> (Halden *et al.*, 2001).

Selecting appropriate instrumentation and test methods should be part of a data quality objective evaluation. Some tips about selecting chromatography methods are provided in Table 4.6.

Reliable concentration data may have a wide range of uses by forensic investigators, from calculating the total mass of certain contaminants for use in remediation cost allocation, to determining diagnostic ratios of certain contaminants indicative of sources. Diagnostic ratio determination can be particularly prone to error caused by analytical variability. For example, ratios of certain PAHs may indicate sources of contamination at manufactured gas plant sites, such as by-product tars (Costa *et al.*, 2004). Carburetted water gas tars typically have fluoranthene/pyrene (Fl/Py) ratios between about 0.5 and

0.9, and coal carbonization and oil gas tars have Fl/Py ratios greater than 1.0 (EPRI, 2000). However, systematic differences in compound ratios due to the analytical process can be misinterpreted as differences in source. Commercial laboratories often produce PAH data that when used in diagnostic ratio determinations are biased due to the method by which each PAH concentration is determined. The concentration of fluoranthene usually is determined relative to the response of the internal standard, phenanthrene-d_{10}, whereas pyrene usually is determined using the internal standard chrysene-d_{12} (EPRI, 2000). Sample-to-sample differences in the relative recovery of these two internal standards can affect the concentrations of fluoranthene and pyrene, and ultimately their ratio. Hence, it is important to assure that variations in diagnostic ratios are not caused by analytical variability.

Laboratory selection is critical to a successful testing program. Equally important, the laboratory must be staffed with experienced senior chemists who can help direct an iterative chemistry program down a pathway that will help the forensic investigator solve their problem (Mishalanie and Ramsey, 1999; Wait and Cook, 1999; Schultz, 2001). As such, parties selecting a laboratory should examine:

- Technical experience
- Technical capabilities and instrumentation
- QA and QC protocols
- Standard operating procedures (SOPs)
- Document control systems
- Automated laboratory practices and electronic record keeping
- Certification, accreditations, audit results
- Training programs
- Ethics policies and practices
- References

Unfortunately, those who contract for laboratory services often naively assume that data produced by environmental laboratories are impeccable. Although there are reasonable steps that the purchaser of laboratory services can implement to minimize the generation of unreliable data, such as on-site laboratory audits and due diligence evaluation of the laboratory's prior performance, there are no guarantees to data quality and integrity, particularly if, as happens on occasion, the laboratory is devious. Accounts of data fraud abound (Wait, 2002; US EPA, 2006). One case that received a lot of press involved the *United States v. Hess Environmental Laboratories* (E.D. Pa. 97-531 [1997]), where a Pennsylvania environmental laboratory admitted defrauding hundreds of clients who were billed over $2.1 million for tests, which were falsified or never

performed. The laboratory paid over $5.5 million in fines as part of its guilty plea for defrauding customers and violating the Clean Water Act.

The reasons for fraud are varied, but often are related to a laboratory's financial situation. Other reasons for misconduct include penalties that are imposed based upon contractual report due dates, sample analysis hold time requirements, poor communication, inadequate training of staff, and questionable management practices (Wait, 2002). The resulting types of fraudulent activities often include bench sheet modifications, alteration of instrument response to an analyte, selective exclusion of data, and complete fabrication of data. Obviously courts should forbid an expert from offering opinions based on a "fictitious set of facts," as noted in *Guillory v. Domtar Industries Inc.* (95 F.3d 1320, 1331, 5thCir.[1996]). Forensic investigators should enact more proactive warnings to laboratories that poor laboratory practices and fraudulent activities will not be tolerated. These approaches may include, in part, preprogram laboratory audits and unannounced visits during the program, mandatory participation in accreditation and certification programs, submission of overt and blind performance evaluations samples, and more sophisticated third-party computer analysis of raw data and electronic records (Wait, 2002).

4.5 ASSESSMENT

Data assessment involves the scientific and statistical evaluation of data to determine whether the results are of the right type, quality, and quantity to defensibly support the decisions that need to be made with a specified confidence. Meaningful assessments require that the assessor be in possession of all documentation that defines data quality objectives and provides details about the methods used. These documents include the laboratory's quality assurance plan, a sampling and analysis plan and/or a quality assurance project plan (QAPP), and all pertinent field and laboratory standard operating procedures (SOPs) (US EPA, 2000b,c, 2001; Wait, 2002). An assessor should gauge their opinions, in part, to the information provided in these documents.

4.5.1 DOCUMENTATION

A tenet of forensic science in all aspects of an investigation need to be documented in a manner that allows one to independently recreate and evaluate the work at a later date. Adequate documentation involves thorough sample custody records, field sampling notes and records that are decipherable, and comprehensive laboratory records for all sample storage, preparation, and analysis procedures and results.

Chain-of-custody is one of the most critical documentation requirements. A sample is under custody if one or more of the following criteria are met:

- It is in the sampler's or transferee's possession
- It is in the sampler's or transferee's view after possession
- It is placed in a designated secure area
- It has intact custody seals

There must be documentation in place with authentic signatures that demonstrate the possession of the sample is traceable throughout the measurement process (Berger *et al.*, 1996). Field sampling records are also an important part of the documentation record and should reflect field measurements, field conditions, and personnel. Whereas accurate and relevant field records allow resolution of field and laboratory data interpretation issues, poor and inaccurate field records may raise more questions than they answer. Similar documentation and custody records must be maintained in the laboratory until the measurement process is complete (ASTM, 1992). Regarding data assessment, all document and raw data must be available to a reviewer to allow for the recalculation of all results. A suggested listing of data package elements needed to assess organic GC/MS results are provided in Table 4.7, and for inductively-coupled plasma (ICP) trace metal results are provided in Table 4.8.

4.5.2 QUALITY CONTROL INTERPRETATION

The purpose of quality control (QC) is to help assess data quality, and identify and measure errors in the measurement process. Quality control is the only method to measure errors, therefore the most important component of data assessment. It is important to understand information provided by specific quality control parameters. Indiscriminate collection of quality control samples may be too excessive for some projects and insufficient for others. If insufficient quality control is performed, the validity of the data and any resultant conclusions are suspect. All too often quality control is considered a checklist item with no thought to interpreting the quality control results. However, itemized quality control assessments, known as *validation*, have an important role in interpreting quality control results (US EPA, 2002b).

Measurable quality control parameters include (1) blanks, which provide information on possible contamination during sampling and analysis activities; (2) replicates, which provide information on precision; and (3) spikes, which indicate bias. Method or matrix blanks, another important quality control measure, indicates the limit of reliable detection of an analytical method. The accuracy of the analytical instrumentation should be established through the use of reference standards (Clement *et al.*, 1997).

Table 4.7

Suggested content for a forensic GC/MS data package.

1. Laboratory narrative summarizing methods and discussing QC issues
2. Cross reference of field sample numbers, laboratory identification numbers and QC batches
3. Chain-of-custody forms/laboratory receipt forms
4. Sample results
5. Method blank results
6. Replicate extraction results (subsampling precision)
7. Replicate analysis results
8. LCS/LCSD results and recoveries, with RPD control limits
9. Surrogate standard recoveries, with control limits
10. MS/MSD results and recoveries, with acceptance and RPD control limits
11. Instrument performance check (tuning)
12. Initial calibration data and acceptance limits
13. Initial and continuing calibration verification, results, recoveries, and acceptance limits
14. Internal standard areas and acceptance limits
15. Raw data for each sample, blank, spike, duplicate, and standard (quantification reports, reconstructed ion chromatograms)
16. Raw and background subtracted mass spectra with library spectra of five best-fit matches
17. Sample preparation bench sheets
18. Standard preparation and traceability logs
19. Analysis run logs
20. Percent moisture for soils/sediments

QC = quality control; GC/MS = gas chromatography/mass spectrometry; LCS/LCSD = laboratory control sample/laboratory control sample duplicate; RPD = relative percent difference; MS/MSD = matrix spike/matrix spike duplicate.

Blank (Trip, field, decontamination) samples are used to determine if any contamination exists in the measurement process (Maney, 1997). The typical interpretation is that if there is no trace of an analyte in the blank samples, there is no contamination in the measurement process. If there is some measurable level of analyte in the blank samples, data interpretation becomes more difficult. It is possible to detect measurable levels of analyte A when analyte B is of interest. In this case the claim can be made that any concentration of analyte B found in the samples is not due to contamination. Another example would be low levels of analyte B in the blank samples, but high levels of analyte B in the field samples. A claim could be made that the contamination, although present, is insignificant and can be ignored. These two examples demonstrate why it is not advisable to have a quality control program that excludes data in which any contamination is found in the blank samples. Each project's blank sample quality control must be interpreted in light of specific facts and goals of that project.

Split sample results are compared to determine if reasonable differences exist. Usually splits are sent to two different laboratories, but can also be used to check two analysts or two methods in the same laboratory. Acceptable differences in split results can be specified as part of the data quality objectives. If the difference in results is greater than the difference allowed, the data are

Table 4.8
Suggested content for a forensic trace metal ICP data package.

1. Laboratory narrative summarizing methods and discussing QC issues
2. Cross reference of field sample numbers, laboratory identification numbers, and QC batches
3. Chain-of-custody forms/laboratory receipt forms
4. Sample results
5. Method blank results
6. Preparation and initial/continuing calibration blank results
7. Replicate digestion results (subsampling precision)
8. Replicate analysis results
9. LCS/LCSD results and recoveries, with acceptance and RPD control limits
10. MS/MSD results and recoveries, with acceptance and RPD control limits
11. Post-digestion spike recovery and acceptance limits
12. Laboratory duplicate results and acceptance limits
13. Initial calibration data and acceptance limits
14. Initial and continuing calibration verification results, recoveries, and acceptance limits
15. ICP interference check sample results
16. ICP serial dilution results
17. Sample preparation logs
18. Analysis run logs
19. ICP interelement correction factors
20. Instrument raw data printouts for all samples and QC
21. Standard preparation and traceability logs
22. Percent moisture for soils/sediments

QC = quality control; ICP = inductively-coupled plasma; LCS/LCSD = laboratory control sample/laboratory control sample duplicate; RPD = relative percent difference; MS/MSD = matrix spike/matrix spike duplicate.

suspect (Grant *et al.*, 1996). Measurement uncertainty should be known to properly evaluate the difference between splits. The goal of the splitting process is to make the splits as similar as possible, such that any differences can be attributed to differences in laboratory performance, not to differences between the splits. If split results are divergent, one needs to determine which data are suspect. This is difficult to determine unless there is a sufficient amount of laboratory reliable quality control information available. Of course it is not scientifically honest to ignore laboratory data that is counter to the desired result.

Field replicates are the only quality control samples that can measure sampling error, thus enhancing the importance of this measurement. For the purpose of concentration, this quality control event should always be included. A replicate is a repetition of the same process used to collect the initial sample from a decision unit. If the sample collection decision adequately represents the decision unit, repeating the collection process should yield the same result.

Laboratory quality control check samples are used to demonstrate precision, bias, and detection limits. A listing of typical quality control check samples is provided in Table 4.9. Calibration is the establishment of a quantitative relationship between the response of a test method to the concentration

Instrument Blank
Calibration Blank
Method Blank
Laboratory Control Sample (LCS) (spike into blank water)
Matrix Spike (spike into site sample)
Laboratory Replicate
Surrogate Standard
Internal Standard

Table 4.9
Common laboratory quality control check samples.

of a target analyte (Berger, 1996). There are various methods, using both internal and external standards, to accomplish acceptable calibration (Popek, 2003). The boundaries of calibration are established at the low concentration end by electronic or chemical noise and at the high concentration end by either detector saturation or the slope "flattening" of the calibration curve. Proper calibration is critical to producing reliable concentration results, and calibration methodology and acceptance criteria should be established during the planning process.

Commercial laboratories use various means and terms to report detection limits to a client (Rosecrance, 2000). Understanding the basis for which a laboratory is reporting a nondetect is crucial to a forensic investigator, and if not properly defined can lead to inappropriate use of a method or a misunderstanding as to whether an analyte is present (Koorse, 1989; Lambert *et al.*, 1991). Although an analytical chemist theoretically may view detection limits strictly as the signal-to-noise ratio of an instrument's output, more practically a detection limit must reflect the vagaries of method performance as well as the influence of a sample matrix (Baldwin *et al.*, 1997). US EPA has defined a Method Detection Limit (MDL) as the "minimum concentration of a substance that can be measured and reported with 99% confidence that the true value is greater than zero" (US EPA, 2003b). US EPA has further defined a Practical Quantification Limit (PQL) as "the lowest level that can be reliably achieved within specified limits of precision and accuracy during routine laboratory operating conditions" (US EPA, 1985). More than 50 different terms have been used to describe detection and quantification capabilities of test methods (Telliard, 2004). The lack of consensus among various government agencies and scientific organizations regarding a single approach has led to much consternation, culminating in a lawsuit brought against US EPA by various trade organizations, which was eventually settled (*Alliance of Automobile Manufacturers* et al. *v. Carol Brunner* [No. 99-1420, D.C. Cir. App. 2001] (US EPA, 2003c). Logically, forensic investigators should ensure that the test method of choice actually achieves the detection limit required for their study.

Assessment of quality control data can sometimes be misleading. Long-term system bias at a particular laboratory may not be apparent because most quality control systems are designed for a short-term evaluation (e.g., matrix

spikes, surrogates). This "snapshot" of a laboratory's performance does not reflect long-term bias. Control charts allow for long-term performance assessment, however they are not diligently used by many laboratories. For instance, most laboratories do not monitor slow long-term deterioration of GC/MS response factors; therefore, laboratories may only recalibrate when they violate regulated acceptable limits.

4.5.3 STATISTICAL ANALYSIS

A statistical evaluation may be performed to determine the adequacy and completeness of collected data relative to the data quality objectives, evaluate whether the data are sufficiently representative of a decision unit, or examine data trends (US EPA, 2000b). When statistical analysis is performed, it must be consistent with the data quality objectives. If the statistical assessment is done independent of the data quality objectives, it is likely that incorrect decisions will result. It is always desirable, though not always feasible, to state which statistical analyses will be performed as part of the data quality objectives.

Multivariate statistical techniques involving chemometric methods, such as Principal Component Analysis (PCA) and Polytopic Vector Analysis (PVA), can assist investigators to evaluate and demonstrate measurement trends and unique attributes of a set of data (Johnson and Ehrlich, 2002; Plumb, 2004; Mudge, 2005). Statistical chemical trend analysis can also be a useful tool to decipher the overall contaminant pattern of a single chemical present in multiple locations or decision units at a site (Gauthier, 2001). Extensive discussions on these methodologies are presented elsewhere in this volume. Investigators should be aware that if statistical methods are not applied or interpreted properly, fatally flawed opinions may result (Gastwirth, 2000; Sutherland, 2001).

4.5.4 DATA USABILITY ASSESSMENT

Data usability assessment is the process of determining and assuring that the quality of data produced for an investigation meets the intended use. An experienced data assessor should consider data usability in a holistic manner, not only examining the adequacy of documentation and quality control performance, but also considering site-specific circumstances, additive effects of quality control exceedances, and overall performance nuances. A holistic assessment should evaluate whether:

- The data have a well-defined purpose
- The data are representative of the sampled matrix
- The data adhere to applicable standards and specifications
- The data are technically and legally defensible

There are two parts to a usability assessment—qualitative and quantitative. Qualitative assessment verifies that everything was done in a manner that it should have been done. Some examples are whether log books were filled out completely or if the correct sampling tools were used. Quantitative assessments examine whether the results and associated quality control met acceptable limits. If only part of the assessment is performed, then the data lack defensibility. Qualitative assessment may seem easier, but aspects of it can be difficult to determine. For instance, how would one know if the log book is complete? How would one know if the correct sampling tool was used, and used properly? Qualitative assessment items need to be determined in advance and systems put in place to allow an assessment at a later date. Quantitative assessment is more than just comparing numbers or performing endless statistical tests. Any type of statistical analysis must be consistent with the data quality objectives, and statistical assumptions must be verified. Regarding quality control interpretation, the assessor must consider the effects of numerous quality control exceedances in their totality. Table 4.10 (modified from US EPA, 1992b) provides some examples of quality control exceedances, and how these exceedances may affect data usability. Experienced chemists should be involved with data usability determinations.

Table 4.10
Quality Control Data Usage

QC Criterion	Effect on Identification	Effect on Concentration	Use
Spikes (high recovery)	–	High bias	Use data as upper limit
Spikes (low recovery)	False negative[1]	Low bias	Use data as lower limit
Duplicate	None, unless analyte found in one duplicate and not in the other. Then either false positive or false negative.	High or low bias[2]	Use data as estimate—poor precision
Blanks	False positive	High bias	Set confidence level 5× blank Use data above confidence level Use data below confidence level as estimate
Calibration	–	High or low bias[2]	Use data as estimate unless problem is extreme
Internal standards (reproducibility)[3]	–	–	Use data as estimate—poor precision
Internal standards (high recovery)	–	Low bias	Use data as lower limit
Internal standards (low recovery)	False negative[1]	High bias	Use data as upper limit

[1]False negative only likely if recovery is near zero.
[2]Effect on bias determined by examination of data for each individual analyte.
[3]Includes surrogates and system monitoring compounds.

Forensic investigators may also decide that for sensitive programs, studies using novel techniques, or investigations with confounding results, a peer review may be useful as part of the assessment process. Peer reviews either internal to the study team, or external to the team, can be beneficial in determining whether the science used was sound and interpreted properly, and whether the results will be admissible in court (US EPA, 2000d; Brilis and Lyon, 2004).

4.6 SUMMARY

Environmental forensic investigations typically rely upon chemical and physical data. Using representative data of known quality and integrity is paramount to making scientifically sound decisions that can be defended and ultimately be admissible in court. Investigators should proactively design sampling and analysis programs with clearly defined data quality objectives, and ensure that the study is conducted according to plan and its implementation properly documented. Admissibility of evidence will more likely succeed when a reliable, relevant and defensible method is implemented correctly by trained and experienced scientists. Lastly, chemistry measurements produced by other parties should be closely scrutinized prior to use in decision-making.

REFERENCES

Allchin, D. (2002) Error types. *Perspectives on Science* 9, 38–58.

American Society of Testing and Materials (ASTM). (1991) Standard practice for comparing test methods, D4855-91.

American Society of Testing and Materials (ASTM). (1992) Standard practice for receiving, documenting, storing, and retrieving evidence in a forensic science laboratory, E1492-92.

American Society of Testing and Materials (ASTM). (1996) Standard practice for generation of environmental data related to waste management activities, development of data quality objectives, D5792-95.

American Society of Testing and Materials (ASTM). (1997) Standard guide for soil gas monitoring in the vadose zone, D5314-92.

American Society of Testing and Materials (ASTM). (1998) Standard guide for generation of environmental data related to waste management activities: Selection and optimization of sampling design, D6311-98.

American Society of Testing and Materials (ASTM). (2001) Standard guide for field collection of organic compounds from surfaces using wipe sampling, D6661-01.

American Society of Testing and Materials (ASTM). (2005) Standard guide for collection of surface dirt by micro-vacuum sampling for subsequent metals determination, D7144-05a.

Baldwin, R., Bethem, R., Boyd, R., Budde, W., Cairns, T., Gibbons, R. *et al.* (1997) 1996 ASMS Fall Workshop: Limits to confirmation, quantification, and detection. *Journal of the American Society of Mass Spectrometry* 8, 1180–1190.

Bayne, C.K., Dindal, A.B., Jenkins, R.A., Crumbling, D.M., and Koglin, E.N. (2001) Meeting data quality objectives with interval information. *Environmental Science & Technology* 35, 3350–3355.

Berger, W., McCarty, H., and Smith, R. (1996) *Environmental Laboratory Data Evaluation*. Genium Publishing Corporation, Amsterdam, New York.

Blumer, M. (1976) Polycyclic aromatic hydrocarbons in nature. *Scientific American* 234, 35–45.

Boehm, P.D. (2006) Polycyclic aromatic hydrocarbons (PAHs). In *Environmental Forensics—Contaminant Specific Guide*. Academic Press, San Diego.

Boehm, P.D., Burns, W.A., Page, D.S., Bence, A.E., Mankiewicz, P.J., Brown, J.S., and Douglas, G.S. (2002) Total organic carbon, an important tool in a holistic approach to hydrocarbon source fingerprinting. *Environmental Forensics* 3, 243–250.

Bonomo, L., Butelli, P., Romele, L., and Saponaro, S. (2002) Uncertainty estimation in a remediation process: A case study. *Soil and Sediment Contamination* 11(6), 827–839.

Bruya, J.E. and Costales, M. (2005) Discussion of the error associated with polycyclic aromatic hydrocarbons (PAH) analyses. *Environmental Forensics* 6, 175–185.

Brilis, G.M. and Lyon, J.G. (2004) Peer review as a QA tool in environmental applications. *Quality Assurance* 11, 111–129.

Clement, R.E., Keith, L.H., and Siu, K.M. (1997) *Reference Materials for Environmental Analysis.* CRC Press, Boca Raton.

Costa, H.J., White, K.A., and Ruspantini, J.J. (2004) Distinguishing PAH background and MGP residues in sediments of a freshwater creek. *Environmental Forensics* 5, 171–182.

Crumbling, D.M., Groenjes, C., Lesnik, B., Lynch, K., Shockley, J., Van EE, J. *et al.* (2001) Managing uncertainty in environmental decisions. *Environmental Science & Technology* 35, 405A–409A.

Currie, L.A. (1999) Detection and quantification limits: Origins and historical overview. *Analytica Chemica Acta* 391, 127–134.

Davidson, J.M. and Creek, D.N. (2000) Using the gasoline additive MTBE in forensic environmental investigations. *Environmental Forensics* 1, 31–36.

Denton, C.M., Sklash, M.G., and Eng, P. (2006) Contaminant fate and transport in the courtroom. In *Contaminated Soils, Sediments and Water. Volume 10: Successes and Challenges.* Springer, New York.

Douglas, G.S., Burns, W.A., Bence, A.E., Page, D.S., and Boehm, P. (2004) Optimizing detection limits for the analysis of petroleum hydrocarbons in complex environmental samples. *Environmental Science & Technology* 38, 3958–3964.

Electric Power Research Institute (EPRI). (2000) *Chemical Source Attribution at Former MGP Sites*. Report No. 1000728, EPRI, NYSEG and RG&E.

Feenstra, S. (2005) Soil sampling in NAPL source zones: Challenges in representativeness. *Environmental Forensics* 6, 57–63.

Gaborek, B.J., Mullikin, J.M., Pitrat, T., Cummings, L., and May, L.M. (2001) Pentagon surface wipe sampling health risk assessment. *Toxicology and Industrial Health* 17, 254–261.

Gastwirth, J.L. (2000) *Statistical Science in the Courtroom.* Springer, New York.

Gauthier, T.D. (2001) Detecting trends using Spearman's Rank correlation coefficient. *Environmental Forensics* 2, 359–362.

Georgian, T. and Osborn, K.E. (2001) A practical strategy for determining and verifying detection limits. *Environmental Testing & Analysis* 10(3), 13–14.

Gerlach, R.W., Dobb, D.E., Raab, G.A., and Nocerino, J.M. (2002) Gy sampling theory in environmental studies, 1. Assessing soil splitting protocols. *Journal of Chemometrics* 16, 321–328.

Gilbert, R.O. and Pulsipher, B.A. (2005) Role of sampling designs in obtaining representative data. *Environmental Forensics* 6, 27–33.

Grant, C.L., Jenkins, T.F., and Mudambi, A.R. (1996) Comparison criteria for environmental chemical analyses of split samples sent to different laboratories, CCREL Special Report 96-6. United States Army Cold Regions Research and Engineering Laboratory.

Gy, P. (1992) *Samples of Heterogeneous and Dynamic Materials Systems.* Elsevier, New York.

Gy, P. (1998) *Sampling for Analytical Purposes*. John Wiley, New York.

Halden, R.U., Hoppel, A.M., and Schoen, S.R. (2001) Evaluation of standard methods for the analysis of methyl tert-butyl ether and related oxygenates in gasoline-contaminated groundwater. *Environmental Science & Technology* 35, 1469–1474.

Hartman, B. (2002) How to collect reliable soil-gas data for risk-based applications, Part 1: active soil-gas method. *LUSTline Bulletin* 42, 17–22.

Hartman, B. (2003) How to collect reliable soil-gas data for risk-based applications, Part 2: surface flux-chamber method. *LUSTline Bulletin* 44, 14–18.

Hewitt, A.D. (1994) Comparison of sampling methods for trichloroethylene in vadose zone soils. *Journal of the Association of Analytical Chemists International* 77(2), 458–463.

Hewitt, A.D. (1997) Chemical preservation of volatile organic compounds in soil. *Environmental Science & Technology* 31, 67–70.

Hewitt, A.D. (1999a) Relationships between soil vapor and soil matrix—measurements for trichloroethene. *Environmental Testing & Analysis* 8(3), 25–32.

Hewitt, A.D. (1999b) Storage and preservation of soil samples for volatile organic compound analysis, CRREL Special Report 99-5. United States Army Cold Regions Research and Engineering Laboratory.

Hewitt, A.D. (1999c) Frozen storage of soil samples for volatile organic compounds analysis. *Environmental Testing & Analysis* 8(5), 18–45.

Hewitt, A.D., Jenkins, T.F., Ranney, T.A., Stark, J.A., Walsh, M.E., Taylor, S. *et al.* (2003) Estimates for explosive residues from the detonation of Army munitions, CRREL Technical Report TR-03-16. United States Army Cold Regions Research and Engineering Laboratory.

International Organization for Standardization and International Electrotechemical Commission (ISO/IEC). (1999) ANS/ISO/IEC 17025—General requirements for the competence of testing and calibration laboratories.

Jenkins, T.F., Grant, C.L., Brar, G.S., Thorne, P.G., Ranney, T.A., and Schumaker, P.W. (1996) Assessment of sampling error associated with collection and analysis of soil samples at explosives–contaminated sites, CRREL Special Report 96-15. United States Army Cold Regions Research and Engineering Laboratory.

Jenkins, T.F., Grant, C.L., Walsh, M.E., Thorne, P.G., Thiboutot, S., Ampleman, G., and Ranney, T.A. (1999) Coping with spatial heterogeneity effects on sampling and analysis at an HMX-contaminated antitank firing range. *Field Analytical Chemistry and Technology* 3(1), 19–28.

Jenkins, T.F., Hewitt, A.D., Walsh, M.E., Ranney, T.A., Ramsey, C.A., Grant, C.L., and Bejella, K.L. (2005) Representative sampling for energetic compounds at military training ranges. *Environmental Forensics* 6, 45–55.

Johnson, G.W. and Ehrlich, R. (2002) State of the art report on multivariate chemometric methods for environmental forensics. *Environmental Forensics* 3, 59–79.

Johnson, R.L. (1992) The dissolution of dense immiscible solvents into groundwater: Implications for site characterization and remediation. In *Groundwater Contamination and Analysis at Hazardous Waste Sites.* Marcel Dekker, New York, 463–477.

Keith, L.H., Brass, H.J., Sullivan, D.J., Boiani, J.A., and Alben, K.T. (2005) An introduction to the National Environmental Methods Index. *Environmental Science and Technology* 39, 173A–176A.

Koorse, S.J. (1989) False positives, detection limits, and other laboratory imperfections: The regulatory implications. *Environmental Law Reporter* 19, 10211–10222.

Kram, M.L., Keller, A.A., Rossabi, J., and Everett, L.G. (2001) DNAPL characterization methods and approaches, Part 1: Performance comparisons. *Ground Water Monitoring & Remediation* 21(4), 109–145.

Laflamme, R.E. and Hites, R.A. (1978) The global distribution of polycyclic aromatic hydrocarbons in recent sediments. *Geochimica et Cosmochimica Acta* 42, 289–303.

Lambert, D., Peterson, B., and Terpenning, I. (1991) Nondetects, detection limits, and the probability of detection. *Journal of American Statistical Association* 86, 266–277.

Lima, A.C., Farrington, J.W., and Reddy, C.M. (2005) Combustion–derived polycyclic aromatic hydrocarbons in the environment—A review. *Environmental Forensics* 6, 109–131.

Linkov, I., Ames, M.R., Crouch, E., and Satterstrom, F.K. (2005) Uncertainty in octanol-water partition coefficient: Implications for risk assessment and remedial costs. *Environmental Science & Technology* 39, 6917–6922.

Maney, J. (1997) Assessing blank data. *Environmental Testing & Analysis* 6(4), 20–22.

Maney, J. (2001) Sample collection design. In *Hazardous and Radioactive Waste Treatment Technologies Handbook*. CRC Press, New York, 2.1.3–2.1.33.

Maney, J. (2002) Optimizing data collection design. *Environmental Science & Technology* 36, 383A–389A.

Maney, J. and Wait, D. (1991) The importance of measurement integrity. *Environmental Lab* 3(5), 20–25.

Maney, J. and Wait, D. (2005) The role of representativeness in the project life cycle. *Environmental Forensics* 6, 17–19.

Marrin, D.L. and Kerfoot, H.B. (1988) Soil-gas surveying techniques. *Environmental Science & Technology* 22, 740–745.

Mercer, J.W. and Cohen, R.M. (1990) A review of immiscible fluids in the subsurface: Properties, models, characterization and remediation. *Journal of Contaminant Hydrology* 6, 107–163.

Mishalanie, E. and Ramsey, C. (1999) Obtaining trustworthy environmental data: Sampling and analysis issues. *Natural Resources & Environment* 13, 522–527.

Morrison, R. (2000) Application of forensic techniques for age dating and source identification in environmental litigation. *Environmental Forensics* 1, 131–153.

Morrison, R. and Murphy, B. (2006) *Environmental Forensics—Contaminant Specific Guide*. Academic Press, San Diego.

Mudge, S.M., Birch, G.F., and Matthai, C. (2003) The effect of grain size and element concentration in identifying contaminant sources. *Environmental Forensics* 4, 305–312.

Mudge, S.M. (2005) Signature analysis by statistical methods: Chemometric methods for source allocation. In *Preceedings of Southeast Asia Environmental Forensics Conference 2005*. 456, 1–5, September 19–20, 2005.

Murray, R.C. and Tedrow, J. (1992) *Forensic Geology*. Prentice Hall, Englewood Cliffs.

Office of Technology Assessment (OTA). (1983) Habitability of the Love Canal area: An analysis of the technical basis for the decision on the habitability of the emergency declaration area, United States Congress, No. PB84-114917.

Olson, S.S. (1999) *International Environmental Standards Handbook.* CRC Press, Boca Raton.

Organization for Economic Co-operation and Development (OECD). (1993) OCED guidelines for testing of chemicals.

Peters, K.E., Walters, C.C., and Moldowan, J.M. (2005) *The Biomarker Guide*, 2e. Cambridge University Press, Cambridge, UK.

Petrisor, I.G. (2006) Use of oxygenates to date a gasoline release. *Environmental Forensics* 7, 103–104.

Pitard, F.F. (1993) *Pierre Gy's Sampling Theory and Sampling Practice*, 2e. CRC Press, Boca Raton.

Plumb, R.H. (2004) Fingerprint analysis of contaminant data: A forensic tool for evaluating environmental contamination, EPA/600/5-04-054. US EPA Office of Research and Development.

Pontolillo, J. and Eganhouse, R. (2001) The search for reliable aqueous solubility (S_w) and octanol-water partition coefficient (K_{ow}) data for hydrophobic organic compounds: DDT and DDE as a case study. U.S. Geological Survey Water-Resources Investigation Report 01-4201.

Popek, E.P. (2003) *Sampling and Analysis of Environmental Chemical Pollutants. A Complete Guide*. Academic Press, San Diego.

Qian, S.S. and Lyons, R.E. (2006) Characterization of background concentrations of contaminants using a mixture of normal distributions. *Environmental Science and Technology* 40, 6021–6025.

Ramsey, C.A. and Hewitt, A.D. (2005) A methodology for assessing sample representativeness. *Environmental Forensics* 6, 71–75.

Ramsey C.A. and Suggs, J. (2001) Improving laboratory performance through scientific subsampling techniques. *Environmental Testing & Analysis* 10(2), 12–16.

Robbins, M.C. (1980) OSHA's wipe sampling policies and procedures. In *Sampling and Analysis of Toxic Organics in the Atmosphere*. ASTM STP 721, 48–55.

Rosecrance, A. (2000) The three "Rs" for relevant detection, reliable quantitation, and respectable reporting limits. *Environmental Testing & Analysis* 9(6), 13.

Ryba, S.A. and Burgess, R.A. (2002) Effects of sample preparation on the measurement of organic carbon, hydrogen, nitrogen, sulfur, and oxygen concentrations in marine sediments. *Chemosphere* 48, 139–147.

Saxton, G.N. and Engel, B. (2005) A survey of soil sample handling procedures of state pesticides agencies. *Environmental Forensics* 6, 105–108.

Schultz, W.G. (2001) Chemists needed for forensic analyses. *Chemical and Engineering News* 79(45), 51.

Shapiro, M. and Charrow, R. (1989) The role of data audits in detecting scientific misconduct: Results of the FDA program. *Journal of American Medical Association* 261, 2505–2511.

Shelton, L.R. (1994) Field guide for collecting and processing stream–water samples for the National Water–Quality Assessment Program. U.S. Geological Survey Open—File Report 94-455.

Siegrist, R.L. and Jenssen, P.D. (1980) Evaluation of sampling method effects on volatile organic compound measurements in contaminated soils. *Environmental Science & Technology* 24, 1387–1392.

Simmons, B. (1999) Are the data legally defensible? *Hydrovisions* 8(2), 8–9.

Slayton, T.M., Valberg, P.A., and Wait, D. (1998) Estimating dermal transfer from PCB-contaminated porous surfaces. *Chemosphere* 36, 3003–3014.

Smith, J., Eng, L., Comeau, J., Rose, C., Schulte, R., Barcelona, M. *et al.* (1996) Volatile organic compounds in soil—Accurate and representative analysis. In *Principles of Environmental Sampling.* ACS Professional Reference Book, American Chemical Society, Washington, DC.

Smith, R. (2000) *Interpretation of Organic Data*, Genium Publishing Corporation, Amsterdam, New York.

Smith, R. (2003) *Guide to Environmental Analytical Methods*, 5e. Genium Publishing Corporation, Amsterdam, New York.

Stelling, M.A. and Sjerps, M. (2005) Heterogeneity of contaminants in solid waste materials sampling for compliance testing. *Environmental Forensics* 6, 231–239.

Stout, S.A., Douglas, G.S., and Uhler, A.D. (2006) Automotive gasoline. In *Environmental Forensics—Contaminant Specific Guide*. Academic Press, San Diego.

Stout, S.A., Uhler, A.D., Naymik, T.G., and McCarthy, K.J. (1998) Environmental forensics—Unraveling site liability. *Environmental Science & Technology* 32, 260A–264A.

Stout, S.A. and Wasielewski, T.N. (2004) Historical and chemical assessment of the sources of PAHs in soils at a former coal-burning power plant, New Haven, Connecticut. *Environmental Forensics* 5, 195–211.

Suggs, J.A., Beam, E.W., Biggs, D.E., Collins, W., Dusenbury, M.R., Macleish, P.P. *et al.* (2002) Guidelines and resources for conducting an environmental crime investigation in the United States. *Environmental Forensics* 3, 91–113.

Sutherland, R.A. (2001) Analysis and commentary on "Statistical Methods and Pitfalls in Environmental Data Analysis" by Yue Rong. *Environmental Forensics* 2, 265–273.

Swartz, M. and Krull, I. (2006) Method validation and robustness. *LCGC North America* 24, 480–490.

Taylor, J.K. (1987) In *Quality Assurance of Chemical Measurements*. Lewis Publishers, Boca Raton, 74–94.

Telliard, W.A. (2004) EPA's assessment of detection and quantification limit concepts. PITTCON 2004, Chicago, March 8.

United States Environmental Protection Agency (US EPA). (1979) Memorandum from Doug Costle, US EPA Administrator, to the other Administrators and General Counsel, May 30, 1979.

United States Environmental Protection Agency (US EPA). (1985) National primary drinking water regulations; volatile synthetic organic chemicals. *Federal Register* 50, 46906–46908.

United States Environmental Protection Agency (US EPA). (1986) Test methods for evaluating solid waste: Physical/chemical methods, 3e. SW-846.

United States Environmental Protection Agency (US EPA). (1989) Sampling of common pesticides and PCBs from inert surfaces, EPA 330/1-90-001.

United States Environmental Protection Agency (US EPA). (1991a) Description and sampling of contaminated soils—A field pocket guide, EPA/625-12-91/002.

United States Environmental Protection Agency (US EPA). (1991b) Chip, wipe, and sweep sampling, SOP #2011. In *Compendium of ERT Waste Sampling Procedures.* OSWER Directive 9360.4-07.

United States Environmental Protection Agency (US EPA). (1992a) Preparation of soil sampling protocols: Sampling techniques and strategies, EPA/600/R-92/128.

United States Environmental Protection Agency (US EPA). (1992b) Guidance for data usability in risk assessment (Part A), Final, PB92-963356.

United States Environmental Protection Agency (US EPA). (1994) DNAPL site characterization, EPA/540/F-94/049.

United States Environmental Protection Agency (US EPA). (1997) National ambient air quality standards for particulate matter: Final rule. *Federal Register* 64, 35714–35744.

United States Environmental Protection Agency (US EPA). (1998a) Passive soil gas sampler. Gore-Sorber® screening survey passive soil gas sampling system. Environmental Technology Verification Program Statement, EPA-VS-SCM-19.

United States Environmental Protection Agency (US EPA). (1998b) Passive soil gas sampler. Emflux® soil gas investigation system. Environmental Technology Verification Program Statement, EPA-VS-SCM-22.

United States Environmental Protection Agency (US EPA). (1998c) Region 1, EPA-New England Sediment Sampling Guidance, draft.

United States Environmental Protection Agency (US EPA). (2000a) Guidance for obtaining representative laboratory analytical subsamples from particulate laboratory samples, EPA/600/R-03/027.

United States Environmental Protection Agency (US EPA). (2000b) Guidance for data quality assessment—Practical methods for data analysis, EPA QA/G-9, QA-00 Update, EPA/600/R-96/084.

United States Environmental Protection Agency (US EPA). (2000c) Guidance for quality assurance project plans, QA/G-5, EPA/240/R-02/009.

United States Environmental Protection Agency (US EPA). (2000d) Science Policy Council Handbook—Peer review, 2e, 100-B-00-001.

United States Environmental Protection Agency (US EPA). (2001) Guidance for preparing standard operating procedures, QA/G-6, EPA/240/B-01/004.

United States Environmental Protection Agency (US EPA). (2002a) Guidance for the data quality objective process, EPA QA/G4, EPA/600/R-96/055.

United States Environmental Protection Agency (US EPA). (2000b) Guidance on environmental data verification and data validation, QA/G-8, EPA/240/R-02/004.

United States Environmental Protection Agency (US EPA). (2003a) Index to EPA test methods. Available at http://www.epa.gov/region01/oarm/testmeth.pdf.

United States Environmental Protection Agency (US EPA). (2003b) Guidelines establishing test procedures for the analysis of pollutants; procedures for detection and quantification (40 CFR, Part 136). *Federal Register* 68, 11770–11793.

United States Environmental Protection Agency (US EPA). (2003c) Technical support document for the assessment of detection and quantification approaches, EPA-821-R-03-005.

United States Environmental Protection Agency (US EPA). (2004) Contract Laboratory Program guidance for field samplers, EPA 540-R-00-003.

United States Environmental Protection Agency (US EPA). (2005) Sample holding time reevaluation, EPA/600/R-05/124.

United States Environmental Protection Agency (US EPA) (2006) Promising techniques identified to improve drinking water laboratory integrity and reduce public health risks, Office of Inspector General Report No. 2006–P–00036, September 21, 2006.

United States Geological Survey (USGS). (1997) Comparison of the suspended sediment splitting capabilities of the churn and cone splitters, USGS Office of Water Quality Technical Memorandum 97.06.

United States Geological Survey (USGS). (2005) A guide to the proper selection and use of federally approved sediment and water quality samplings, USGS Open File Report 2005-1087.

United States Supreme Court (USSC). (1993) *Daubert v. Merrill Dow Pharmaceuticals,* 509 U.S. 579.

Wait, D. (2000) Evolution of organic analytical methods in environmental forensic chemistry. *Environmental Forensics* 1, 37–46.

Wait, D. (2002) Challenges in producing defensible environmental chemistry measurements for litigation. *Environmental Claims Journal* 14, 415–454.

Wait, D. and Cook, L. (1999) Opportunities in environmental forensic chemistry analysis. *Environmental Testing & Analysis* 8(4), 31–32.

Wait, D. and Douglas, G. (1995) QA for nonstandard analytical methods. *Environmental Lab* 7(6), 30–34.

Wait, A.D. and Maney, J.P. (2006) Regulatory science and the Data Quality Act. *Environmental Claims Journal* 18(2), 145–162.

Walsh, M.E., Ramsey, C.A., and Jenkins, T.F. (2002) The effect of particle size reduction by grinding on subsampling variance for explosives residues in soil. *Chemosphere* 49, 1267–1273.

Walsh, M.R. (2004) Field sampling tools for explosive residues developed at CRREL, CRREL Technical Note TN-04-1. United States Army Cold Regions Research and Engineering Laboratory.

Wang, Z., Stout, S.A., and Fingas, M. (2006) Forensic fingerprinting of biomarkers for oil characterization. *Environmental Forensics* 7, 105–146.

Weiner, E.R. (2000) In *Applications of Environmental Chemistry*. Lewis Publishers, Boca Raton, 124–145.

Whitaker, T.B. and Dickens, J.W. (1974) Variability of aflatoxin test results. *Journal of American Oil Chemists' Society* 51, 214–218.

Wight, G.D. (1994) *Fundamentals of Air Sampling*. Lewis Publishers, Boca Raton.

CHAPTER 5

STATISTICAL METHODS

Thomas D. Gauthier and Mark E. Hawley

5.1 Introduction 130
5.2 Background 132
5.2.1 Population Parameters and Sample Statistics 132
5.2.2 Population Distributions 134
5.2.3 Confidence Limits and Hypothesis Tests 137
5.2.3.1 Confidence Limits 138
5.2.3.2 Hypothesis Testing 140
5.2.3.3 Normality, Representativeness, and Independence 142
5.3 Applications in Environmental Forensics 144
5.3.1 Comparing Sample Means 144
5.3.1.1 Student's *t*-Test 144
5.3.1.2 Wilcoxon's Rank Sum Test 150
5.3.1.3 Paired *t*-Test 152
5.3.1.4 Analysis of Variance (ANOVA) 154
5.3.1.5 Kruskal-Wallis Test 157
5.3.1.6 Issues in Sampling Design 158
5.3.2 Linear Correlation and Linear Regression Analysis 160
5.3.2.1 Linear Regression Analysis 161
5.3.2.2 Pearson's Product Moment Correlation Coefficient 164
5.3.2.3 Spearman Rank Correlation Coefficient 168
5.3.2.4 Multiple Linear Regression and Correlation 170
5.3.3 Mapping: Trends and Spatial Persistence 171
5.3.3.1 Trends, Spatial Persistence, and Autocorrelation 171
5.3.3.2 Isopleth Mapping and Interpolation Methods 173
5.3.3.3 Estimation of Areal Averages 177
5.4 Conclusions 179
References 180
Further Reading 182

5.1 INTRODUCTION

Environmental forensic investigations can involve a variety of statistical analysis techniques. Most examples in the literature involve rather sophisticated applications including principal components analysis and chemical mass balance receptor modeling (discussed elsewhere in this book). These techniques are powerful analytical tools and provide useful insights for data interpretation but the results can be difficult to explain to a judge or jury.

In some cases, the simpler analyses described in this chapter may be sufficient for demonstrating a point or determining the statistical significance of a conclusion. For example, whether or not a facility has adversely affected the surrounding environment may be assessed by comparing concentrations detected on-site with background levels. Similarly, upwind/downwind or upstream/downstream measurements may be compared. The underlying concepts are easily understood and the results are amenable to a simple two-dimensional presentation (as opposed to the multidimensional framework required for some of the more advanced techniques). Example techniques for comparing two populations are Student's *t*-test, the paired *t*-test, and Wilcoxon's rank sum test.

Statistical techniques can also be used to identify relationships among contaminants at a site. The simplest application involves determination of characteristic ratios between contaminants—a precursor to fingerprinting.[1] In criminal forensics, fingerprints can be used for identification purposes and to place a person at the scene of a crime. Similarly, in environmental forensics, fingerprinting can be used for identification purposes and to link contamination to a given source. An example is the case when two or more sources are suspected of having contributed to the contamination at a site. The waste streams may contain similar constituents, but if the constituents are present in proportions that are unique to a given waste stream, it may be possible to distinguish contributions using statistical analysis. Correlation coefficients provide a simple measure of the linear association between two variables. If multiple contaminants are present, construction of a correlation matrix can be a useful tool for exploratory data analysis.

Statistical analysis can also be used to evaluate larger patterns of contamination or "footprints" in the environment. Relatively simple techniques are available for evaluating trends and establishing whether concentrations are increasing or decreasing over time or space. In some cases it may be possible to associate a unique pattern of contamination with a particular waste disposal practice. For example, a well-established groundwater contamination plume emanating from an underground cistern that connects to a laboratory sink drain is strong evidence that contaminants were routinely dumped down the

[1] Additional ratio analysis methods are discussed in Chapter 6.

drain. Statistics can be used to confirm the presence of any patterns or trends within the plume. Evidence of this nature is strong corroboration to eyewitness reports.

Finally, as theories about a case evolve and models are developed, it may be possible to test the significance or validity of the model if site-specific data are available. For example, the timing of a release may be based on results predicted from a reverse ground water model. The uncertainty in the model can be evaluated by comparing model predictions with observations. In many cases, simple linear regression analysis provides a sufficient basis for evaluating predictive models.

A significant benefit of statistical analysis is that one can associate some degree of confidence with the results. The potential applications are numerous because forensic investigations typically are based on a limited set of data, and experts often must arrive at conclusions based on incomplete evidence.

Engineers and scientists are trained to report results with a high degree of certainty (typically with a 90% or 95% confidence level); however, in the context of litigation, a confidence level of "more probable than not" (corresponding to greater than 50% probability) or "within a reasonable degree of scientific certainty" may be all that is required (Charrow and Bernstein, 1994). In any event, some measure of uncertainty usually is warranted in an environmental forensic investigation and methods are available to calculate these confidence levels.

In this chapter we discuss a variety of basic statistical analysis techniques and demonstrate their use in environmental forensics. The types of statistical tests discussed in this chapter are listed in Table 5.1. Following this introduction, we provide a brief discussion of some background materials

Table 5.1
Statistical tests discussed in this chapter.

Type of Analysis	Parametric Techniques[a]	Nonparametric Techniques
Comparing populations	Student's *t*-test Paired *t*-test Analysis of variance	Wilcoxon's rank sum test Kruskal Wallis test
Evaluating relationships	Pearson's product moment correlation coefficient	Spearman's rank correlation coefficient
Making predictions	Regression analysis	
Testing distributions		Chi-square goodness of fit test

[a]Parametric techniques require that the data follow some underlying distribution—typically a normal distribution.

and then discuss each technique and how it is applied. The chapter concludes with a brief summary and caution regarding misinterpretation of the results.

5.2 BACKGROUND

5.2.1 POPULATION PARAMETERS AND SAMPLE STATISTICS

Statistical inference is the process of drawing conclusions about an underlying population based on a sample or subset of the data. In most cases, it is not practical to obtain all the measurements in a given population. For example, if we were interested in knowing the average concentration of arsenic in the top two feet of soil at a one-acre site, and each measurement required a 100-gram sample, we would have to collect and analyze 37 million samples[2] to know the true average. This, of course, is impractical. As a tradeoff, we accept some uncertainty in our estimate of the true average in exchange for making fewer measurements.

The **population** consists of all the conceivable items, observations, or measurements in a group. In this example, the population consists of the total number of 100-gram quantities of soil contained in the top two feet of the one-acre site (i.e., 3.7×10^7 items). A **sample** is a subset of observations or measurements used to characterize the population. In the preceding example, we might collect and analyze 20 100-gram quantities of soil to estimate the average arsenic concentration. Thus, the sample would consist of those 20 measurements.

In this case, the **population parameter** of interest is the arithmetic mean or average of the 3.7×10^7 arsenic measurements. Parameters used to describe characteristics of the underlying population usually are represented by Greek letters. The arithmetic mean, denoted by the Greek letter μ (mu), is a measure of central tendency. Another parameter of interest is the standard deviation, a measure of the dispersion or variability in the population, denoted by the Greek letter σ (sigma).

Estimates of population parameters derived from a subset of the measurements in a sample drawn from the underlying population are called **sample statistics**. Latin letters are used to represent sample statistics. For example the sample mean is denoted by $\overline{x}$ (x-bar) and the sample standard deviation is denoted by s.

The arithmetic average or mean of the population, μ, is equal to the sum of all observations, x_i (where x_i is the ith observation), divided by the total number of observations, N.

[2] The total volume involved in the top two feet of a four-acre site is 87,120 cubic feet = 2.47×10^9 cubic centimeters. Assuming a soil bulk density of 1.5 g/cm^3, the total mass is 3.7×10^9 grams, which yields a total of 3.7×10^7 100-gram samples.

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} \tag{5.1}$$

Since we never really know the true population mean (unless we sample N times for all x_i), our best estimate of this value is the sample mean. The sample mean is equal to the sum of n values in the sample divided by the number of values.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{5.2}$$

One way to characterize the dispersion or variability in a population is to note the lowest and highest measurements. But this yields no information about how the data are distributed in relation to the mean. A better measure of dispersion is to see how the values vary, on average, in relation to the mean value. The average of the square of the deviations about the mean is called the mean square deviation or the variance. The variance is denoted by the Greek letter σ^2 (sigma squared) and is defined in Equation 5.3.

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N} \tag{5.3}$$

The population standard deviation is equal to the square root of the variance. It is also known as the root mean square deviation.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}} \tag{5.4}$$

The sample standard deviation, s, is an estimate of the population standard deviation, and is defined in Equation 5.5.

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} \tag{5.5}$$

Notice that the denominator in the square root term in Equation 5.5 is $n - 1$ instead of n. This is because one degree of freedom[3] is used to estimate the arithmetic mean ($\bar{x}$) in the sample.

The Central Limit Theorem states: If a variable x has a distribution with a mean μ, and a standard deviation σ then the sampling distribution of the mean ($\bar{x}$), based on random samples of size n, will have a mean equal to μ and a standard deviation ($\sigma_{\bar{x}}$) where:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{5.6}$$

and will tend to be normal as the sample size, n, becomes large (Kachigan, 1991). The standard deviation of the population divided by the square root of n is known as the standard error of the mean (SEM) and is an important parameter for estimating confidence limits. What is meant by a normal distribution is defined next; confidence limits are defined later in the text.

5.2.2 POPULATION DISTRIBUTIONS

Most statistical tests are conducted under the assumption that measurements in the underlying population follow some known distribution. These are referred to as parametric techniques. The most common distribution assumed is the Gaussian or normal distribution. The expression describing the frequency or probability distribution function (PDF) for a normal distribution is presented in Equation 5.7:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5.7}$$

where $f(x)$ is the expected frequency for any value of x. A normal distribution with a mean of zero and standard deviation equal to one is referred to as the standard normal distribution. In this special case, Equation 5.7 is reduced to the expression in Equation 5.8, and x is represented by the letter z, which is known as the normal deviate, also referred to as the z-score or standard score.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{5.8}$$

[3] Kachigan (1991) defines degrees of freedom as "the number of observations in the data collection that are *free* to vary after the sample statistics have been calculated. For example, in calculating a sample standard deviation, we must subtract the sample mean from each of the n data observations in order to get the deviations from the mean. But once we have completed the second last subtraction, the final deviation is automatically determined, since the deviations prior to squaring *must* sum to zero. So the last deviation is not free to vary, only $n - 1$ are free to vary."

The z-score is equal to the observed value of x minus the mean value, all divided by the standard deviation as indicated in Equation 5.9. It indicates where a given value is located on the normal curve in terms of the number of standard deviations from the mean.

$$z = \frac{(x - \mu)}{\sigma} \tag{5.9}$$

A plot of $f(z)$ vs. z-score, illustrated in Figure 5.1, reveals the familiar bell-shaped curve of the normal distribution.

The area under the standard normal curve represents 100% of the measurements in a population. The cumulative distribution function (CDF) yields the probability that some random value drawn from the population will be less than or equal to the specified value of x. Unfortunately, the CDF for the normal distribution cannot be evaluated analytically; however, most statistics texts contain tables of the proportion of the area under the standard normal curve as a function of z-score (for example, see Zar, 1984; Kachigan, 1991; or Gilbert, 1987). From these tables, it is easily determined that, for example, 68.26% of the area under the curve lies within ±1 standard deviation and that 95.46% lies within ±2 standard deviations. Other measures are reported in Table 5.2.

A related distribution is Student's t distribution. The t distribution is used when the true mean and standard deviation of the population are unknown—which is the case in most situations. The distribution is bell-shaped and symmetrical like the normal distribution but somewhat broader to account for uncertainty in estimation of the variance. The exact shape of the distribution varies as a function of the degrees of freedom (v, where $v = n - 1$), used to estimate the variance. As the number of degrees of freedom gets larger, the shape of the t distribution approaches the shape of the normal distribution.

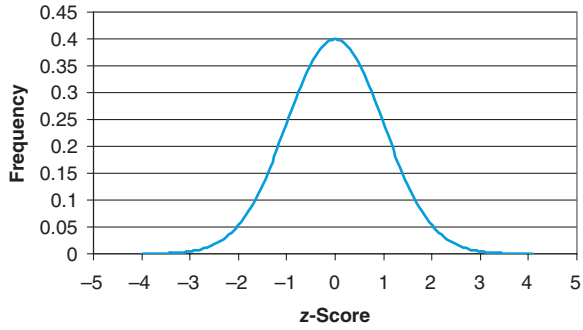


Figure 5.1
Frequency curve for a standard normal distribution with $\mu = 0$ and $\sigma = 1$.

Table 5.2
Percent of the area under the standard normal curve as a function of z*-score.*

±*z*-Score	% Area under the Curve
0.675	50
1.0	68.26
1.645	90
1.96	95
2.0	95.46
2.58	99
3.0	99.73

Table 5.3
Two-tailed probabilities from Student's t *distribution.*

Degrees of Freedom	*P* (Two-tailed probabilities)		
	0.1	0.05	0.01
1	6.314	12.706	63.657
5	2.015	2.571	4.032
7	1.895	2.365	3.499
10	1.812	2.228	3.169
15	1.753	2.131	2.947
19	1.729	2.093	2.861
20	1.725	2.086	2.845
40	1.684	2.021	2.704
60	1.671	2.000	2.660
120	1.658	1.980	2.617
Infinite	1.645	1.960	2.580

Listed in Table 5.3 are t statistics (analogous to z-scores) corresponding to 10%, 5%, and 1% of the area under the curve as a function of the number of degrees of freedom. In this table, the area under the curve corresponds to the two-tailed probabilities. Thus, for example, at $p = 0.1$ and $v = 5$, 90% of the area under the curve lies within ±2.015 standard deviations, 5% lies greater than 2.015 standard deviations, and 5% lies less than −2.015 standard deviations away. Note that for $v = \infty$ the normal distribution results are obtained. More extensive tables of t values are found in most statistics texts (for example, see Zar, 1984; Kachigan, 1991; or Gilbert, 1987).

Although most statistical tests are based on the assumption that the underlying distribution is normal, most environmental data appear to have frequency distributions that are log-normal. Two advantages of the log-normal distribution in describing environmental data are that it always gives positive values (there are no negative concentrations) and it can account for a small fraction of higher values (hot spot contamination) in the right side or "tail" of the curve. Ott (1990) has shown that a series of successive random dilutions

results in a distribution of values that is log-normal and thus provides a physical explanation for why environmental data often appear to be distributed this way.

By definition, data are said to be log-normally distributed if the log-transformed values are normally distributed. This means log-normal data can be analyzed using statistical techniques requiring normally distributed data if the data are first natural log-transformed. This is an important point because most parametric techniques require an assumption of normality or data that can be transformed to fit a normal distribution. The probability distribution function (PDF) for the log-normal distribution is given in Equation 5.10, where μ_y and σ_y are the mean and standard deviation of the transformed variable $y = \ln(x)$.

$$f(x) = \frac{1}{x\sigma_y\sqrt{2\pi}} e^{-\frac{(\ln x - \mu_y)^2}{2\sigma_y^2}} \tag{5.10}$$

The transformation step can introduce some complexities, as discussed by Gilbert (1987). For example, estimated quantities in the transformed scale (e.g., μ_y and σ_y) can lead to biased estimates when they are transformed back into the original scale (Wallace, 1997). Also, data presented in the transformed scale can be difficult to interpret. A number of authors have addressed these issues specific to the log-normal distribution (Land, 1975; Parkin *et al.*, 1988; Singh *et al.*, 1997).

5.2.3 CONFIDENCE LIMITS AND HYPOTHESIS TESTS

It is evident from the preceding discussion that with a given distribution assumption, it is possible to determine the probability of randomly selecting an object with a value less than or greater than a selected value (or between two selected values) from that distribution. For example, given a normal distribution with a mean of 10 and standard deviation of 2, we can determine the probability of randomly selecting an object with a value of 7 or less. Using Equation 5.9 we can calculate a z-score of $(7 - 10)/2 = -1.50$ and look up the one-sided probability corresponding to a z-score of –1.5. From a standard table of z-scores we find probability = 0.0668. Thus, there is a 6.68% probability of selecting an object with a value of 7 or less from a population of values fitting a normal distribution with a mean of 10 and standard deviation of 2.

Because of the Central Limit Theorem, we can make similar statements about the probability of estimating a mean value from a sample of size n

drawn from a normal distribution with a mean of 10 and standard deviation of 2. For example, what is the probability of calculating a sample mean greater than 11 in a sample of 8 objects drawn from that distribution? Using Equation 5.6, we can construct an expression similar to the expression in Equation 5.9.

$$z = \frac{(\bar{x} - \mu)}{\sigma_{\bar{x}}} = \frac{(\bar{x} - \mu)}{\frac{\sigma}{\sqrt{n}}} \tag{5.11}$$

Thus, using Equation 5.11 we can calculate a *z*-score of 1.41 and look up a corresponding probability of 0.0793. Thus, there is a 7.93 percent probability of calculating a sample mean of 11 or greater in a sample of eight objects randomly selected from a population of values fitting a normal distribution with a mean of 10 and standard deviation of 2. The significance of the Central Limit Theorem is that the underlying population doesn't have to be normal for this estimate to be valid, particularly if n is large.

5.2.3.1 Confidence Limits

This basic procedure can also be used to calculate symmetrical confidence limits about the true mean of the underlying population. For example, from Table 5.2 we know that 95% of the area under a standard normal curve is bound by $z = \pm 1.96$. If we plug this value into Equation 5.11 and solve for $\bar{x}$ we get

$$\bar{x} = \mu \pm 1.96\frac{\sigma}{\sqrt{n}} \tag{5.12}$$

Thus there is a 95% probability that a sample mean of 8 objects randomly selected from the underlying population will lie between 8.62 and 11.38 (i.e., 10 − 1.38 and 10 + 1.38). These upper and lower limits are called confidence limits, and in this case would be referred to as lower and upper 95% confidence limits about the mean. This is commonly written as:

$$P\left\{\mu - 1.96\frac{\sigma}{\sqrt{n}} \le \bar{x} \le \mu + 1.96\frac{\sigma}{\sqrt{n}}\right\} = 0.95 \tag{5.13}$$

But we never really know the true mean and true standard deviation (μ and σ) of a given population. In practice, a sample mean is calculated from a given data set and then confidence limits are determined. Then we can make the statement that there is a given probability that the true mean lies some-

where within these limits. Since we do not know the true population standard deviation, σ, we need to use the sample standard deviation, s, to calculate confidence limits. We also need to use the t distribution instead of the normal distribution to account for the uncertainty in our estimation of the standard deviation.

$$P\left\{\bar{x} - t\frac{s}{\sqrt{n}} \le \mu \le \bar{x} + t\frac{s}{\sqrt{n}}\right\} = p \tag{5.14}$$

If we continue with our example, the appropriate t statistic corresponding to a 95% probability level with $v = n - 1 = 7$ degrees of freedom is 2.365 (compared to a corresponding z-score = 1.96 if our sample contained all possible measurements). If we assume the sample standard deviation is 2.0, then we can say with 95% probability that the true mean lies somewhere between 8.33 and 11.67 (10 − 1.67 and 10 + 1.67). Note from Equation 5.14 that we can narrow our confidence limits about the mean simply by increasing the number of measurements in our sample.

Thus, confidence limits provide a measure of the uncertainty associated with estimates derived from data. Characterizing the levels of risk or contamination that people may have been exposed to is a common problem in forensics projects. The US EPA procedures for assessing human health risks at Superfund sites often are applied to this task. These procedures use the 95% upper confidence limit (UCL) on the mean as an estimate of the concentration within a defined exposure area. The 95% UCL is used to account for the uncertainty associated with estimates of the mean concentrations (US EPA, 1992).

Given a set of concentration data, confidence intervals and limits can be derived by a number of methods. The confidence interval formulas provided earlier are based on assumptions regarding the characteristics of the populations represented by the sample data; this subject is addressed in more detail in Section 5.2.3.3. If these assumptions are not consistent with the characteristics of the population or the data set, the confidence limits derived with these formulas will not be accurate—that is, the true likelihood that the unknown population mean is within the confidence interval may be quite different from the probability used to derive the interval. Although confidence intervals and limits derived using the z statistic (Equation 5.13) and the t statistic (Equation 5.14) are widely used, these intervals and limits are not exact unless the population of interest has a normal distribution.

As noted earlier, environmental data sets are commonly bounded at zero and positively skewed. These characteristics are consistent with log-normal distributions, and a formula developed specifically for deriving UCLs for log-

normal populations was recommended in earlier US EPA guidance (US EPA, 1992). The UCLs derived with this formula were often much higher than the highest observed values, which limited their utility; therefore, other methods of deriving UCLs for environmental data sets were considered.

A variety of methods of calculating the UCL of skewed data sets are discussed in articles in scientific journals (e.g., Schulz and Griffin, 1999; Porter *et al.*, 1997) and US EPA documents (e.g., US EPA, 2002a). US EPA has developed a software package called ProUCL that calculates UCLs by a number of methods, many of which are nonparametric (i.e., they are not based on assumptions about the distribution of the underlying population). Example calculations performed with this software indicate that most of the methods produce similar values when the data set is symmetric, and that the 95% UCL calculated using the *t* distribution (Equation 5.14) is not affected by substantial deviations from normality in small data sets.

Table 5.4 illustrates these points using three similar data sets. All three data sets contain eight observations with a sample mean of 10 and a standard deviation of 2. Set 1 is drawn from a symmetrical population; the data are consistent with a normal distribution and the skewness coefficient[4] is zero. Set 2 is drawn from a population that has a moderate amount of positive skew; the data set has a skewness coefficient of 0.61 and is consistent with a log-normal distribution.[5] Set 3 represents a population with a much higher degree of positive skew; the data set has a skewness coefficient of 2.02 and is not consistent with a log-normal distribution.

As shown in Table 5.4, the 95% UCL calculated with the *t* statistic is the same for all three data sets. This is expected; Equation 10.14 calculates the UCL as a function of the mean and standard deviation, which are the same for all three data sets. Note that the *t* statistic used to calculate the 95% UCL corresponds to a one-tailed probability. The UCLs calculated with the *H* statistic (which was recommended for log-normal distributions in US EPA 1992) get smaller and approach the *t*-statistic UCL as the degree of skewness increases. The nonparametric methods exhibit varying degrees of sensitivity to skewness, but all provide UCLs close to the *t*-statistic UCL when the data set is symmetric.

[4] The skewness coefficient is a measure of the symmetry of a distribution. Symmetrical distributions, such as the normal distribution, are not skewed and have a skewness coefficient of zero. Positively skewed distributions, such as the log-normal distribution, tail off to the right, and have skewness coefficients greater than zero. Negatively skewed distributions tail off to the left and have skewness coefficients less than zero.

[5] Hypothesis tests indicate that data sets 1 and 2 are consistent with both normal and log-normal distributions because these tests are not powerful enough to detect substantial deviations from these distributions when the data sets are small.

5.2.3.2 Hypothesis Testing

In hypothesis testing, the convention is to assume that there is no difference between two values (this is referred to as the *null hypothesis*, denoted H_0). This assumption holds unless we can determine that the probability of seeing such a large difference is so small that it is more likely that our initial assumption is wrong than it is that we have encountered this rare occurrence. The conven-

Table 5.4

Comparison of 95% upper confidence limits derived from three small data sets.

Summary statistics for three small data sets			
Statistic:	**Set 1**	**Set 2**	**Set 3**
Sample mean	10.00	10.00	10.00
Sample standard deviation	2.00	2.00	2.00
Sample skewness	0.00	0.61	2.02
Test of normality at 5%[a]	normal (p = 0.9308)	normal (p = 0.8130)	not normal (p = 0.0012)
Mean logarithm of sample	2.29	2.29	2.29
Standard deviation of logarithms	0.209	0.199	0.178
Skewness of logarithms	–0.648	–0.007	1.859
Test of log-normality at 5%	log-normal (p = 0.7740)	log-normal (p = 0.9368)	not log-normal (p = 0.0023)
95% Upper confidence limits (UCLs) developed by selected methods			
UCL method:	**Set 1**	**Set 2**	**Set 3**
t-statistic UCL[b]	11.34	11.34	11.34
H-statistic UCL	11.72	11.61	11.39
Jackknife UCL	11.34	11.34	11.34
Percentile bootstrap UCL	11.10	11.20	11.30
BCA bootstrap UCL	11.03	11.21	11.70
Bootstrap-*t* UCL	11.32	11.58	17.94

[a]Hypothesis tests on the distributions of the underlying populations were performed as Shapiro-Wilk tests using JMP version 3.2.6 (SAS Institute Inc., 1999). The *p*-values indicate the probability of error if the null hypothesis is rejected and the tests are interpreted by comparing the *p*-value to 0.05 (the five percent level of significance).

[b]All UCLs were calculated using ProUCL version 3.0 (US EPA, 2004), which also provides additional UCL methods that are not presented in this table.

tion is to conclude that our initial assumption is wrong when the probability associated with the null hypothesis is 5% or less.

When we establish our null hypothesis, it is also important to establish an alternate hypothesis (denoted H_a). For a null hypothesis of no difference between two sample means (i.e., H_0: $\mu_a = \mu_b$), there are three different alternate hypotheses: H_a: $\mu_a \neq \mu_b$; H_a: $\mu_a > \mu_b$; or H_a: $\mu_a < \mu_b$. The *a priori* decision of which alternate hypothesis is selected determines whether we are performing a one-tailed or two-tailed test. If our decision level is set at 5%, and our alternate hypothesis is H_a: $\mu_a \neq \mu_b$, then we don't care if $\mu_a > \mu_b$ or $\mu_a < \mu_b$. In this case, our 5% decision level is interpreted as a 2.5% chance of rejecting our null hypothesis when in fact $\mu_a > \mu_b$, and a 2.5% chance of rejecting the null hypothesis when in fact $\mu_a < \mu_b$. This is referred to as a two-tailed test. If however, we really would like to know if the average concentration on-site is greater than the average background level (i.e., $\mu_a > \mu_b$), and we are willing

to accept a 5% chance of wrongly rejecting our null hypothesis, then we would be performing a one-tailed test.

In making this decision to accept or reject the null hypothesis, two types of error are possible, Type I and Type II. A Type I error occurs when the null hypothesis is rejected (we cannot conclude that the values are drawn from the same population) when in fact they have been. A Type II error occurs when we accept the null hypothesis of no difference between means when in fact there is a difference.

The probability of committing a Type I error is equal to α (alpha), the significance level. Alpha is a value that is chosen by the investigator and usually set equal to 0.05. With alpha set equal to 0.05, we are willing to accept a 5% chance of rejecting our null hypothesis of no difference between means when in fact the null hypothesis is true. The probability of committing a Type II error is equal to β (beta). The quantity (I-β) is referred to as the power of a test and is more difficult to determine. It is a function of alpha, the standard error of the difference between the two means, and the size of the effect that we are trying to detect.

The power of the test is often neglected, which can lead to a false confidence in decision-making. For example, if $\alpha = 0.05$, and $\beta = 0.5$, there is a 5% chance of rejecting the null hypothesis when in fact it is true; but, there is a 50% chance of accepting the null hypothesis when it is false. For this reason, some statisticians believe that the only significant conclusion is when the null hypothesis is rejected (for example see Oakes, 1986; Reckhow *et al.*, 1990). If the null hypothesis cannot be rejected, the reason may be that the test had insufficient power to detect a difference (e.g., the sample size was too small or the data are highly variable). Thus, a poorly designed test based on too few measurements can be biased toward accepting the null hypothesis.

5.2.3.3 Normality, Representativeness, and Independence

The classical hypothesis tests and confidence intervals just described are based on assumptions regarding the characteristics of both the data set and the underlying population. When these assumptions are not met, the resulting hypothesis tests and confidence intervals are not exact. Deviations from assumptions regarding the underlying population are common in environmental data sets. For example, methods that use the z and t statistics are based on the assumption that the underlying population is normally distributed. As noted previously, environmental data sets are commonly bounded at zero and skewed to the right; in general, they do not appear to be consistent with a normal distribution. Methods based on the z and t statistics may be used even when the assumption of normality is not met, however, especially if the number of samples is large. This practice is justified by the Central Limit Theorem,

which states that the distribution of the sample mean approaches a normal distribution as the sample size increases. Alternately, nonparametric tests such as those described in Sections 5.3.1.2 and 5.3.1.5 can be used; these tests are valid even when the underlying population is not normally distributed. Methods that can be used to test whether a data set represents a particular type of distribution (such as normal or log-normal) are discussed in Section 5.3.1.1 and in many statistics books (e.g., Zar, 1984; Gilbert, 1987).

Assumptions regarding the characteristics of the data set often are ignored in environmental work. The parametric methods based on the z and t statistics and the common nonparametric alternatives described later in this chapter all require the assumption that the data set is composed of n independent observations that are representative of a single underlying population. In the statistics literature, this assumption is sometimes stated as a requirement that all the observations are independent and identically distributed random variables. This assumption ensures that each observation included in the data set provides the same amount of information regarding the underlying population. For a variety of reasons, many environmental data sets are not consistent with this assumption.

One common cause of this problem is biased sampling design. Many environmental data sets are collected to identify or characterize conditions in areas that are or may have been affected by releases of chemicals. The spatial distribution of the chemicals of interest frequently includes hot spots (i.e., small areas with relatively high concentrations); when the purpose of the investigation is to locate these features, the areas where the concentrations are highest are often overrepresented in the resulting data set. When the data set is used for other purposes, such as characterizing the mean concentration over a property or exposure area that extends well beyond the area affected by the chemical release, the data set may be less representative of the population of interest than it would be if the sampling locations were selected at random. The issue of bias in sampling designs is discussed in more detail in Section 5.3.1.6.

Another feature of environmental data sets that may lead to inconsistency with the underlying assumptions is the use of multiple observations collected from the same or closely spaced sampling locations. As an extreme example, consider that many environmental sampling programs include field duplicate samples. The data from both duplicates should not be treated as independent observations because the second observation at a sampling location generally provides much less information about the underlying population than a separate observation made some distance away. This problem is easily avoided by using the average of the duplicate sample values to represent their common sampling location (the fact that the resulting average is expected to be more

accurate than the single measurements obtained from other locations is usually ignored). Similarly, when the data set includes observations from multiple depths in the same borehole, the use of a single value that best represents the characteristic of interest (e.g., the average or the maximum value) for each sampling location provides a more realistic estimate of the number of independent observations.

Although the interdependence of data obtained from duplicate samples or multiple depth intervals is easily recognized, the correlation of observations made at adjacent sampling locations or at the same location over time often is ignored. Environmental variables often exhibit autocorrelation over distances that are greater than the distance between adjacent sampling locations. When this is true, knowledge of the value at one sampling location provides information regarding the values at other nearby sampling locations. Similarly, concentrations in groundwater made in the same well over time may be autocorrelated; knowledge of the concentration during one sampling event may provide information on the concentration at the next sampling event. As a result of spatial or temporal autocorrelation, the number of independent observations of the population of interest may be substantially lower than the number of observations available for statistical analysis. Methods for deriving confidence limits for correlated data sets are discussed in Section 11.12 of Gilbert (1987). Unless the effects of autocorrelation are addressed, confidence limits and hypothesis tests developed from autocorrelated data sets will appear to provide a greater level of certainty than the data actually will support.

5.3 APPLICATIONS IN ENVIRONMENTAL FORENSICS

5.3.1 COMPARING SAMPLE MEANS

Evidence that a facility may have contributed to the level of contamination at a site can be demonstrated by comparing data collected at the site with an appropriate set of background measurements. In comparing two data sets, it is important to keep in mind that it is not possible to prove that two data sets are different; rather, we are able to assess only the likelihood that they are drawn from the same population.

5.3.1.1 Student's *t*-Test

Perhaps the most common test for comparing samples is Student's *t*-test, also known as the *t*-test. The *t*-test is a parametric technique, which as previously

described means that some underlying distribution is assumed for the test. For the *t*-test (and most parametric techniques) it is assumed that each sample is drawn from a normal distribution. It is also assumed that the measurements in the sample were selected randomly from the distribution and that the variances of the two sampled populations are equal. (The *t*-test can also be used to compare a data set with a single value, for example to compare the mean of a data set to some clean-up level.)

In conducting a *t*-test for comparing means, the underlying hypothesis is that there is no difference between sample means. In other words, we establish the null hypothesis that the difference in the means is zero, that is, H_0: $\mu_1 = \mu_2$ or, H_0: $\mu_1 - \mu_2 = 0$. In most cases, the alternate hypothesis is that the means are not the same, that is, H_a: $\mu_1 \neq \mu_2$. The test statistic is given as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\frac{s_1}{\sqrt{n_1}} + \frac{s_2}{\sqrt{n_2}}} \tag{5.15}$$

If the population variances are equal, as they are assumed to be in the *t*-test, then we can calculate a pooled variance, s_p^2, and pooled standard deviation, s_p, where s_p = square root of s_p^2.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{5.16}$$

The variance ratio test or *F* test can be used to evaluate whether or not the variances are in fact equal; that is, whether or not it is likely that the two samples were drawn from populations with the same variance. We test the hypothesis H_0: $\sigma_1^2 = \sigma_2^2$ and assign some probability alpha, of committing a Type I error. The *F* statistic is calculated by simply dividing the larger variance by the smaller variance. This ratio is then compared to the tabular value of $F_{\alpha,v1,v2}$ for a given alpha and the degrees of freedom associated with each data set (tables of *F* values are found in most statistics texts). If we accept the null hypothesis and conclude no difference in variances, the *t* statistic can be calculated using the pooled standard deviation as indicated here:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \tag{5.17}$$

Table 5.5
Example data set used to demonstrate the t*-test.*

On-site Concentrations (ppm)	Background Concentrations (ppm)
353	86
300	111
288	306
459	111
345	346
324	142
473	73
339	217
236	149
310	417
359	111
420	584
357	68
581	263
762	171
315	176
406	266
279	146
504	155
233	81
299	336
371	106
452	80
384	101
480	178

An example data set to demonstrate application of the *t*-test is provided in Table 5.5. Summary statistics are listed in Table 5.6. The data consist of 25 measurements of lead in soil collected on-site (on-site data) and 25 measurements collected off-site, representing background measurements (background data).

A frequency distribution of the data is presented in Figure 5.2. The values on the *x*-axis represent the upper end of the bin range. Thus, the first bin range covers the values from 0–75, and the second runs from 76–150, and so on.

Test assumptions about normality and equal variance should be checked before performing the *t*-test. As indicated in Figure 5.2, both data sets appear to have a distribution that is skewed to the right, resembling a log-normal distribution. The coefficient of variation test, described by EPA (1989), provides a simple and quick check to detect gross nonnormality in the data set. The coefficient of variation is equal to the standard deviation divided by the mean. A value greater than 1.0 is evidence that the data are not normally dis-

	On-site Data	Background Data
Number of Samples	25	25
Minimum Value	233.4	67.8
Maximum Value	761.8	583.6
Sample Average	385	191
Sample Variance	13,515	15,925
Sample Standard Deviation	116	126
Standard Error of the Mean	23.3	25.2

Table 5.6
Summary statistics for the data presented in Table 5.4.

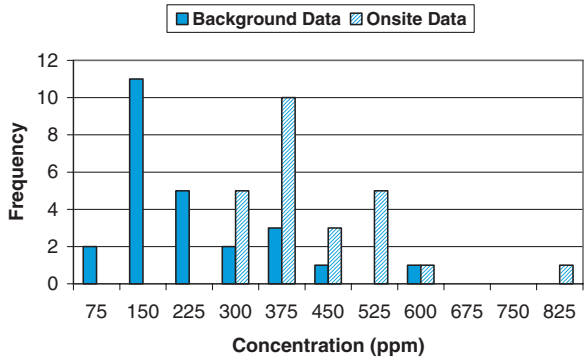


Figure 5.2
Frequency distribution for data presented in Table 5.5.

tributed. In this example, both coefficients of variation are less than 1.0 so there is no evidence of gross nonnormality.

A more powerful test that can detect whether or not a data set fits a specified distribution is the χ^2 (chi-square) goodness of fit test.[6] The χ^2 test compares observed frequencies in a sample data set with expected frequencies if the data set fit some known distribution (e.g., normal, log-normal, or some other specified distribution). The χ^2 statistic is calculated as

$$\chi^2 = \sum_{i=1}^{k} \frac{(N_i - E_i)^2}{E_i} \tag{5.18}$$

where k is the number of bin ranges, N is the observed number of measurements falling within a given bin range, and E is the expected number of measurements falling within the range. If there were no difference between the observed and expected frequencies, the value for χ^2 would be zero. As the calculated value for χ^2 increases, we become less certain that the distributions are similar. If the χ^2 statistic exceeds the tabular value (found in most statistics texts) for a given alpha level with k-3 degrees of freedom,[7] we must reject the null hypothesis of no difference between distributions.

[6] Note that there are other techniques for testing distribution assumptions such as the Kolmogorov Smirnov test and the Shapiro Wilk test (see Gilbert, 1987; Zar, 1984).

[7] Normally, the degrees of freedom for the chi-square test is equal to k-1; however, in this case we use k-3 rather than k-1 because we lose two additional degrees of freedom for estimating the mean and standard deviation when we convert the data to z-scores.

Table 5.7
Observed and expected frequencies calculated in the χ^2 test for normality for the data set in Table 5.4.

Bin Range for $k = 6$	Expected Frequency (E)	Observed Frequency (N) On-site	Observed Frequency (N) Background
≤0.97	4.17	2	1
−0.97−−0.43	4.17	7	9
−0.43−0	4.17	7	7
0−0.43	4.17	2	1
0.43−0.97	4.17	4	3
>0.97	4.17	3	4

The number of bin ranges should be selected so that E is at least 4. Thus, with $n = 25$, k should be less than $25/4 = 6.25$, so we set $k = 6$. The easiest way to perform this test by hand is to select bin ranges so that $E_1 = E_2 = E_3$ and so on. Thus, for $n = 25$ and $k = 6$, E is equal to $25/6 = 4.17$. EPA (1989) contains a table of appropriate bin ranges for a normal distribution that can be used for any data set if the data are first converted to z-scores. Bin ranges for $k = 6$ are presented in Table 5.7 along with the observed frequencies for the on-site and background data sets.

Using Equation 5.18, the calculated χ^2 statistics for the on-site and background data sets are 6.43 and 12.7, respectively. The tabular value for χ^2 with 3 degrees of freedom at the 95% probability level ($\alpha = 0.05$) is 7.815. Since the χ^2 statistic for the background data exceeds the tabular value, we cannot conclude that the data were drawn from a normal distribution.

When the data are found to be significantly different from normal, it may be possible to normalize the data by taking the natural log of each data point. This is a common transformation since most environmental data sets are adequately described by a log-normal distribution. Summary statistics for the log-transformed data are presented in Table 5.8. Observed and expected frequencies for the χ^2 test are found in Table 5.9.

Note that the bin ranges and expected frequencies in Table 5.9 remain the same if the test is performed on data that have been converted to z-scores. The calculated χ^2 statistics for the log-transformed on-site and background data sets are 1.16 and 2.11, respectively. The tabular value for χ^2 with 3 degrees of freedom at the 95% probability level is the same as before, 7.815. Since the χ^2 statistic for both the on-site and background data are less than the tabular value, we can conclude that the log-transformed data sets are not significantly different from normal. The frequency distributions of the natural log-transformed data are presented in Figure 5.3. Plotting the data is always recommended to get a feel for the data set. The on-site data clearly resemble a normal distribution; however, the background data still

	On-Site Data	Background Data
Number of Samples	25	25
Minimum Value	5.45	4.21
Maximum Value	6.64	6.37
Sample Average	5.92	5.08
Sample Variance	0.0751	0.345
Sample Standard Deviation	0.274	0.587
Standard Error of the Mean	0.0548	0.117

Table 5.8

Summary statistics for the natural log-transformation of the data presented in Table 5.4.

Bin Range for $k = 6$	Expected Frequency (E)	Observed Frequency (N) On-site	Observed Frequency (N) Background
≤0.97	4.17	5	3
–0.97–-0.43	4.17	5	6
–0.43–0	4.17	4	5
0–0.43	4.17	3	3
0.43–0.97	4.17	3	5
>0.97	4.17	5	3

Table 5.9

Observed and expected frequencies calculated in the χ^2 test for normality for the natural log-transformation of the data set in Table 5.4.

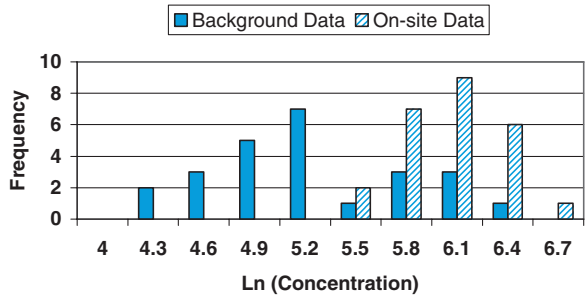


Figure 5.3

Frequency distribution for data presented in Table 5.4 after log transformation.

exhibit some deviation from normal, but nothing significant according to the χ^2 test.

The ratio of the variances in the log-transformed data set is 0.345/0.0751 = 4.59. This ratio can be compared to the tabular F value with $n - 1 = 24$ degrees of freedom for the numerator and denominator $F_{.05,\,24,\,24} = 2.27$. Since the variance ratio exceeds the tabular F value, we cannot conclude that the variances are equal at the 95% probability level. (This is clearly evident in Figure 5.3.) The scenario where both data sets are not significantly different from normal but the variances are unequal is referred to as the Behrens-Fisher problem (Zar, 1984). One solution to this problem is known as Cochran's approximation to the Behrens-Fisher problem (Zar, 1984; Ross, 1997). The t statistic is calculated according to Equation 5.19 as illustrated.

$$t = \frac{5.92 - 5.08}{\frac{0.274}{\sqrt{25}} + \frac{0.587}{\sqrt{25}}} = \frac{0.84}{0.172} = 4.88 \tag{5.19}$$

However, the tabular value is Student's t with degrees of freedom calculated for the combined data set according to Equation 5.20.

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \tag{5.20}$$

In this case, $v = 17.7$ (or 17, when rounded to the next smallest integer). The tabular value for Student's t with alpha = 0.05 and 17 degrees of freedom is 2.11. Since the calculated value of 4.88 exceeds the tabular value of 2.11, we must reject the null hypothesis and conclude that there is less than a 5% probability that the observed difference between the two means could occur by chance. This is strong evidence that lead concentrations on-site are higher than background levels and suggests that the facility has indeed affected the soils on-site. In addition, the t-test has allowed us to arrive at this conclusion with a quantitative estimate of the confidence in our conclusion.

5.3.1.2 Wilcoxon's Rank Sum Test

Data sets that cannot be transformed to approximate a normal distribution can be compared using a nonparametric technique such as Wilcoxon's Rank Sum test (also known as the Mann Whitney U test). With a nonparametric test, there is no requirement that the data fit some underlying distribution. Wilcoxon's Rank Sum test is designed to test the null hypothesis that the two data sets are drawn from the same distribution (whatever it is), or similar distributions with the same central tendency.

The test is conducted by combining the data points from the two samples, while maintaining their identity, and ranking the data from lowest to highest (or highest to lowest). Ties (data points with the same value) are assigned the average of the ranks they would have been assigned had they not been tied. Gilbert (1987) notes an advantage of the Wilcoxon Rank Sum test is that it can handle a moderate number of nondetect values by treating them as ties.

The sum of the ranks from each sample (R_1 and R_2) are tallied and the smaller sum is selected for comparison (n_R is the number of data points associated with the sample with the smaller sum). For samples with $n \leq 20$, R may be compared directly to a table of R values (for example see Langley, 1970).

For samples with more than 20 data points, the significance of R can be determined by computing Z_{rs}.

$$Z_{rs} = \frac{R - \frac{n_R(n_1 + n_2 + 1)}{2}}{\sqrt{\frac{n_1 \cdot n_2(n_1 + n_2 + 1)}{12}}} \tag{5.21}$$

Because R is approximately normal for large n, all we really are doing is calculating a standard normal deviate analogous to Equation 5.9 where:

$$Z_{rs} = \frac{R - E(R)}{\sqrt{Var(R)}} \tag{5.22}$$

the expected value of R, $E(R)$ is given as

$$E(R) = \frac{n_R(n_1 + n_2 + 1)}{2} \tag{5.23}$$

and the variance $Var(R)$ is estimated using the following equation (Reckhow *et al.*, 1990):

$$Var(R) = \frac{n_1 \cdot n_2(n_1 + n_2 + 1)}{12} \tag{5.24}$$

Gilbert (1987) also gives an equation for calculating Z_{rs} when ties are present:

$$Z_{rs} = \frac{R - \frac{n_R(n_1 + n_2 + 1)}{2}}{\left\{\frac{n_1 n_2}{12}\left[n_1 + n_2 + 1 - \frac{\sum_{j=1}^{g} t_j(t_j - 1)^2}{(n_1 + n_2)(n_1 + n_2 - 1)}\right]\right\}} \tag{5.25}$$

where g is the number of tied groups, and t_j is the number of tied data in the jth group.

In our example data set presented in Table 5.5, the smaller sum of ranks is equal to 396 and there is one group of ties with $t = 3$. The resulting value of Z_{rs} calculated using Equation 5.25 is 4.68—corresponding to a <0.001 prob-

ability that the two data sets are drawn from the same underlying distribution. The results of this test confirm our earlier conclusion based on Student's *t*-test.

Higgs *et al.* (1999) used the rank sum test to compare patterns of soil lead contamination observed on elementary school properties. The authors found that soils on inner-city school properties had significantly higher lead concentrations than soils on outer-city school properties, consistent with earlier findings that lead tends to cluster within the interior of the largest cities.

5.3.1.3 Paired *t*-Test

If paired data sets are available, for example, upwind–downwind measurements, then the paired *t*-test can be used. In the paired *t*-test, the null hypothesis is that the population mean of the differences between each data pair (μ_d) is equal to zero (i.e., H_0: $\mu_d = 0$), where d_i is the difference between the ith data pair and $\overline{d}$ is the sample mean of the differences between each data pair. The paired *t*-test assumes that the values of d_i are normally distributed. The test statistic is calculated as

$$t = \frac{\overline{d}}{\frac{s_d}{\sqrt{n}}} \tag{5.26}$$

where:

$$s_d = \sqrt{\frac{\sum (d_i - \overline{d})^2}{n-1}} \tag{5.27}$$

The calculated t statistic is compared with the tabular t value with $n - 1$ degrees of freedom. If the calculated value exceeds the tabular value then the null hypothesis is rejected.

Daily upwind–downwind total suspended particulate (TSP) measurements collected at the site of a former gold mine in California during the months of July, August, and September are presented in Table 5.10. The downwind–upwind difference between each data pair reflects a contribution from the site. A scatter plot of downwind–upwind measurements for the three months (see Figure 5.4) indicates that downwind TSP concentrations are generally greater than upwind measurements.

A frequency distribution plot of the downwind–upwind difference concentrations is shown in Figure 5.5. Note that the first bin includes data in the

Day	July		August		September	
	Upwind	Downwind	Upwind	Downwind	Upwind	Downwind
1	56	116	59	112		
2	64	78			84	68
3	73	86			75	56
4			53	85	78	69
5			66	165	82	53
6			68	167		
7	36	53	58	145		
8	45	98	56	95	166	56
9	72	115			105	73
10	62	70			84	63
11					90	39
12					180	63
13						
14	51	147	89	148		
15	63	171	75	161	87	38
16	56	157			105	46
17	66	150			35	38
18	47	125	48	78	67	36
19			55	53		
20			53	137		
21	125	113	55	103		
22	51	88	54	76	123	36
23	50	50			93	108
24	32	118			45	15
25	41	169	61	75	69	33
26			65	102	83	36
27			61	55		
28	53	166	62	82		
29	59	75	57	68	41	34
30	86	76				

Table 5.10
Upwind–downwind TSP concentrations ($\mu g/m^3$) for July, August, and September at a former gold mining site in California.

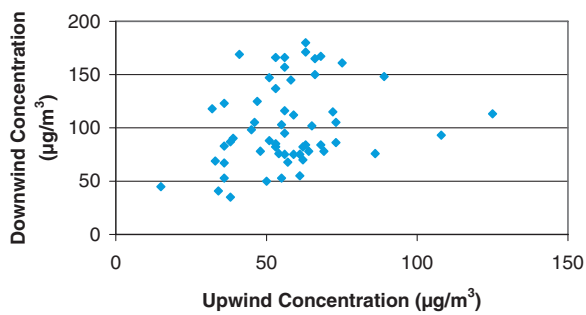


Figure 5.4
Scatter plot of upwind–downwind TSP measurements for July, August, and September.

range of −20 to 0 and represents data pairs where upwind concentrations exceed downwind values.

Once again, we can test the assumption of normality with a χ^2 test. Using Equation 5.18, and dividing the 57 difference measurements into 8 equal bins, the calculated χ^2 statistic is 6.30 compared to a tabular value of 11.07 for $\alpha = 0.05$ and $v = 5$. Thus, we can accept our null hypothesis of no difference

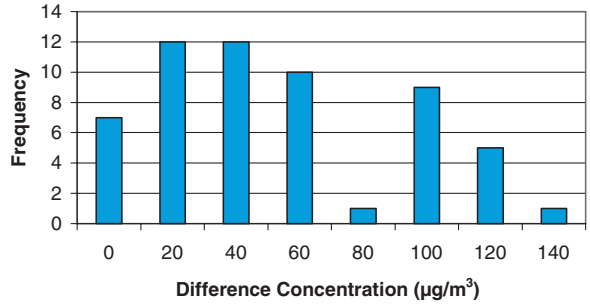


Figure 5.5
Frequency distribution of downwind–upwind difference concentrations.

between the observed distribution of difference values and a normal distribution. We can now proceed with the paired *t*-test. The average downwind–upwind difference between the 57 data pairs is 45.2 and the standard deviation is 38.2. Thus using equation 5.26, we calculate $t = 8.94$.

In comparison, the tabular value for $t_{.05,56} = 2.003$, so we reject our null hypothesis that the mean difference of 45.2 is not significantly different than zero. We can take this analysis one step further and calculate confidence limits about the mean difference using Equation 5.14.

$$P\left\{45.2 - 2.003\frac{38.2}{\sqrt{57}} \le \mu_d \le 45.2 + 2.003\frac{38.2}{\sqrt{57}}\right\} = 0.95 \tag{5.28}$$

Thus there is a 95% probability that the true mean difference for the months of July, August, and September lies somewhere between 35.1 and 55.3 $\mu g/m^3$. Note that only three months out of the year have been selected for this example. This analysis could also be performed for all the available data. If the investigation was being conducted in support of a toxic tort case, then the time period of exposure could be analyzed. It is also possible that exposure could be greater in certain months. This hypothesis could be tested using an analysis of variance technique.

5.3.1.4 Analysis of Variance (ANOVA)

If we wish to compare more than two samples, for example the months of July, August, and September, we could perform multiple *t*-tests. However, as the number of samples to be tested increases, the number of possible comparisons becomes large. For example, comparing five samples would require $5(5 - 1)/2 = 10$ *t*-tests; and 10 samples would require 45 *t*-tests. Moreover, as we perform more tests, we increase our chances of committing a Type I error, even though each test might be conducted at the 0.05 level. This is because, at the 0.05 level, we would expect to observe a significant difference between samples in 1 out of 20 tests strictly by chance.

	July	August	September
Minimum	–12	–6	–15
Maximum	128	99	117
Mean	51.65	45.11	38.52
Variance	2013	1155	1238
Number of Samples	20	18	19

Table 5.11
Summary statistics grouped by month for downwind–upwind difference measurements.

Analysis of Variance (ANOVA) is a technique that can be used to compare the means of two or more samples. (If there are only two samples, then ANOVA is identical to the two-sample *t*-test.) It is a parametric technique so all the same assumptions apply. It is assumed that the data are drawn at random from a normal population, and that sample variances are equal.

The one-way analysis of variance technique involves, just as the name implies, an analysis of the variability or variance in a set of measurements. The test usually is conducted on experimental data designed to test the effect of various treatments on the outcome of a *single* variable (hence the one-way). In this case, the null hypothesis is that the means for each treatment group are the same (i.e., H_0: $\mu_1 = \mu_2 = \mu_3$, etc). In our example, we could test the hypothesis that there was no significant difference in operations at the facility between the months of July, August, and September that would affect the contribution of TSP to ambient air. To test the hypothesis, we determine if the variability between months is any greater than the variability within each month of measurements. The difference measurements, grouped by month, are summarized in Table 5.11.

Two estimates of the population variance are calculated. One estimate is called the *within groups variance*, also called the error variance. It is simply a pooling of the variances within each group, analogous to the pooled variance described in Equation 5.16. In this case, the pooled variance is equal to:

$$\textit{Error Variance} = \frac{19(2013)+17(1155)+18(1238)}{19+17+18} = 1484 \qquad (5.29)$$

The other estimate is called the *between groups variance* and is estimated by comparing the mean of each group, $\overline{x}_i$, with the grand mean, $\overline{x}$, for all data points:

$$\textit{Between Groups Variance} = \frac{\sum_{i=1}^{k} n_i(\overline{x}_i - \overline{x})^2}{k-1} \qquad (5.30)$$

Table 5.12
Analysis of variance table for the null hypothesis of no difference among average downwind–upwind difference measurements for the months of July, August, and September.

Source of Variation	SS	DF	MS	F
Between Groups	1,680	2	840	0.57
Within Groups (Error)	80,136	54	1,484	
Total	81,816	56		

In our example the between groups variance is equal to

$$\frac{20(51.65-45.21)^2+18(45.11-45.21)^2+19(38.52-45.21)^2}{2}=840 \qquad (5.31)$$

If the null hypothesis is true (i.e., H_0: $\mu_1=\mu_2=\mu_3$), then the between groups variance should be essentially equal to the within groups variance and the ratio of these two quantities (between groups/within groups) should be close to one. If the group means are not equal, then the between groups variance will be larger than the within groups variance and the ratio will exceed one. The one-tailed variance ratio test or F test with (k-1) and (N-k) degrees of freedom is used to compare variances. A one-tailed test is performed because our alternate hypothesis is that between groups variance is greater than the within groups variance, not just that it is different.

In our example, the calculated F statistic is equal to 840/1484 = 0.57. Thus the variance between groups is smaller than the variance within groups. The tabular value for F with $\alpha = 0.05$, 2 degrees of freedom in the numerator and 54 degrees of freedom in the denominator is 3.17. Thus we accept the null hypothesis of no difference between means.

In performing an ANOVA, the convention is to construct an ANOVA table where the variance is separated into two parts, the sum of squares (*SS*) and the degrees of freedom (*DF*). The between groups *SS* is equal to the numerator in Equation 5.30 and the degrees of freedom is equal to the denominator. The within groups *SS* is defined in Equation 5.32.

$$\textit{Within Groups } SS=\sum_{i=1}^{k}\left[\sum_{j=1}^{n}(x_{ij}-\bar{x}_i)^2\right] \qquad (5.32)$$

and the degrees of freedom is equal to $N-k$. An ANOVA table for the TSP example is illustrated in Table 5.12.

Finally, we should check our assumptions of equal variance using the F test. The ratio of the larger to smaller variance is 2013/1155 = 1.74. Compared to the tabular value of F = 2.63 for α = 0.05 and with 19 and 17 degrees of freedom, we can accept our null hypothesis of no difference. Since the vari-

ances are constant, we can use the pooled, within groups variance for estimating confidence limits about each of the group means. Thus, for example, the 95% confidence limits about the mean downwind–upwind difference observed in July is

$$P\left\{51.6 - 2.093\frac{\sqrt{1484}}{\sqrt{20}} \le \mu_d \le 51.6 + 2.093\frac{\sqrt{1484}}{\sqrt{20}}\right\} = 0.95 \qquad (5.33)$$

where the value 2.093 is the Student's two-tailed t statistic for $\alpha = 0.05$ and 19 degrees of freedom. Thus we can say that there is a 95% probability that the true mean difference in July lies between 33.6 and 69.6 $\mu g/m^3$.

If the calculated F statistic exceeded the tabular value, we would reject our null hypothesis and conclude that there is a significant difference between groups; but, we would not know which groups are different. There are procedures like the Tukey multiple comparison test that can be used to determine which groups are different. The Tukey test is discussed by Zar (1984) and Berthouex and Brown (1994).

Boehm *et al.* (1998) used analysis of variance to compare polycyclic aromatic hydrocarbons (PAHs) in sediments from different locations. The authors analyzed the distribution of PAHs in sediments collected from Prince William Sound, Alaska to distinguish petrogenic, pyrogenic, and biogenic sources. PAH levels in sediments collected from the Bay of Isles, a bay affected by the *Exxon Valdez* oil spill, were then compared to PAH levels in sediment samples collected from a bay receiving minimal impact from the spill using analysis of variance techniques. The authors were able to show that there is no evidence of large-scale offshore transport of spill oil to the subtidal sediments.

The ANOVA is robust to small deviations from equal variance and normality assumptions (ZAR, 1984). However, if the deviations are significant, there are nonparametric techniques available such as the Kruskal-Wallis test.

5.3.1.5 Kruskal-Wallis Test

The Kruskal-Wallis test is similar to Wilcoxon's Rank Sum test in that we are comparing the sum of ranks applied to the data. The test statistic is calculated as

$$K = \frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i} - 3(N+1) \qquad (5.34)$$

where R_i is the sum of ranks for the ith group. For the TSP example, the sum of ranks for July, August, and September are 616, 533, and 504, respectively,

and the calculated K value is 0.68. There is a correction for ties. The correction factor, C, is given in Equation 5.35.

$$C = 1 - \frac{\sum_{j=1}^{g} t_j^3 - t_j}{N^3 - N} \tag{5.35}$$

where g is the number of tied groups, and t_j is the number of tied data in the jth group. The value of K corrected for ties, K_c, is equal to K/C. For large data sets (large N), the correction factor is minimal (in our example 0.9996 with $g = 10$ and $t_j = 2$). For larger samples, the calculated K statistic is compared to the tabular value for χ^2 with k-1 degrees of freedom. At $\alpha = 0.05$ and $v = 2$, $\chi^2 = 5.991$. Thus, as expected, we arrive at the same conclusion of no difference between average downwind–upwind difference measurements for the months of July, August, and September.

The most common application of these parametric and nonparametric techniques is for the comparison of concentrations on-site with background levels. Spatial and temporal variations in background can complicate the analysis but these issues can be addressed with proper sampling design and modifications to these basic procedures. Numerous techniques have been published (for example see Liggett, 1984; Gilbert and Simpson, 1990).

In addition to comparing on-site levels with background, these statistical techniques can also be used to evaluate cost allocation schemes if it is claimed that two distinct plumes are present (by comparing distributions of constituents in each plume, for example), or test natural resource damage claims if it is claimed that there are a statistically significant fewer number of wildlife species in a given area.

5.3.1.6 Issues in Sampling Design

In many forensics cases, all the data that can be used to address the questions of interest have already been collected. Although this suggests that sampling design issues are not relevant, an understanding of sampling design issues allows the forensic investigator to make the most of the available data. Therefore, sampling design issues that may limit the utility of an available data set are discussed in this section.

The goal of most statistical sampling designs is to provide independent observations that are representative of the populations of interest in sufficient number to allow reliable tests of the hypotheses of interest. A common question in sampling design is, how many samples are needed to perform a hypothesis test with the desired error rates? The answer depends on many factors. The most important factors in determining the number of samples needed

for comparisons of mean values are the variances of the populations, the size of the difference in means to be detected, and the acceptable levels of the two types of potential errors that may occur in hypothesis testing (as discussed in Section 5.2.3.2). Sampling design methods for environmental projects are reviewed and discussed in a number of USEPA documents (e.g., USEPA, 2000; USEPA, 2002b) as well as in Gilbert (1987).

Biased sampling is commonly encountered in forensics projects where the data have already been collected. When the bias is recognized, it is often possible to select or subdivide the data set in ways that allow the questions of interest to be addressed without bias. If the bias is not recognized, the results of the hypothesis tests (or other statistical calculations) may be erroneous. An example that illustrates these points involves the levels of metals in soils in the community around a smelter. The data set is biased because many of the data were obtained from a site investigation study that was designed to delineate and characterize a contaminated zone; the sampling density in this zone is much higher than the density over the rest of the community. When hypothesis tests are used to compare the mean of all the concentration data from the smelter community to background levels, a significant difference is detected. The data indicate, however, that the contaminated zone is a small and distinct portion of the larger community. Once this is recognized, the data set can be divided into two subsets (one that represents the contaminated zone immediately adjacent to the smelter and another that represents the rest of the community). Hypothesis tests on these subsets show that the levels of metals in soil in the contaminated zone are elevated, whereas the levels throughout most of the community are not significantly greater than background.

Another sampling design issue that is often encountered in environmental work involves the independence of the available observations. As discussed in Section 5.2.3.3, many environmental variables are spatially and/or temporally persistent; the value at one location is closely related to the values at nearby locations, and less closely related to the values at locations farther away. If the observations are related, the amount of information they provide is lower than the amount that would be available from the same number of independent observations. If this problem is ignored, the level of confidence associated with the results of the hypothesis tests is usually erroneous. Alternately, if the lack of independence is recognized, it may be compensated for by performing hypothesis tests on a selected subset composed of independent observations or by using methods designed for autocorrelated data.

One situation that illustrates these sampling design issues involves blood lead levels. Suppose a study based on voluntary sampling produces blood lead values from a large number of people in a smelter community, and similar

sampling is conducted in a control community. The question of interest is whether blood lead levels have been affected by the smelter, and this question will be addressed by hypothesis testing. The null hypothesis is that the mean blood lead levels in the two communities are equal; the alternate is that the mean blood lead level in the smelter community is higher than that in the control community.

Because participation in the study is voluntary, the possibility of biased sampling must be considered. The demographic characteristics of people from the control community who choose to participate in the study may be quite different from those of the smelter community participants, especially if the possibility of health effects is well publicized. Blood lead levels are related to demographic characteristics such as age, so the groups used in the hypothesis tests should be closely matched. In addition, the smelter community group may include individuals who are exposed in ways that do not occur in the control community (such as occupational exposure at the smelter). Furthermore, the independence of the available measurements is questionable if many of the individuals whose blood was sampled live with other individuals whose blood was sampled. Cohabitation implies exposure to the same levels of lead in common sources such as dust, water, air, and diet. Thus, if one member of a household has an elevated blood lead level for his/her age group, other members of the household are also likely to have elevated levels. This suggests that the number of independent observations in the data set is somewhat smaller than the number of measurements. If recognized, these sampling design issues (i.e., potential bias and lack of independence) can be accounted for by selective subsampling. Ignoring these issues may lead to conclusions that are not supported by the data when more appropriate hypothesis tests are performed.

5.3.2 LINEAR CORRELATION AND LINEAR REGRESSION ANALYSIS

Thus far we have discussed techniques for comparing a single variable—for example, lead in soil, TSP in ambient air, and so on—between or among samples. However, most analytical techniques are designed to measure multiple constituents at once, such as volatile organics in air or metals in soil. Simple techniques for evaluating the relationships among multiple variables in a sample are discussed in this section. The two basic techniques are correlation and regression. Other more sophisticated techniques like Principal Components Analysis and Receptor Modeling are related to these basic techniques and discussed elsewhere in the book.

There are slight but important differences between linear regression and linear correlation. Simple linear regression seeks to quantify the functional

dependence of one variable (the dependent variable, usually designated y) on another (the independent variable, usually designated x). The relationship is assumed to be linear; and for each value of x there is assumed to be a normal distribution of y values. A good example of regression analysis is the construction of an analytical calibration curve where the functional relationship between sample concentration and instrument response is quantified. In this case, the instrument response (the dependent variable) depends upon the concentration of analyte (independent variable) sampled by the instrument.

In contrast, linear correlation analysis is used to assess the linear association between two independent variables. In correlation analysis it is assumed that *both* variables are drawn from underlying normal distributions. This distinction does not preclude one from performing a regression analysis on two independent variables—in fact, this is often done. However, one must be careful about interpretation of the results.

5.3.2.1 Linear Regression Analysis

The general equation for a straight line is given in Equation (5.36)

$$y_i = \alpha + \beta x_i \tag{5.36}$$

where α and β are both population parameters, and α is the y intercept and β is the slope of the line. In any given data set, there are a number of straight lines that can be drawn through the data. Rarely do all the values fit on a single straight line. In regression analysis, the best fit line through the data is determined, where best fit line is defined as the line that minimizes the square of the deviations between the y_i values predicted by the equation for the line (denoted by $\hat{y}_i$) and the observed y_i values, that is,

$$\sum_{i=1}^{n} (\hat{y}_i - y_i)^2 = Minimum \tag{5.37}$$

Thus, we are trying to minimize the vertical deviations between the observed data points and the fitted line. If we could sample and determine all values of x_i and y_i we could determine both α and β. Since this is rarely possible, we estimate the slope and intercept (denoted b and a, respectively) using the following equations:

$$b = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{5.38}$$

Table 5.13
Construction of a regression analysis table.

Source of Variation	*SS*	*DF*	*MS*
Total	$\sum_{i=1}^{n}(y_i-\bar{y})^2$	$n-1$	
Regression	$\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2$	1	$\frac{Regression\ SS}{Regression\ DF}$
Residual	$\sum_{i=1}^{n}(\hat{y}_i-y_i)^2$	$n-2$	$\frac{Residual\ SS}{Residual\ DF}$

and

$$a = \bar{y} - b\bar{x} \tag{5.39}$$

The uncertainty in our estimates of the slope and intercept can be addressed using an analysis of variance approach. We are actually testing the null hypothesis that the true slope is zero, H_0: $\beta = 0$. A regression analysis table is constructed where the variances are broken down into a sum of squares (*SS*) term and a degrees of freedom (*DF*) term. Construction of a regression analysis table is illustrated in Table 5.13.

The analysis is similar to that in ANOVA except that instead of dealing with discreet groups, the variable x is continuous. If the slope is zero, this is the same as saying that there is no difference between group means for an infinite number of groups. Once again, we have two variance estimates, the regression *MS* and the residual *MS*. The one-tailed *F* test with (1) and $(n-1)$ degrees of freedom is used to compare the two estimates. If there is any slope, positive or negative, the contributions from both ends of the fitted line will increase the regression *SS* term and the ratio will exceed one.

The coefficient of determination, r^2, is a measure of the percentage of the total variation in y that can be explained by the regression line. The coefficient of determination is equal to the regression *SS* divided by the total *SS* as indicated in Equation 5.40.

$$r^2 = \frac{Regression\ SS}{Total\ SS} \tag{5.40}$$

The residual *MS* is the variance of y after taking into account the dependence of y on x. It is denoted $s^2_{y \cdot x}$. The square root of this term is known as the standard error of the regression, or standard error of the estimate. It is a measure of how well the fitted equation predicts the dependence of y on x.

The standard error of the slope, s_b, can be calculated from the standard error of the estimate according to Equation 5.41.

$$s_b = \sqrt{\frac{s_{y \cdot x}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}} \tag{5.41}$$

From the standard error of the slope, we can estimate confidence intervals about the true slope.

$$P\{b - t_{\alpha,n-2} s_b \leq \beta \leq b + t_{\alpha,n-2} s_b\} = p \tag{5.42}$$

This is an important result because in many environmental forensic investigations we are interested in discovering trends in the data. Trends in concentration that increase or decrease over time or distance can usually be evaluated using simple linear regression techniques. For example, contaminant concentrations that decrease over time at a rate that is proportional to the concentration are said to follow first order kinetics. Most natural attenuation processes in the environment such as biodegradation, photolysis, and volatilization are modeled assuming first order or pseudo first order kinetics. In one example, Kaplan *et al.* (1997) describe how the ratio of (benzene + toluene)/(ethylbenzene + xylenes) can be used to monitor changes in a dissolved gasoline plume over time. The ratio decreases over time due to greater solubility and mobility of B + T compared to E + X. At one site, the decrease over time was modeled as a pseudo first order process with a half-life estimated to be 2.3 years.

The basic equation is:

$$C = C_0 e^{-kt} \tag{5.43}$$

where

C = contaminant concentration at time t
C_0 = contaminant concentration at $t = 0$
k = first order rate constant (t^{-1})

Equation 5.43 can be rearranged into a form resembling Equation 5.36 by taking the natural log of both sides of the equation. The resulting expression is:

$$\ln C = \ln C_0 - kt \tag{5.44}$$

where the slope is equal to the negative of the first order rate constant and the y intercept is equal to the natural log of the concentration at time $t = 0$. The half-life (the time required for the concentration to be decreased by one-half) is equal to the value 0.693 divided by the first order rate constant.

$$t_{1/2} = \frac{\ln(2)}{k} = \frac{0.693}{k} \tag{5.45}$$

A common practice in forensic investigations is to estimate the date of release of a contaminant to the environment based on half-life values reported in the literature. If sufficient data exist, site data may be evaluated using linear regression techniques to establish a site-specific rate constant. In addition, we can place confidence limits about our estimate to evaluate the uncertainty in release dates. If the confidence limits about the slope include zero, then the trend is not significant.

5.3.2.2 Pearson's Product Moment Correlation Coefficient

Pearson's product moment correlation coefficient, r, also referred to as simply the correlation coefficient, is a dimensionless value that can range from –1 for a perfect negative linear correlation to +1 for a perfect positive linear correlation. A value of zero indicates no linear relationship between variables. It is important to remember that the correlation coefficient is a measure of *linear* association between two variables. Data arranged in a perfect circle are clearly associated with each other, but these data would yield a correlation coefficient equal to zero. The correlation coefficient is a parametric statistic and it is assumed that both variables are drawn from an underlying normal distribution.

The correlation coefficient is calculated as

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{5.46}$$

The correlation coefficient can also be written as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n} z_x z_y}{n-1} \tag{5.47}$$

In this form, it is apparent that the correlation coefficient is like an average of the products of paired z-scores (Kachigan, 1991). It is also easy to see that

when positive *z*-scores are paired with positive *z*-scores and negative *z*-scores are paired with negative *z*-scores, the correlation coefficient will be positive indicating a positive correlation. And, when positive *z*-scores are paired with negative *z*-scores, the correlation coefficient will be negative indicating a negative correlation.

Note that the correlation coefficient is also related to the slope of the best fit line through the data.

$$r\frac{s_y}{s_x} = b \tag{5.48}$$

The square of the correlation coefficient is called the coefficient of determination and reveals the proportion of the variance in *y* that can be predicted from *x* or vice versa, the proportion of the variance in *x* that can be predicted from *y*.

Peers (1996) notes that the correlation coefficient can be affected by the presence of outliers, and suggests that it is most appropriate when *n* is greater than 30. In calculating *r*, we have estimated the population correlation, ρ, which we could calculate using Equation 5.46 if we knew all values of *x* and *y*. We can measure the confidence in our estimate by either testing the null hypothesis that the two variables are independent (i.e., H_0: $\rho = 0$) or calculating confidence limits about *r*. The test statistic for the null hypothesis is calculated according to Equation 5.49 and compared to the *t* value with *n*-2 degrees of freedom.

$$t = \frac{r\sqrt{n-2}}{\sqrt{(1-r^2)}} \tag{5.49}$$

Confidence limits on *r* can be estimated by first transforming the correlation coefficient to Fisher's *z* value (which is different than the normal distribution *z* value) where:

$$\text{Fisher's } z = 0.5\ln\frac{(1+r)}{(1-r)} \tag{5.50}$$

and the standard error (*SE*) is given as:

$$SE = \frac{1}{\sqrt{n-3}} \tag{5.51}$$

The 95% confidence limits are Fisher's $z \pm 1.96$ (*SE*); where the 1.96 is derived from the standard normal curve (i.e., 95% of the standard normal

Table 5.14
Correlation matrix for BTEX concentrations detected in indoor air using natural log-transformed data.

	Benzene	Toluene	Ethylbenzene	m/p-Xylene	o-Xylene
Benzene	1				
Toluene	0.82	1			
Ethylbenzene	0.61	0.5	1		
m/p-Xylene	0.56	0.44	0.99	1	
o-Xylene	0.6	0.48	0.96	0.97	1

curve lies within ±1.96 standard deviations). Once the confidence limits (*CL*) about Fisher's *z* value are calculated, they must be converted back to correlation coefficients using the formula:

$$r = \frac{e^{2CL} - 1}{e^{2CL} + 1} \tag{5.52}$$

The correlation coefficient is a useful tool for exploratory data analysis. For example, Kimbrough and Suffet (1995) calculated correlation coefficients for metals (lead, antimony, arsenic, cadmium, chromium, copper, and nickel) in air, soil, and dust near two secondary lead smelters in California in order to determine those metals best suited for establishing characteristic ratios associated with impacts from the operation. Lead, antimony, and arsenic were selected and ratios of these three species were calculated in on- and off-site samples to determine the source of off-site concentrations. A similar analysis was performed by Amter and Eckel (1996).

When multiple species are compared, it is useful to prepare a correlation matrix. It is prepared by calculating correlation coefficients for each possible pair of variables and displaying them in a matrix as shown in Table 5.14.

The correlation matrix displayed in Table 5.14 was prepared from residential indoor air samples collected at a site in Louisiana where a prior spill had left a significant quantity of jet fuel floating on the water table. Concentrations of benzene, toluene, ethylbenzene, and xylenes (BTEX chemicals) detected in indoor air in some homes caused regulators to suspect the floating product on the water table as a source.

The highest correlation coefficients were reported for the association of ethylbenzene with the xylenes, and for the association of o-xylene with m/p-xylene. Benzene concentrations were most closely correlated with toluene. Soil gas data were also collected at the site. Significant correlation coefficients were also recorded for BTEX chemicals in soil gas as indicated in Table 5.15.

	Benzene	Toluene	Ethylbenzene	Xylenes
Benzene	1			
Toluene	0.66	1		
Ethylbenzene	0.52	0.94	1	
Xylenes	0.47	0.86	0.88	1

Table 5.15

Correlation matrix for BTEX concentrations detected in soil gas using natural log-transformed data.

Statistic	Benzene/Toluene		Ethylbenzene/Xylenes	
	Indoor Air	Soil Gas	Indoor Air	Soil Gas
Average	0.59	3.38	0.19	0.70
Std. Dev.	0.23	3.46	0.03	0.57
Minimum	0.16	0.44	0.14	0.20
Maximum	0.98	17.0	0.26	4.0
# Samples	17	60	26	48

Table 5.16

Comparison of benzene/toluene and ethylbenzene/total xylenes ratios in indoor air and soil gas.

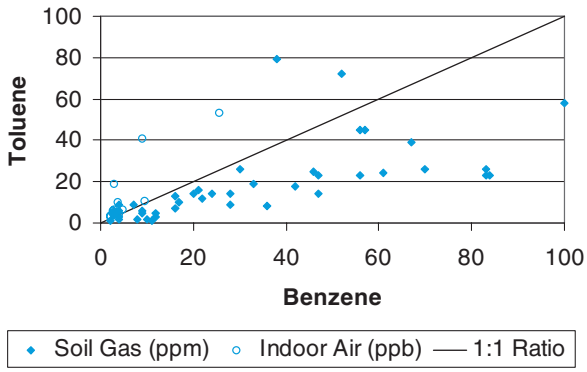


Figure 5.6

Scatter plot of benzene and toluene concentrations in indoor air and soil gas.

Soil gas beneath the site was evaluated as a potential source of BTEX chemicals in indoor air by comparing benzene/toluene (B/T) and ethylbenzene/total xylenes (E/X) ratios. Summary statistics are reported in Table 5.16. Benzene/toluene and ethylbenzene/xylene concentrations are plotted in Figures 5.6 and 5.7. Note that the soil gas data are reported in units of parts per million (ppm) and the indoor air data are reported in units of parts per billion (ppb). The strong positive linear correlation between ethylbenzene and total xylenes in indoor air is evident in Figure 5.7 as the data clearly can be fit to a straight line.

B/T ratios in indoor air ranged from 0.16 to 0.98 with an average of 0.59, whereas soil gas B/T ratios ranged from 0.44 to 17 with an average value of

Figure 5.7

Log–log scatter plot of ethylbenzene and total xylenes concentrations in indoor air and soil gas.

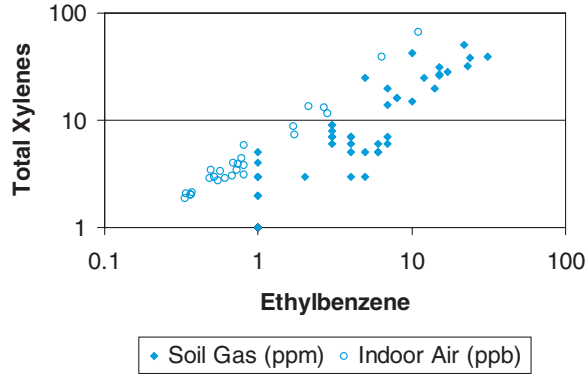
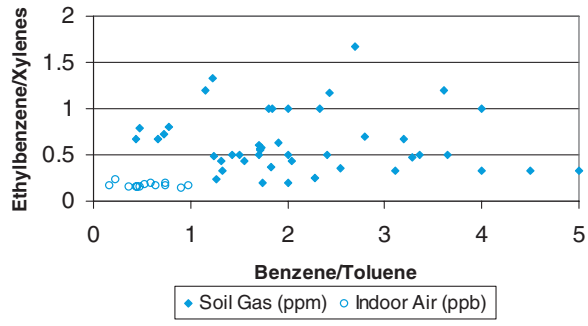


Figure 5.8

Double ratio plot of ethylbenzene/total xylenes vs. benzene/toluene for indoor air and soil gas.



3.38. A Student's *t*-test performed on log-transformed data indicated a significant difference between data sets at the 99% confidence level. A similar conclusion may be drawn by comparing E/T ratios. E/T ratios in indoor air ranged from 0.14 to 0.26 with an average of 0.19 compared to soil gas E/T ratios with a range of 0.2 to 4.0 with an average of 0.7.

Once significant ratios have been identified, double ratio plots can be an effective means of displaying the results. A double ratio plot of B/T vs E/X is shown in Figure 5.8. The indoor air data are neatly clustered in the lower left corner of the plot, and there appears to be no relationship between BTEX concentrations detected in indoor air and soil gas. In contrast, characteristic ratios reported in the literature for gasoline related sources are remarkably similar (see Table 5.17).

5.3.2.3 Spearman Rank Correlation Coefficient

If the data are obviously not normal, then a nonparametric technique may be used. The Spearman Rank Correlation Coefficient, r_s, is a correlation technique that operates on the ranks of the data. Each variable is ranked separately and the difference between ranks for each data pair is recorded. The

Table 5.17

Benzene/toluene and ethylbenzene/total xylenes ratios reported in gasoline related samples, indoor air and soil gas.

Sample Type	Benzene/Toluene Ratio	Ethylbenzene/Xylenes Ratio
Indoor Air (Louisiana Site)	0.59	0.19
Ambient Air Sampled Near Busy Roadway[a]	0.41	0.21
Headspace Sampled above Gasoline[a]	0.50	0.20
Gasoline Engine Exhaust (catalyst)[b]	0.42	0.27
Gasoline Engine Exhaust (no catalyst)[b]	0.62	0.21
Whole Liquid Gasoline[b]	0.19	0.16
Soil Gas (Louisiana Site)	3.38	0.70

[a]Conner *et al.* (1995).
[b]Harley *et al.* (1992).

Spearman rank correlation coefficient is calculated according to the following equation:

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n} \tag{5.53}$$

where d_i is the difference between ranks for each data pair = x_i, y_i. If ties are involved, the equation is more complicated but makes only an appreciable difference if there are a large number of ties.

$$r_s = \frac{\dfrac{(n^3-3)}{6} - \sum_{i=1}^{n} d_i^2 - \sum T_x - \sum T_y}{\sqrt{\left[\dfrac{(n^3-3)}{6} - 2\sum T_x\right]\left[\dfrac{(n^3-3)}{6} - 2\sum T_y\right]}} \tag{5.54}$$

where g is the number of tied groups, t_j is the number of tied data in the jth group, and

$$\sum T_x = \frac{\sum_{j=1}^{g}(t_j^3 - t_j)}{12} \quad \text{(for } x \text{ values)} \tag{5.55}$$

and

$$\sum T_y = \frac{\sum_{j=1}^{g}(t_j^3 - t_j)}{12} \quad \text{(for } y \text{ values)} \tag{5.56}$$

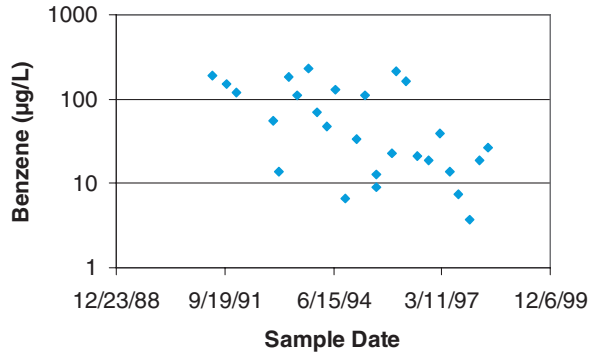


Figure 5.9
Benzene concentration (log scale) vs. time (Zemo, 2000).

Hypothesis testing and setting confidence limits are performed as described for the Pearson product moment correlation coefficient.

In addition to exploratory data analysis, the Spearman Rank Correlation coefficient is a useful statistic for detecting trends in concentration with time and distance (El-Shaarawi *et al.*, 1983). There are no underlying distribution assumptions and the distance and time measurements need not be uniformly spaced.

As an example, benzene concentration data presented by Zemo (2000) are plotted in Figure 5.9. The data in Figure 5.9 suggest that benzene levels are decreasing. The Spearman rank correlation coefficient can be used to test if the decrease is significant. Using Equation 5.53, the calculated rank correlation coefficient is −0.586 and, using equation 5.49, the calculated t statistic is 3.688. The corresponding t value with $\alpha = 0.05$ and 26 degrees of freedom is 2.056. Since the calculated t statistic exceeds the tabular value, we can reject our null hypothesis of $\rho = 0$ and conclude at the 95% probability level that there is a negative linear relationship between benzene concentrations and time. The same type of analysis could also be used to evaluate trends with distance from a source.

5.3.2.4 Multiple Linear Regression and Correlation

The logical extensions to simple linear regression and correlation models are multiple linear regression and correlation techniques. In most cases, we have simultaneous measurements of more than two variables. We can extend the concepts of linear regression and quantify the dependence of a single variable on multiple independent variables. Equation 5.57 describes the general equation:

$$\hat{y}_i = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \ldots + \beta_m x_{m,i} \tag{5.57}$$

It is assumed that y is linearly dependent on x_1 and also linearly dependent on x_2. The β terms are called partial regression coefficients and provide a measure of the relationship between y and a single x variable after removing the effects of all other variables. For example, β_1 is a measure of the relationship between y and x_1 after accounting for effects from all other x variables. The α term is equal to the y intercept when all the β terms are set equal to zero.

Using the basic approach illustrated in Table 5.13, it is possible to evaluate the fit of the regression curve, calculate a coefficient of multiple determination analogous to Equation 5.40, and calculate a multiple correlation coefficient reflecting the overall relationship between all m variables. The equations for performing a multiple regression analysis and calculating the α and β terms are beyond the scope of this chapter (for a discussion of techniques, see Zar, 1984).

One example where linear regression techniques were used to evaluate spatial patterns of soil contamination is presented by Small *et al.* (1995). The authors modeled residential soil lead concentrations near a former automotive battery recycling facility in Pennsylvania to distinguish contributions from the site from other sources such as lead-based paint. A weighted least squares multiple linear regression technique was used. Contributions from the site were represented by relative deposition rates estimated with an atmospheric dispersion model. The authors used indicator variables to represent potential contributions from lead-based paint and other sources. For example, whether or not the sample was collected within five feet of a home built prior to 1955 was selected as an indicator of potential contributions from lead-based paint. Effects from drainage features that might result in concentration enhancements were also included in the model. With this model the authors were able to predict the spatial distribution of lead in soils and help identify additional properties likely to contain lead in soils above the specified cleanup level.

5.3.3 MAPPING: TRENDS AND SPATIAL PERSISTENCE

Many forensics cases involve maps that illustrate the spatial distribution of the environmental variables of interest. These maps may be used to identify and investigate potential contaminant sources and transport patterns, and also to calculate areal averages. Statistical methods that may be used in forensics cases to create maps and characterize spatial variation are discussed in this section.

5.3.3.1 Trends, Spatial Persistence, and Autocorrelation

Many of the environmental variables studied in forensics projects vary gradually over the area of interest. The variation of these variables may be thought

of as a combination of a regional trend and local deviations from the trend. A trend is a systematic (predictable) pattern of variation. Although some data sets do not exhibit significant regional trends, the regional trend component of a spatial data set usually represents a substantial portion of the variation over the area of interest and provides a basis for estimating the value at unsampled locations. Local deviations are the differences between the actual value at each location and the value estimated by the trend. The decomposition of the total variation into these two components is sometimes subjective, but both of these components of the total variation should be considered in developing maps that involve the estimation of values at unsampled locations. Even in cases where the trend is of more interest than the local deviation, both components should be evaluated if the project involves estimation of areal averages or values at specific locations.

The local deviations from a regional trend often exhibit a substantial degree of spatial persistence. A spatially persistent variable is one for which the value at any location is closely related to the values at nearby locations. Although spatial persistence may complicate the derivation of accurate confidence intervals and hypothesis tests (as discussed in Section 5.2.3.3), the spatial persistence of the local deviations from a large-scale trend can be used to derive more accurate isopleth maps. When many observations are available, the degree of spatial persistence can be characterized by calculating the degree of autocorrelation over the distances between the sampling locations. If the variable of interest is autocorrelated over distances that are greater than the distance between adjacent sampling locations, knowledge of the value at one sampling location provides information regarding the values at other nearby sampling locations. The correlation structure of a data set may be characterized by constructing an autocorrelogram or a semivariogram. Both of these graphs illustrate the strength of the association of values as a function of separation distance. In general, the strength of this association declines gradually as separation distance increases. Both methods are explained in detail in Davis (1986). The premise underlying the construction of these graphs is that the difference in value between two locations separated by a given distance is independent of their location in the field. This premise is not consistent with the presence of trends, so spatial persistence should not be evaluated without investigating the possibility of significant spatial trends.

Regional trends are commonly identified by developing a series of numerical models that estimate the observed values as a function of location, then selecting one model to represent the trend. Goodness-of-fit statistics often are used to select the best trend model, but other characteristics of the variable

of interest should also be considered. For concentration data, the nature and scale of the variation may be associated with the mechanisms by which the substances of interest are released and transported. For example, the concentrations in surface soils of substances released to the atmosphere as a plume from a single point (e.g., a stack) may vary gradually over a very large area in a pattern aligned with the prevailing winds as discussed in Chapter 12. Alternately, the concentrations of substances that are released to soil and tend to sorb onto soil particles may be determined primarily by drainage patterns and the degree of soil erosion that has occurred since the release; the area affected by such releases may be relatively small and sharply defined. The following section describes a statistical method of identifying regional trends (trend surface analysis), but nonstatistical factors should also be considered in selecting the most appropriate trend for a specific project and data set.

5.3.3.2 Isopleth Mapping and Interpolation Methods

Isopleths (i.e., lines that connect points of equal value, often referred to as contour lines) are added to maps to illustrate spatial trends and provide estimates of values at unsampled locations. A variety of algorithms is available for developing isopleths from environmental data sets. These methods vary in the degree to which they honor the available data and make use of all relevant information to estimate unknown values.

Trend surface analysis is a method that is used most commonly to characterize data sets that exhibit strong regional trends. This method uses the regression techniques described earlier (see Section 5.3.2) to calibrate equations (models) that predict the value of the variable of interest at each point as a function of location coordinates. First, second, and third order trend surfaces are commonly considered; the general equations for these models are provided in Table 5.18. Higher order models may be developed if sufficient data are available to allow reliable calibration. The significance of various trend surface models can be tested to identify the surface that provides

Table 5.18
General equations for trend surface models.

Order of Trend	Trend Surface Equation
First	$Z = b_0 + b_1X + b_2Y$
Second	$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2$
Third	$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + b_6X^3 + b_7X^2Y + b_8XY^2 + b_9Y^3$

X and Y are the Cartesian location coordinates of the observed values and Z is the property being evaluated for spatial trends.

the best representation of the trend without unnecessary model complexity. The selected surface can then be used to construct isopleths that illustrate the regional trend. The method is explained in more detail by Davis (1986). The surfaces developed by this method are not constrained to match the values observed at the sampling locations used in calibrating the model, but the proportion of the total variation that is explained by these models may be very high. Thus, even though the model does not predict the observed values at the sampled locations exactly, the estimates represented by the isopleths may be quite accurate. In cases where the regional trends are of more interest than the values at individual points, the variation that is not explained by the model (i.e., the local deviation) frequently is ignored. Trend surface analysis is best suited to data sets that are dominated by large-scale regional trends, such as concentration patterns resulting from aerial dispersion and deposition.

The mapping methods that are most appropriate for data sets that exhibit no trend or a high degree of local variation and spatial persistence generally are based on weighted moving averages. In preparation for drawing isopleths, the available data are used to estimate the values at the nodes of a rectangular grid. The value at each grid node is estimated as a function (usually a weighted average) of the values observed at nearby sampling locations. In some of these interpolation methods, the weights are assigned as a function of distance to the node for which the estimate is being developed. Although a variety of moving average methods is available for selecting and weighting the nearby observations, only a few of these algorithms assign weights by statistical analysis of the available data.

The mapping method that makes the greatest use of statistics is kriging. In this procedure, the value at each node is calculated as a weighted average of all the observations that are located within the radius of influence. The radius of influence is the distance over which the variable of interest exhibits significant spatial persistence, which is determined by semivariogram analysis. The weights assigned to the observations are based on the expected difference in value at the distance between each observation and the node. The method is described in greater detail by Davis (1986). Because the expected difference in value for each point in the moving average is known, the expected error of the estimate derived for each node can be calculated. Thus, kriging produces estimated values that match the available observations and provides an index of the accuracy of the interpolated estimates at unsampled locations.

Examples of isopleth maps generated by trend surface analysis and moving average methods are provided in Figures 5.10 through 5.13.[8] Each of these figures is an isopleth map derived from concentration data for metals measured at 180 locations. The first three figures were generated by trend surface

[8] All of these figures were generated using Surfer® 8 software (Golden Software, Inc., Denver, CO), which also provides other options and mapping methods.

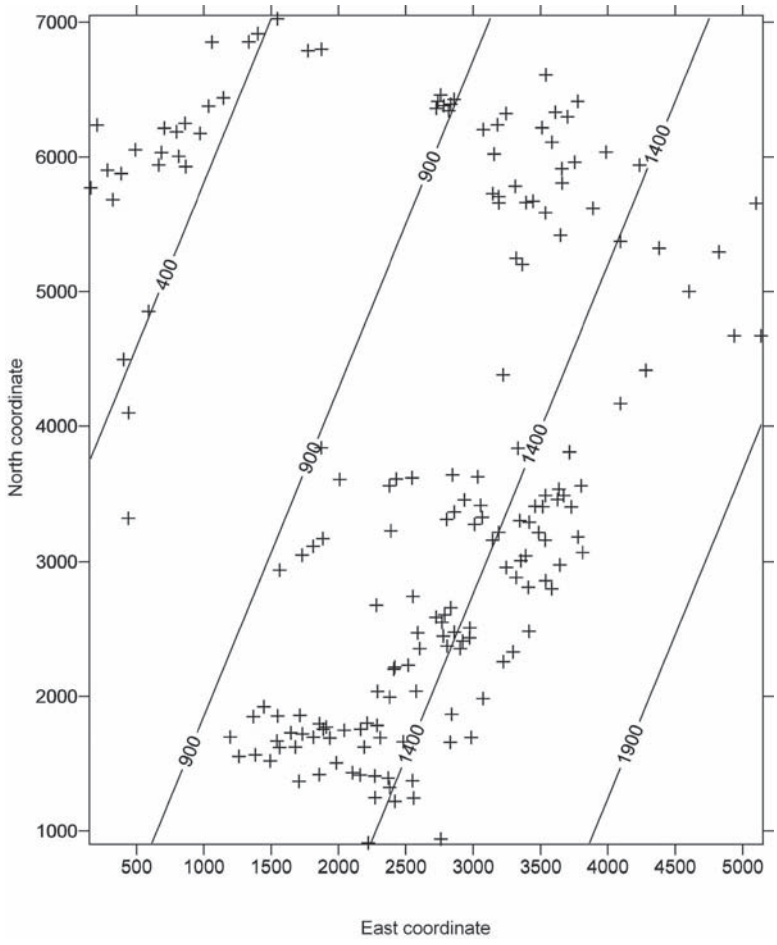


Figure 5.10
Isopleth map for first order trend surface.

analysis. The first order surface shown in Figure 5.10 represents the variation as a plane sloping downward to the west-northwest. The second order trend surface shown in Figure 5.11 indicates a region of relatively high values near the center of the field with a gradual decline in value in all directions. The third order trend surface in Figure 5.12 reveals the existence of a hot spot with a steep decline in most directions, but not to the north-northeast. Such a pattern could be explained by a point source (e.g., a stack) with prevailing winds to the north-northeast. The last figure in this series (see Figure 5.13), which was derived by kriging, suggests that there may be two hot spots rather than a single point source with an elongated footprint in the downwind direction.

These figures are provided to illustrate the characteristics of isopleth maps developed by two types of statistical mapping methods; trend surface analysis, which is based on regression, and kriging, which is a weighted moving

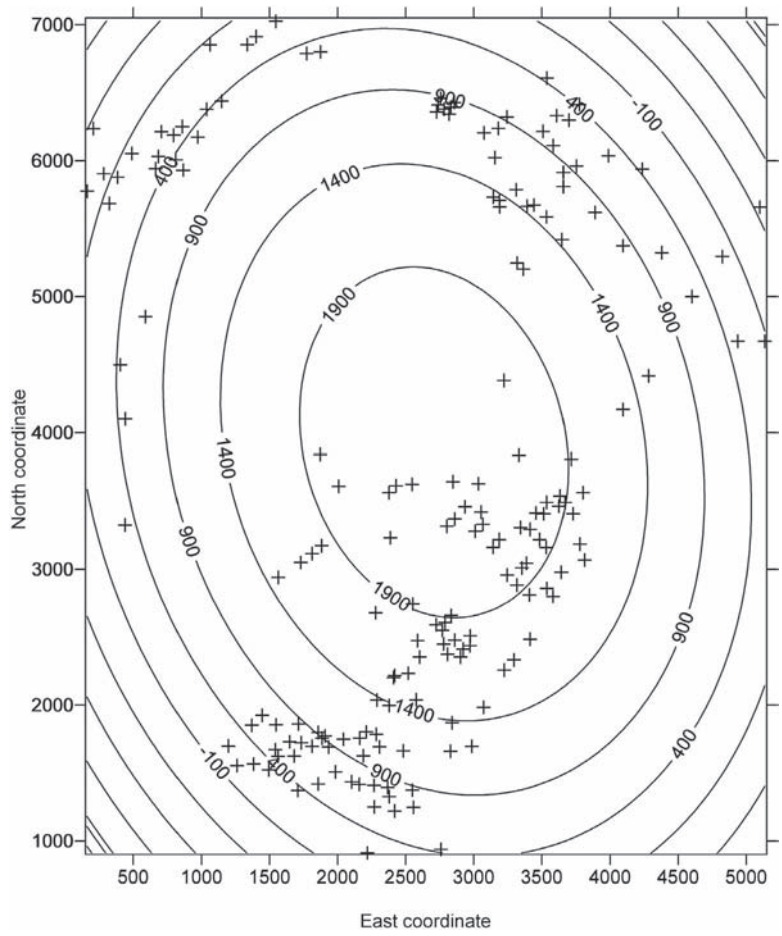


Figure 5.11
Isopleth map for second order trend surface.

average approach. In trend surface analysis, all the available data are used to calibrate a model that estimates the value at each point as a function of its location coordinates. All the observed values are treated as independent observations of the regional pattern, and each observation is used in estimating the value at each point in the field of interest. This approach is appropriate when the objective of the investigation is to identify a regional trend and there is little interest in the local deviation from this trend. Although trend surface models usually do not match the observed values precisely, they can be derived using relatively small data sets. By way of contrast, kriging focuses on the local variation; the values at unsampled locations are estimated using only the values measured at nearby points and the interdependence of values at nearby points is used to improve the accuracy of the estimates. When local patterns and values at specific points are of interest, kriging is the appropriate method. Maps produced by kriging honor the

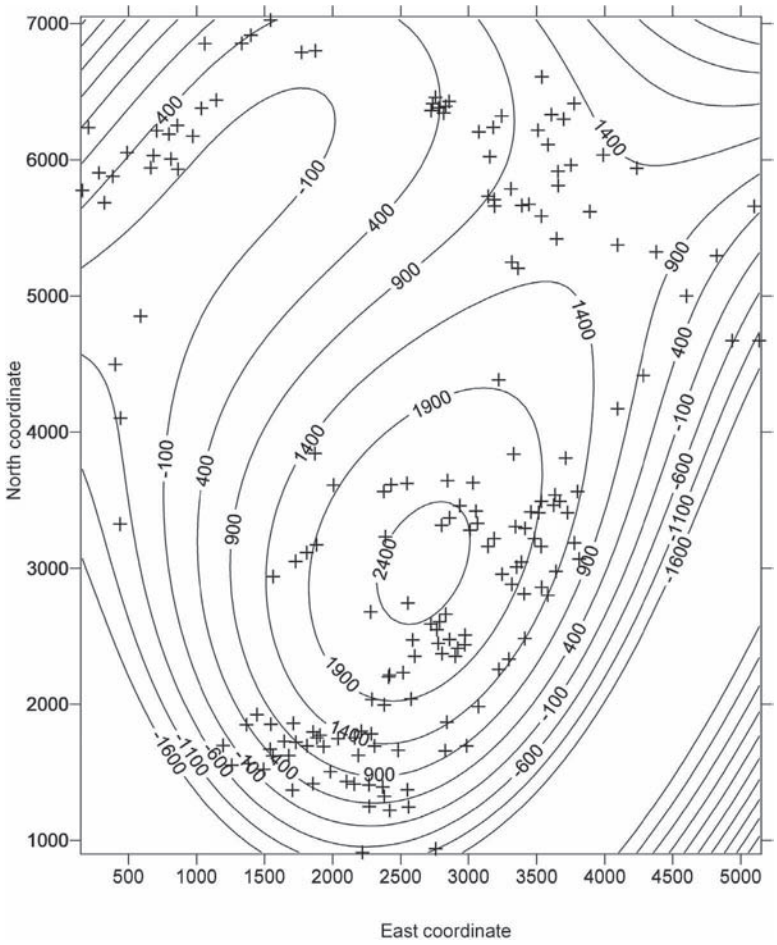


Figure 5.12
Isopleth map for third order trend surface.

observed values, but relatively large data sets are usually needed to develop reliable semivariograms.

5.3.3.3 Estimation of Areal Averages

Environmental forensics projects frequently involve estimation of average values for areas that are defined by political or property boundaries. The simplest way to derive these estimates is to average all the observations collected within the area of interest. This is consistent with the assumptions that underlie most classical statistical methods; that is, that the observed values are independent observations that are all equally representative of the population of interest. In cases where the data set is characterized by significant trends or spatial persistence, however, weighted averaging methods should be considered.

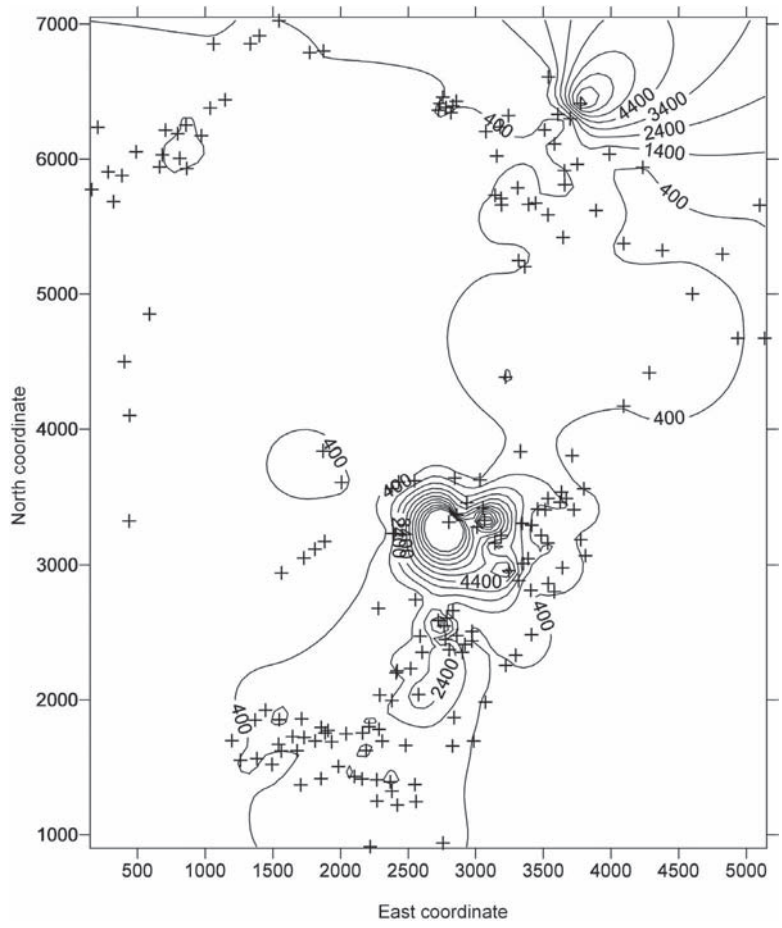


Figure 5.13
Isopleth map developed by kriging.

Block kriging is a statistical method of computing areal averages that can be used with data sets that exhibit both regional trends and spatial persistence. This method generally provides average values for rectangular areas and is appropriate for use with large data sets. When the data set is small, and especially when regional trends account for nearly all the total variation, weighted averages calculated from isopleth maps are appropriate. The area is divided into subareas by the isopleths; each subarea is represented by the average of the isopleths that define it; and the weights are based on the proportion of the total area in each subarea. In cases where significant trends are not evident, the method of Thiessen polygons may be more appropriate. This method assumes that the value at each unsampled location is equal to the value at the nearest sampled location. The area of interest is divided into subareas by the perpendicular bisectors of the lines that connect the various sampling locations, and the proportion of the total area represented by each

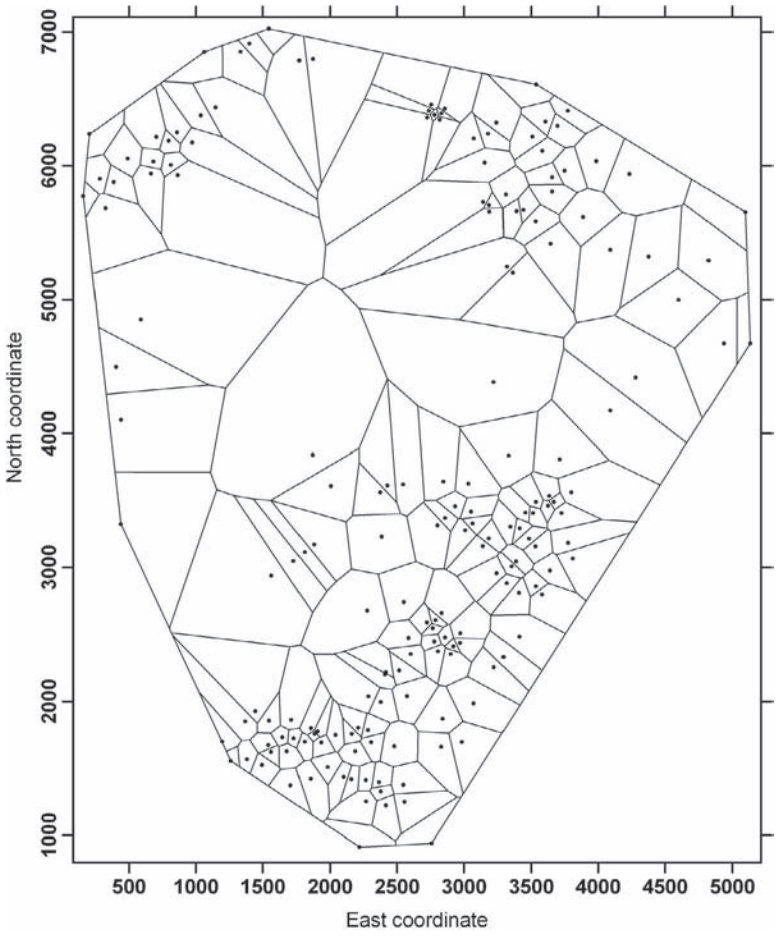


Figure 5.14
Thiessen polygon plot.

observation is used to assign weights in the averaging process. An example of a field of values divided into Thiessen polygons is shown in Figure 5.14.

5.4 CONCLUSIONS

Relatively simple statistical analysis techniques can be used in environmental forensic investigations to compare data sets, characterize associations between variables, evaluate trends, and make predictions. Moreover, it is often possible to assign a degree of confidence to the results. This advantage is particularly useful in litigation scenarios where experts often are asked to assign a probability to the correctness of their opinions.

Of course, the results are not always unambiguous. After all it was Benjamin Disraeli that said "There are three kinds of lies. Lies, damn lies, and statistics." The outcome of a test depends upon which data are included. For example,

the decrease in benzene concentrations in Figure 5.9 is not significant if we look only at data between 1991 and 1996. Thus, the results of any statistical analysis should be interpreted in relation to which data are included (and which data are *not* included), how the test was performed, what assumptions were made, and the validity of those assumptions.

REFERENCES

Amter, S. and Eckel, W.P. (1996) Comment on "Off-site forensic determination of airborne elemental emissions by multi-media analysis: A case study at two secondary lead smelters." *Environmental Science & Technology* 30(7), 2417–2418.

Berthouex, P. and Brown, L. (1994) *Statistics for Environmental Engineers.* CRC Press, Boca Raton, FL.

Boehm, P.D., Page, D.S., Gilfillan, E.S., Bence, A.E., Burns, W.A., and Mankiewicz, P.J. (1998) Study of the fates and effects of the *Exxon Valdez* oil spill on benthic sediments in two bays in Prince William Sound, Alaska. 1. Study design, chemistry, and source fingerprinting. *Environmental Science & Technology* 32, 567–576.

Charrow, R.P. and Bernstein, D.E. (1994) *Scientific Evidence in the Courtroom: Admissibility and Statistical Significance after Daubert.* Washington Legal Foundation, Washington, DC.

Conner, T.L., Lonneman, W.A., and Seila, R.L. (1995) Transportation-related volatile hydrocarbon source profiles measured in Atlanta. *Journal of the Air & Waste Management Association* 45, 383–394.

Davis, J.C. (1986) *Statistics and Data Analysis in Geology*, 2e. John Wiley & Sons, New York.

El-Shaarawi, A.H., Esterby, S.R., and Kuntz, K.W. (1983) A statistical evaluation of trends in the water quality of the Niagara River. *Journal of Great Lakes Research* 9(2), 234–240.

Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring.* Van Nostrand Reinhold, New York.

Gilbert, R.O. and Simpson, J.C. (1990) Statistical sampling and analysis issues and needs for testing attainment of background-based cleanup standards. Presented at the Workshop on Superfund Hazardous Waste, Arlington, VA. February 21–22.

Harley, R.A., Hannigan, M.P., and Cass, G.R. (1992) Respeciation of organic emissions and the detection of excess unburned gasoline in the atmosphere. *Environmental Science & Technology* 26, 2395–2408.

Higgs, F.J., Mielke, H.W., and Brisco, M. (1999) Soil lead at elementary public schools: Comparison between school properties and residential neighborhoods of New Orleans. *Environmental Geochemistry and Health*, 21, 27–36.

Kachigan, S.K. (1991) *Multivariate Statistical Analysis: A Conceptual Introduction.* Radius Press, New York.

Kaplan, I.R., Galperin, Y., Lu, S., and Lee, R. (1997) Forensic environmental geochemistry: Differentiation of fuel types, their source and release time. *Organic Geochemistry* 27(5/6), 289–317.

Kimbrough, D.E. and Suffet, I.H. (1995) Off-site forensic determination of airborne elemental emissions by multi-media analysis: A case study at two secondary lead smelters. *Environmental Science & Technology* 29, 2217–2221.

Land, C.E. (1975) Tables of confidence limits for linear functions of the normal mean and variance. *Selected Tables in Mathematical Statistics* III, 385–419.

Langley, R. (1970) *Practical Statistics*. Dover Publications, Inc., New York.

Liggett, W. (1984) Detecting elevated contamination by comparisons with background. *Environmental Sampling for Hazardous Wastes.* American Chemical Society, Washington, 119–128.

Oakes, M. (1986) *Statistical Inference*. Epidemiology Resources Inc. Chestnut Hill.

Ott, W.R. (1990) A physical explanation of the log-normality of pollutant concentrations. *Journal of the Air & Waste Management Association* 40, 1378–1383.

Parkin, T.B., Meisinger, J.B., Chester, S.T., Starr, J.L., and Robinson, J.A. (1988) Evaluation of statistical estimation methods for log-normally distributed variables. *Soil Science Society of America Journal* 52, 323–329.

Peers, I. (1996) *Statistical Analysis for Education & Psychology Researchers.* The Falmer Press, London.

Porter, P.S., Rao, S.T., Ku, J-Y, Poirot, R.L., and Dakins, M. (1997) Small sample properties of nonparametric bootstrap *t* confidence intervals. *Journal of the Air & Waste Management Association* 47, 1197–1203.

Reckhow, K.H., Clements, J.T., and Dodd, R.C. (1990) Statistical evaluation of mechanistic water-quality models. *Journal of Environmental Engineering* 116(2), 250–268.

Ross, D.L. (1997) Multivariate statistical analysis of environmental monitoring data. *Ground Water* 35(6), 1050–1057.

Schulz, T.W. and Griffin, S. (1999) Estimating risk assessment exposure point concentrations when data are not normal or log-normal. *Risk Analysis* 19(4), 577–584.

Singh, A.K., Singh, A., and Engelhardt, M. (1997) The log-normal distribution in environmental applications. U.S. EPA Office of Research and Development, EPA/600/R-97/006, December.

Small, M.J., Nunn, A.B., Forslund, B.L., and Daily, D.A. (1995) Source attribution of elevated residential soil lead near a battery recycling facility. *Environmental Science & Technology* 29, 883–895.

US EPA (1989) Statistical analysis of ground-water monitoring data at RCRA facilities—Interim final guidance. Office of Solid Waste, Washington, DC, February.

US EPA (1992) Supplemental guidance to RAGS: Calculating the concentration term. Office of Solid Waste and Emergency Response, Washington, DC. Publication 9285.7-081, May.

US EPA (2000) Data quality objectives process for hazardous waste site investigations (QA/G-4HW). Office of Environmental Information, Washington, DC. Publication EPA/600/R-00/007, January.

US EPA (2002a) Calculating upper confidence limits for exposure point concentrations at hazardous waste sites. Office of Solid Waste and Emergency Response, Washington, DC. Publication 9285.6-10, December.

US EPA (2002b) Guidance on choosing a sampling design for environmental data collection. Office of Environmental Information, Washington, DC. Publication EPA/240/R-02/005, December.

Wallace, J. (1997) Calculating removal efficiencies: Confidence limits and hypothesis tests. *Journal of the Air & Waste Management Association* 47, 976–982.

Zar, J.H. (1984) *Biostatistical Analysis*, 2e. Prentice Hall, Englewood Cliffs, NJ.

Zemo, D. (2000) Environmental litigation: Advanced forensics and legal strategies. University of Wisconsin Short Course. April 14.

FURTHER READING

Beers, Y. (1953) *Introduction to the Theory of Error.* Addison–Wesley Publishing Company Inc., Cambridge, MA.

Box, G.E., Hunter, W.G., and Hunter, J.S. (1978) *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* John Wiley & Sons, Inc. New York.

Dunn, O.J. and Clark, V.A. (1974) *Applied Statistics: Analysis of Variance and Regression.* John Wiley & Sons, Inc. New York.

Mandel, J. (1964) *The Statistical Analysis of Experimental Data.* Dover Publications, Inc. New York.

Ott, W.R. (1995) *Environmental Statistics and Data Analysis.* CRC Press, Boca Raton.

Thorndike, R.M. (1978) *Correlational Procedures for Research.* Gardner Press, Inc., New York.

CHAPTER 6

STATISTICAL TOOLS FOR RATIO DATA

Michael E. Ginevan

6.1 Introduction 185
6.2 Chemical Indicator Concentrations and Log-Log Regression Models 186
6.2.1 Methodology 187
6.2.2 Examples 188
6.2.3 Discussion 190
6.2.4 Log-Log Plots: Conclusions 192
6.3 Log-Ratio-Log Plots 193
6.3.1 Log-Ratio-Log Plots and Interesting Patterns 196
6.3.2 Hypothesis Log Testing 196
6.3.3 Log-Ratio-Log Plots: Conclusions 198
6.4 Evaluating Chemical Profiles Using Log(ratio) (LR) Data 198
6.4.1 The Problem 199
6.4.2 A Two-Sample Solution 199
6.4.3 Computing Probabilities for Acceptable Intervals 202
6.4.4 Two-Sample Log-Ratio Difference Tests: A Bootstrap Alternative 203
6.4.5 Many Samples 204
6.4.6 Log(ratio) Data: Conclusions 205
6.5 Final Conclusions 205
References 205

6.1 INTRODUCTION

The importance of ratio data in environmental forensic investigations is well known. For example, Johnson *et al.* (see Chapter 7) note that "Inference of source by pattern recognition techniques is concerned more with the relative proportions between analytes than with absolute concentrations." This focus is the result of the idea that while concentrations may be modified by factors like dilution, relative concentrations are likely to be more stable. It is also well known that chemical measurements from environmental samples tend to follow a log-normal distribution. The major reason for this log-normality is

that the random factors operating on chemical concentration data tend to be multiplicative (Ott, 1994). For example, a given set of chemicals in a waste stream tend to be diluted as one samples further and further from the source. Similarly, if one samples in areas where there is periodic flooding, chemicals dissolved on flood waters may tend to be concentrated by evaporation. More generally, if variation of concentration of a chemical is the product of a large number of random multiplicative factors, one would expect that, because of the central limit theorem, the concentration of the chemical will follow a log-normal distribution; that is, the logarithm of the chemical concentration can be expressed as the sum of the logarithms of the random multiplicative factors. Thus the central limit theorem (Hogg and Craig, 1995) says that we would expect the logarithm to follow a normal distribution. Note that in our discussions we are using log base 10 rather than log base e. All the comments presented here are independent of the logarithmic base used.

The importance of ratio data in environmental forensics coupled with the tendency of measurements of chemical concentration data to follow log-normal distributions provide the basis for a number of useful tools, which are discussed in subsequent sections.

6.2 CHEMICAL INDICATOR CONCENTRATIONS AND LOG-LOG REGRESSION MODELS

When assessing the origin of a mix of hazardous chemicals, it is helpful to be able to determine if measurements of one chemical tend to be a constant proportion of measurements of another. First, if their ratio is constant across a number of measurements, this may be taken as evidence that the samples have a common origin. Second, one might also want to use one chemical as a surrogate or indicator for another chemical in the mixture because the first is at higher concentration and can be used to predict levels of the second for samples in which its concentration is below the limit of detection. That is, if we can show that the concentration of chemical A is some constant fraction, F, of chemical B, we can measure the concentration of B, C_B, and infer the concentration of A, C_A, as:

$$C_A = F \bullet C_B \tag{6.1}$$

The question addressed here is how to use the actual measurements of chemicals A and B to determine whether a relationship such as that shown in Equation 6.1, in fact, exists.

6.2.1 METHODOLOGY

If the concentration of a chemical follows a log-normal distribution, the log of the concentration will follow a normal distribution. For two chemicals, we might expect a bivariate log-normal distribution, which would translate to a bivariate normal distribution for the log-transformed concentrations. If we translate Equation 6.1 to logarithmic units we obtain:

$$\mathrm{Log}(C_A) = \mathrm{Log}(F) + \mathrm{Log}(C_B) \tag{6.2}$$

This is similar to the regression equation of the logarithm of C_A on the logarithm of C_B. That is, when $\mathrm{Log}(C_A)$ is the dependent variable and $\mathrm{Log}(C_B)$ is the independent variable, the regression equation is:

$$\mathrm{Log}(C_A) = \alpha + \beta \bullet \mathrm{Log}(C_B) \tag{6.3}$$

where α is the intercept term and β is the slope of the regression model. If we let $\mathrm{Log}(F) = \alpha$ (i.e., $F = 10^{\alpha}$), and back-transform Equation 6.3 to original units by taking exponentials (e.g., 10^X, where X is any regression term of interest), we obtain:

$$C_A = F \bullet C_B^{\beta} \tag{6.4}$$

Equation 6.4 is the same as 6.1 except for the β exponential term on C_B, and 6.4 would be identical to 6.1 for the case $\beta = 1$.

This brings us to a general methodology for assessing the utility of using one chemical (chemical B in this case) as an indicator for another chemical (chemical A in this case). One can simply regress the log-transformed concentrations of one chemical on the log-transformed concentration of the other chemical (assuming that the pairs of concentrations are from the same physical sample). One can then use the results of this calculation to evaluate the utility of chemical B as an indicator for chemical A by statistically testing whether $\beta = 1$.

For a formal test of whether Equation 6.1 actually describes the relationship between chemical A and chemical B, one proceeds as follows:

1. Find the regression coefficient (β) for Log (chemical B) regressed on Log (chemical A) together with the standard error of this coefficient (SE_β). (See the examples in the tables.)

2. Construct a formal hypothesis test of whether β equals one as follows:

$$Z = (1 - \beta)/SE_{\beta} \tag{6.5}$$

3. Compare Z to the standard normal distribution.

For significance (i.e., rejecting the hypothesis that $\beta = 1$) at the $p = 0.05$ level on a two-sided test (null hypothesis H_0: $\beta = 1$ vs. the alternate hypothesis H_1: $\beta \neq 1$), the absolute value of Z must be greater than 1.96. In the event that we fail to reject H_0 (i.e., we accept that $\beta = 1$), it follows that Equation 6.1 is a reasonable description of the regression of A on B and that chemical B may thus be a reasonable linear indicator for chemical A.

Table 6.1

Regression calculations for evaluating the utility of ultraviolet fluorescing particulate matter (UVPM) as an indicator for nicotine.[a]

Predictor Variables	Coefficient	Standard Error	Student's *t*	*P*-Value[b]
Constant (α)	–1.044	0.034	–30.8	0.00
Log (UVPM) (β)	0.935	0.034	27.9	0.00

[a]Data are from a study of exposure to environmental tobacco smoke and are the base 10 logarithms of concentrations (in $\mu g/m^3$) observed in workplaces where smoking occurred.
[b]This is the *p*-value for the test of H_0: $\alpha = 0$ or H_0: $\beta = 0$ (i.e., this is not the *p*-value for the test of H_0: $\beta = 1$).
R-squared = 0.63. Cases included: 451.

6.2.2 EXAMPLES

The examples in the tables are taken from a study of exposure to environmental tobacco smoke in workplaces where smoking occurred (LaKind *et al.*, 1999a,b,c) and are chosen to illustrate the use of the Z statistic in conjunction with other statistical information. The first example (see Table 6.1) considers the log-log regression of the nicotine concentration in air (in $\mu g/m^3$) on the ultraviolet fluorescing particulate matter concentration in air (UVPM; also in $\mu g/m^3$). Here we see that the Z statistic described in Equation 6.5 is only 1.91 ($p = 0.06$). Thus, we cannot formally reject H_0, and might wish to consider UVPM as an indicator for nicotine. This might be desirable because nicotine is somewhat harder to measure than UVPM. However, in this case, the R^2 of the regression model given in Table 6.1 is only 0.63. That is, regression of Log (nicotine) on Log (UVPM) explains only 63% of the variation in the log-transformed nicotine concentration. The general regression equation suggests that, on average, nicotine is a constant proportion of UVPM. This proportion is given by $F = 10^{\alpha} = 10^{-1.044} = 0.090$. However, the lack of a relatively

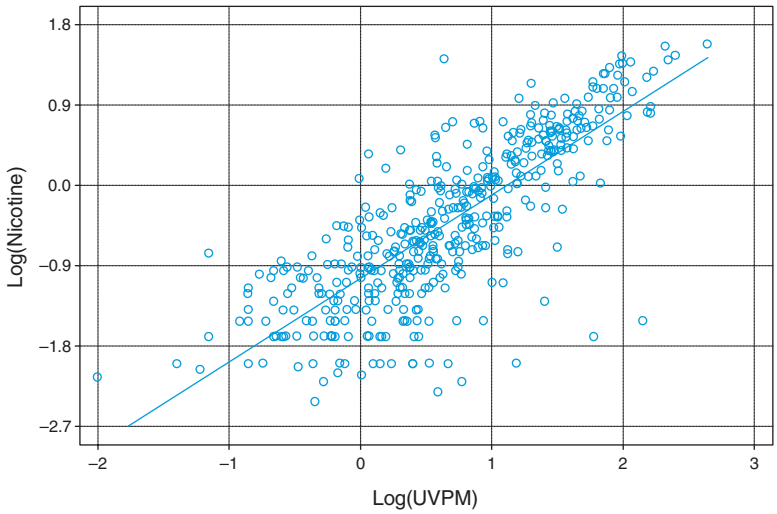


Figure 6.1
Regression of log(nicotine) on log(UVPM).

Table 6.2
Regression calculations for evaluating the utility of ultraviolet fluorescing particulate matter (UVPM) as an indicator for visible light fluorescing particulate matter (FPM).[a]

Predictor Variables	Coefficient	Standard Error	Student's *t*	*P*-Value[b]
Constant (α)	−0.169	0.010	−16.5	0.00
Log (UVPM) (β)	1.059	0.010	103.4	0.00

[a]Data are from a study of exposure to environmental tobacco smoke and are the base 10 logarithms of concentrations (in $\mu g/m^3$) observed in workplaces where smoking occurred.
[b]This is the *p*-value for the test of H_0: $\alpha = 0$ or H_0: $\beta = 0$ (i.e., this is not the *p*-value for the test of H_0: $\beta = 1$).
R-squared = 0.96. Cases included: 459.

high R^2 suggests that for individual observations, the UVPM concentration may or may not be a reliable predictor of the nicotine concentration in air. That is, on average the bias is small, but the difference between an individual nicotine level and the prediction from the regression model may be large. The log-log regression of nicotine on UVPM is shown in Figure 6.1.

The log-log regression of the concentration of visible light fluorescing particulate matter (FPM; in $\mu g/m^3$) on the concentration of UVPM presents a different picture (see Table 6.2). Here the Z statistic is 5.90 ($p < 0.0001$), suggesting that there is reason to believe that β does not equal 1 (i.e., that the relationship between FPM and UVPM is not strictly linear). However, the estimated value of β is 1.059, which is numerically if not statistically close to 1, and Figure 6.2 shows that, in fact, the linear model is a very good description of the data. The bulk of the scatter in Figure 6.2 occurs at low concentrations where the reliability of the laboratory analysis is lowest. Overall, the R^2 for the model is 0.96. That is, the regression of Log (FPM) on Log (UVPM) explains 96% of the variation in log-transformed FPM.

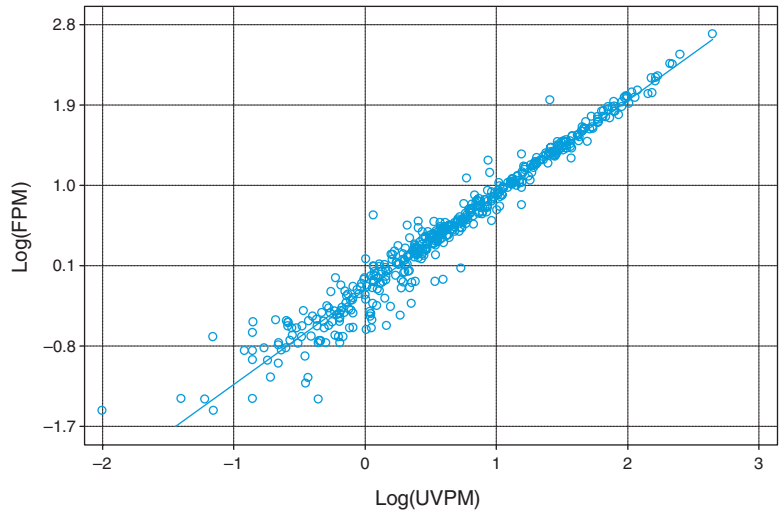


Figure 6.2
Regression of log(FPM) on log(UVPM).

If we calculate the proportion of UVPM that is FPM at unit concentration, we get $F = 10^{-0.169} = 0.678$. If we apply this result to a situation where there were 10 units of UVPM, we would predict 6.78 units of FPM. The exact calculation given by the full regression equation (i.e., using both α and β terms) is from Equation 6.4:

$$\text{FPM} = 0.678 \bullet 10^{1.059} = 7.77 \tag{6.6}$$

That is, we would expect 7.77 units of FPM given 10 units of UVPM. Thus the simple estimator underpredicts FPM by about 13% (1–[6.78/7.77]). The real value of the log-log model suggested here is that it allows one to easily make these kinds of comparisons and thus to generally assess the utility of one chemical as an indicator for another.

6.2.3 DISCUSSION

One caveat is appropriate to the method suggested here. In regression models, it is explicitly assumed that the predictor variable (in this case chemical B) is measured without error. Since measured concentrations are in fact estimates based on the outcome of laboratory procedures, this assumption is not met in our method. When the predictor variable is measured with error, the slope estimate (β) is biased toward zero. That is, if the predictor chemical is measured with error, the β value in our model will tend to be less than 1. However, for many situations the degree of this bias is not large, and we may, in fact, be able to correct for it. The general problem, usually referred to as the "errors

in variables problem," is discussed in Rawlings *et al.* (1998) and in greater detail in Fuller (1987).

One useful way to look at the issue is to assume that each predictor X_i can be decomposed into its "true value", Z_i, and an error component, U_i. The U_i's are assumed to have zero mean and variance σ_u^2. One useful result occurs if we assume that (1) the Z_i's are normally distributed with mean 0 and variance σ_z^2, (2) the U_i's are normally distributed with with mean 0 and variance σ_u^2, and (3) the Z_i's and U_i's are independent. Then:

$$\beta_C = \beta_E \bullet (\sigma_z^2 + \sigma_u^2)/\sigma_z^2 \tag{6.7}$$

where β_C is the correct estimate of β, and β_E is the value estimated from the data. It is clear that if σ_z^2 is large compared to σ_u^2 then:

$$(\sigma_z^2 + \sigma_u^2)/\sigma_z^2 \approx 1 \quad \text{and} \quad \beta_C \approx \beta_E \tag{6.8}$$

Moreover, we typically have a fairly good idea of σ_u^2 because this is the logarithmic variance of the error in the analytic technique used to analyze for the chemical being used as the predictor in our regression. Also because we assume Z_i and U_i to be uncorrelated, it follows that:

$$\sigma_x^2 = \sigma_z^2 + \sigma_u^2 \tag{6.9}$$

Thus we can rewrite Equation 6.7 as:

$$\beta_C = \beta_E \bullet \sigma_x^2/(\sigma_x^2 - \sigma_u^2) \tag{6.10}$$

How large might this correction be? Well, for environmental measurements, it is typical that 95% of the measurements are within a factor of 10 of the geometric mean, and for laboratory measurements we would hope that 95% of the measurements would be within 20% of the true value.

For log-normal distributions this would imply that on the environmental side:

$$UB_{env,\,0.975} = GM \bullet 10 \tag{6.11}$$

That is, the 97.5 percentile of the environmental concentration distribution, $UB_{env,\,0.975}$, is given by the geometric mean, GM, times ten. If we rewrite Equation 6.11 in terms of logarithms, we get:

$$\text{Log}\, UB_{env,\,0.975} = \text{Log}\, GM + \text{Log}(10) \tag{6.12}$$

Here, Log GM is the log of the geometric mean, and, of course, in Base 10 Log(10) = 1. It is also true that:

$$\text{Log}\,\text{UB}_{\text{env, 0.975}} = \text{Log}\,\text{GM} + 1.96\sigma_x \tag{6.13}$$

Thus, equating Equations (6.12) and (6.13),

$$\sigma_x = \text{Log}(10)/1.96 = 0.5102 \quad \text{and, thus,} \quad \sigma_x^2 = 0.2603 \tag{6.14}$$

By similar reasoning, for the error distribution attributable to laboratory analysis,

$$\text{UB}_{\text{lab, 0.975}} = \text{GM} \bullet 1.2 \tag{6.15}$$

This results in:

$$\sigma_u = \text{Log}(1.2)/1.96 = 0.0404 \quad \text{and} \quad \sigma_u^2 = 0.0016 \tag{6.16}$$

When we substitute the values from Equations (6.14) and (6.16) into (6.10) we obtain:

$$\beta_C = \beta_E \bullet 1.0062 \tag{6.17}$$

Thus, if 95% of the concentration measurements are within a factor of 10 of the geometric mean and the laboratory measurements are within 20% of the true values, then the bias in β_E is less than 1%.

The first important point that follows from this discussion is that measurement errors usually result in negligible bias. However, if σ_x^2 is small, which would imply that there is little variability in the chemical concentration data, or σ_u^2 is large, which would imply large measurement errors, β_E may be seriously biased toward zero. The points to remember are that if the measurements have little variability or analytic laboratory variation is large, the approach discussed here will not work well. However, for many cases, σ_x^2 is large and σ_u^2 is small, and the bias in β_E is therefore also small.

6.2.4 LOG-LOG PLOTS: CONCLUSIONS

When the variance of measurements is large and the laboratory variance is small, the log-log regression technique yields a great deal of useful information about whether one chemical is a constant concentration ratio of another chemical. The regression equation provides the basis for a hypothesis test of

whether or not chemical A is a constant fraction of chemical B. The R^2 statistic provides a measure of the predictive strength of chemical B as an indicator for A, and the regression equation provides a simple ratio predictor as a basis for assessing bias in the range of likely concentration levels of chemical B. Moreover, the form of the log-log plot may provide insight as to the physical process relating chemicals A and B. That is, if β is less than one, it suggests that as the concentration of B increases, A becomes smaller and smaller in relation to B. Conversely if β is greater than 1, it implies that A becomes large relative to B as concentrations increase. For example, work on environmental tobacco smoke (ETS; LaKind *et al.*, 1999a,b,c) showed that respirable particulate matter (RSP) became small in relation to ultraviolet fluorescing particulate matter (UVPM) as the concentration of UVPM increased. In retrospect, this was an expected outcome; RSP has a variety of sources and is not strongly related to ETS, whereas UVPM is more specific to ETS. Thus RSP concentration becomes small relative to UVPM as the latter increases in concentration.

6.3 LOG-RATIO-LOG PLOTS

Log-log concentration plots can be helpful in evaluating the hypothesis of constant composition, but the plots by themselves are hard to evaluate because, given the form of Equations 6.3 and 6.4, any power relationship will appear linear and if there is a consistent power relationship the association will appear quite strong. Figure 6.3 illustrates this problem—the ratio of C1 to C2 is not constant (panel A) but is linear on a log-log plot (panel B).

Figure 6.4 shows a different visualization of the concentration relationship between pairs of chemicals. This shows the same data as in Figure 6.3, but plots ln(C1/C2) against ln(C2). Here the inconstant ratio in concentration is readily apparent. Moreover this is a plot that is generally useful in visualizing pairwise relationships. That is, if we have a log-log concentration relationship that shows a strong linear association between ln(C1) and ln(C2), the constant composition relationship is distinguished from all power relationships by the fact that it is the only one that will have a slope of zero on a plot of ln(C1/C2) vs. ln(C2). This will be termed a log-ratio-log or LRL plot. It has considerable utility for visualizing concentration relationships.

Looking at Figures 6.3 and 6.4, one might ask why one would plot data on a logarithmic scale. The answer is that Figures 6.3 and 6.4 use hypothetical data contrived to make a point. Real data (see Figures 6.1 and 6.2) tend to follow severely right skewed distributions. Thus, an untransformed bivariate plot would show an indecipherable mass of points near the origin with a scattering of higher valued observations above this mass, which is not an ideal

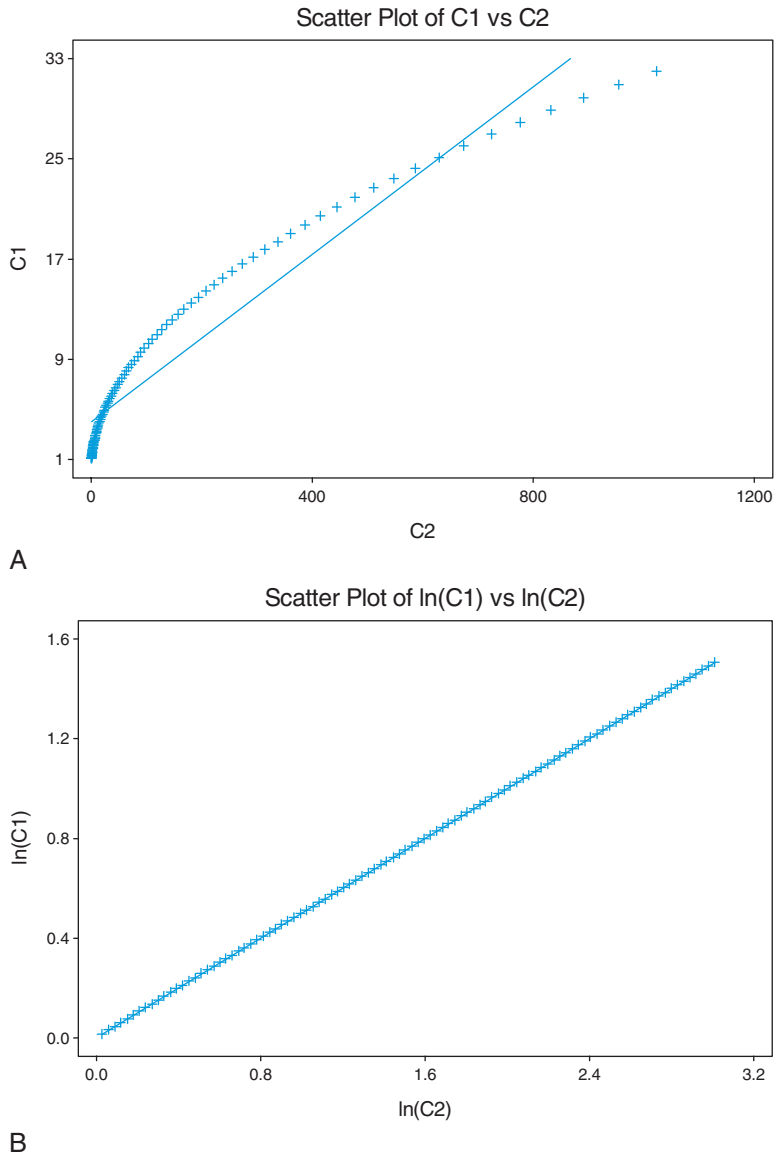


Figure 6.3
The same relationship on ***A****: linear, and* ***B****: logarithmic scales.*

situation for pattern recognition. One might also question whether data that do not have any strong correlation might not show a roughly horizontal scatter plot similar to data that show a fairly constant association. The short answer to this is that there would be no significant pairwise association if C1 and C2 are, in fact, independent. More importantly, if C1 and C2 are independent and have equal logarithmic variances (which is necessarily true for the case of a constant concentration ratio), the plot of ln(C1/C2) versus ln(C2) will show a quite strong negative relationship with a correlation of approximately

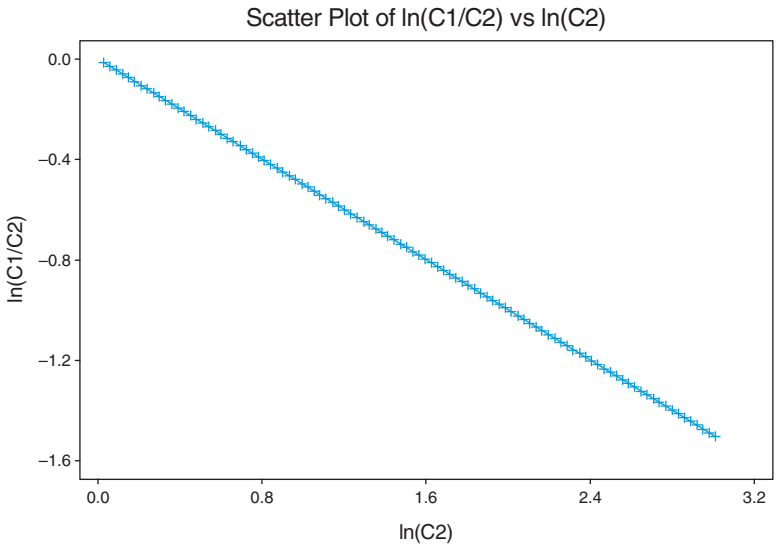


Figure 6.4
A log-ratio-log plot of the same data shown in Figure 6.3.

−0.707. To show this, we can assume, without loss of generality, that ln(C1) and ln(C2) have been standardized to zero mean. Take ln(C1) as Y and ln(C2) as X. The correlation coefficient (r) between ln(C1/C2) and ln(C2) is:

$$\frac{\sum X(Y-X)}{\left[\sum X^2 \sum (X-Y)^2\right]^{1/2}} \tag{6.18}$$

or

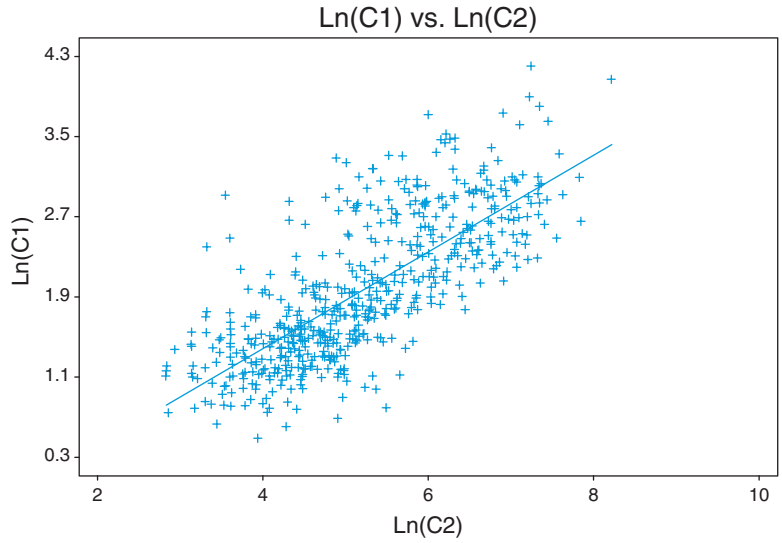
$$\frac{\sum -X^2 - \sum XY}{\left[\sum X^2 \sum (X^2 - 2XY + Y^2)\right]^{1/2}} \tag{6.19}$$

If ln(C1) and ln(C2) uncorrelated, then Equation (6.19) simplifies to:

$$\frac{-\sum X^2}{\left[\sum X^2 \sum (X^2 + Y^2)\right]^{1/2}} \tag{6.20}$$

Thus, if the logarithmic variances of C1 and C2 are equal, keeping in mind that X and Y have zero mean, the expected correlation is −0.707. As noted earlier, equal logarithmic variances are a necessary, but not sufficient, precondition for a constant concentration ratio. A final point is that standardizing log transformed concentration data to unit variance in effect applies a power

Figure 6.5
Real chemical data on a log-log plot—is the relationship consistent?



transformation in the original variable space, and can create an impression of a constant ratio for data that actually have a nonconstant ratio. Thus, such transformations should be avoided.

6.3.1 LOG-RATIO-LOG PLOTS AND INTERESTING PATTERNS

Figure 6.5 shows a log-log plot of two chemicals across many samples. There is obviously a strong relationship but is it consistent? Figure 6.6 shows the same data in an LRL plot. The nonconstant ratio is obvious. One can also use LRL plots as an alternative to or augmentation of simple scatter plot matrices. Figure 6.7 shows pairs of chemicals from the same data set as used for Figures 6.5 and 6.6. It is obvious that the only ratio that is close to constant is C2 and C3. Exploratory plots like Figure 6.7 are a useful screening tool. For example, environmental forensics investigations often use double ratio plots to try and identify groups of samples that have similar composition (Murphy and Morrison, 2001). An LRL scatter plot matrix can be useful in picking the pairs for a double ratio plot. This point is reinforced in Figure 6.8, which shows an LRL plot with two populations with differing but constant ratios. LRL plots make such relationships obvious.

6.3.2 HYPOTHESIS LOG TESTING

The present discussion has focused on data visualization, but one might ask whether one can use the regression lines in LRL plots as a basis for hypothesis tests. That is, if the ratio relationship is not constant one would expect the

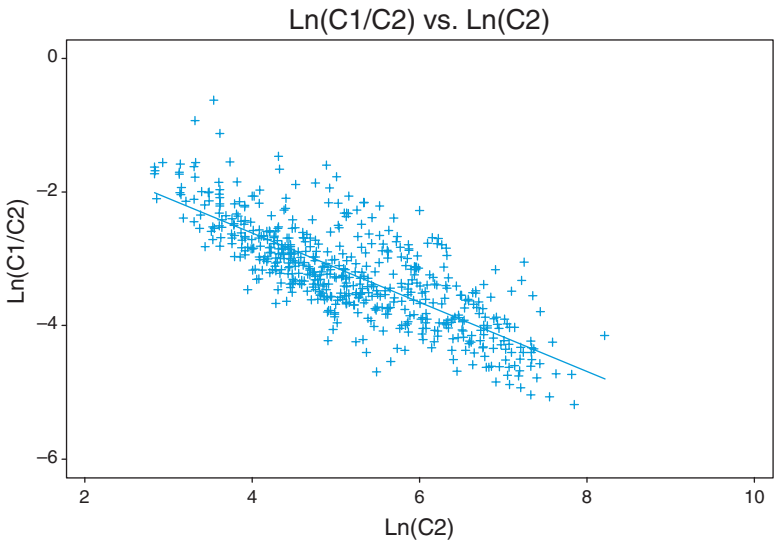
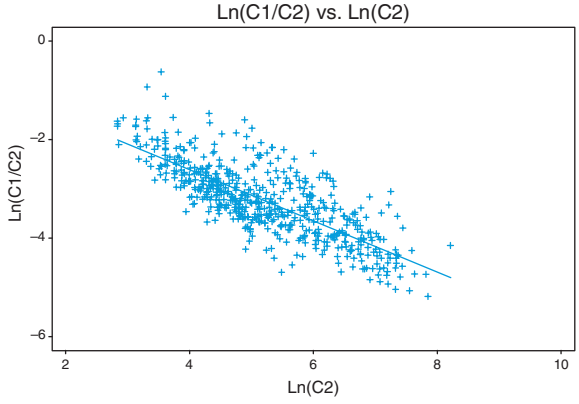
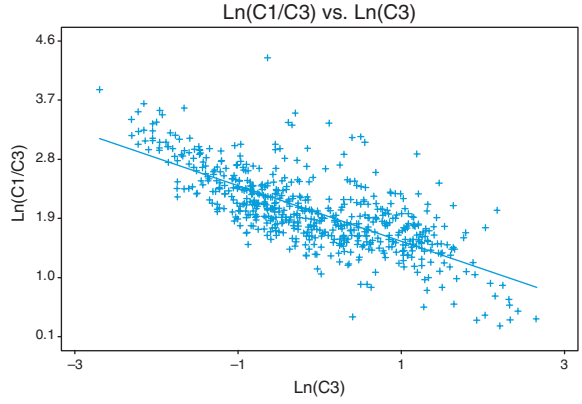


Figure 6.6

The same data as in Figure 6.5 in an LRL plot. The ratio is not consistent.





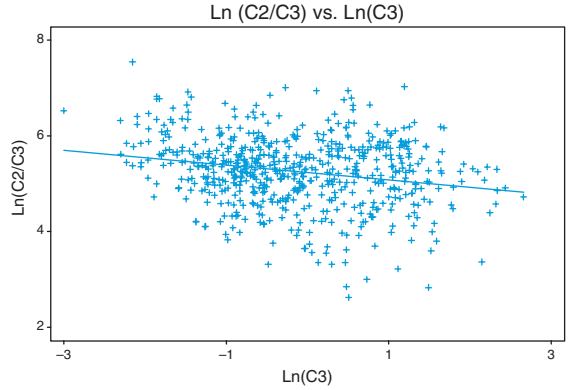


Figure 6.7

LRL scatter plot matrices. C2 and C3 are closest to a constant ratio.

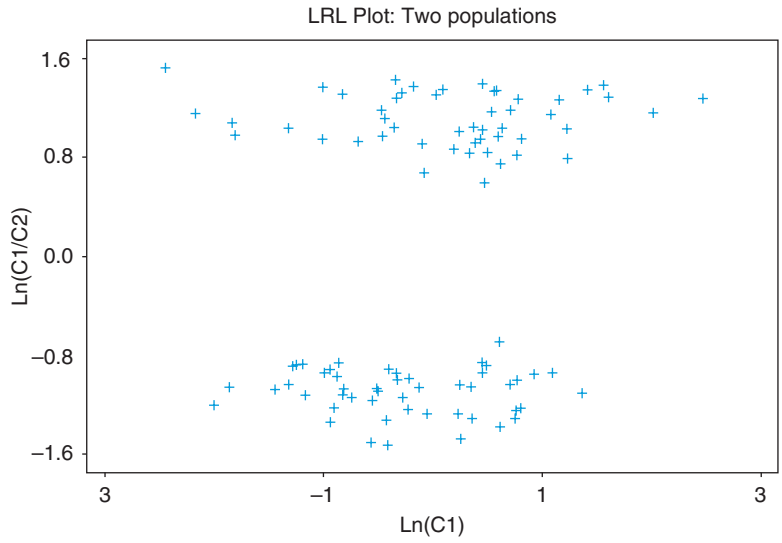


Figure 6.8
An LRL plot of two populations with different C1/C2 ratios.

correlation coefficient between ln(C1/C2) and ln(C2) to differ significantly from zero. This is certainly true, but the previously presented approach of simply regressing ln(C1) vs. ln(C2) results in a more straightforward statistical evaluation. To review, one must answer two questions. First, is the relationship between C1 and C2 strong (high correlation), and second, does the regression coefficient for C1 on C2 differ significantly from one? If there is a strong correlation and the regression coefficient does not differ significantly from one, we have good evidence of a constant ratio (Ginevan and Splitstone, 2004).

6.3.3 LOG-RATIO-LOG PLOTS: CONCLUSIONS

The LRL plots shown here are a simple, but useful, way of looking at chemical concentration data. They can be used to dispute the idea that a good log-log regression fit implies a constant ratio relationship. More important, they can allow one to screen a data set to determining which pairs of chemical concentrations do exhibit constant relationships (see Figure 6.7) and can also help identify multiple groups of observations with constant ratios within each group but with different ratios between groups (see Figure 6.8). In summary, LRL plots provide a unique perspective in the search for patterns in environmental data.

6.4 EVALUATING CHEMICAL PROFILES USING LOG(RATIO) (LR) DATA

The preceding sections considered ways at looking at the question of a constant ratio between two chemicals across a number of environmental samples. This section will examine the somewhat different problem of testing whether

two or more samples have similar composition across a number of chemical species. It has already been established that the logarithm of a concentration ratio of a pair of chemicals across samples (e.g., a sample-by-sample ratio) often will tend to follow a log-normal distribution, because chemical concentration data tend to be log-normally distributed. It is also true that there will be a tendency for the log transformed concentrations different chemicals to have approximately equal variances because as noted at the beginning of this chapter, the random factors that act on chemical concentrations in the environment tend to be multiplicative, and will tend to have the same relative effect on all chemicals involved in the process. Of course, there are exceptions to this as when one chemical reacts with other chemicals to form new chemical species where another more stable compound does not. Nonetheless it is true that log-transformed chemical concentration data tend to have much more similar variances across chemical species than do the same data untransformed.

6.4.1 THE PROBLEM

Assume there are two samples, each of which has quantified measurements (that is, all observations are above the detection limit) for each of N chemical species (dioxin congeners, metals, pcb's, etc.). One can identify K "interesting" pairs of congeners and take the ratio of the first chemical to the second, and then take the logarithm of this ratio. Note that "interesting" is an environmental forensics issue, and will be determined by things like source composition characteristics and relevant concentration ranges from a risk or regulatory standpoint.

The result of this procedure is K LR values for each sample, which one might reasonably expect to follow normal distributions across samples. Also note that, if the ratio is similar across samples, we would expect the variance of the LR values to be small.

6.4.2 A TWO-SAMPLE SOLUTION

Table 6.3 shows such data for actual environmental samples with 11 pairs each. The question to address is whether or not samples 1 and 2 have similar composition, and similarly, whether samples 3 and 4 have similar composition. Figure 6.9 shows the log-ratio data plotted for samples 1 and 2, and 3 and 4, respectively. One could argue that the profiles look similar but it is hard to determine exactly how similar the two pairs of samples really are.

One way of looking at the problem is to note that the log transformation tends to make variances across chemicals in a given set of samples more nearly equal and that log ratios will tend to be normally distributed. This leads

Table 6.3

A sample data set for log ratio difference tests.

Chemicals	Sample Ratios				Sample Log Ratios				Log Ratio Difference	
	S1	S2	S3	S4	S1	S2	S3	S4	S1–S2	S3–S4
C1/C2	1850.000	90.727	232.432	53.175	7.523	4.508	5.449	3.974	3.015	1.475
C3/C4	1.868	24.698	5.665	49.594	0.625	3.207	1.734	3.904	−2.582	−2.169
C5/C6	0.063	0.100	0.062	0.051	−2.763	−2.303	−2.774	−2.984	−0.461	0.209
C7/C8	0.526	0.660	2.131	3.968	−0.643	−0.415	0.757	1.378	−0.228	−0.622
C9/C10	118.239	13.926	34.539	94.359	4.773	2.634	3.542	4.547	2.139	−1.005
C11/C12	2.174	0.962	1.509	1.257	0.777	−0.039	0.412	0.229	0.816	0.183
C13/C14	0.619	3.043	0.467	1.216	−0.480	1.113	−0.762	0.196	−1.593	−0.958
C15/C16	1.425	2.289	0.813	1.803	0.354	0.828	−0.208	0.590	−0.474	−0.797
C17/C18	0.692	2.222	0.933	1.923	−0.368	0.799	−0.069	0.654	−1.167	−0.723
C19/C20	0.910	0.070	0.652	0.769	−0.094	−2.662	−0.427	−0.262	2.567	−0.165
C21/C22	1.215	1.500	1.214	1.263	0.194	0.405	0.194	0.234	−0.211	−0.039

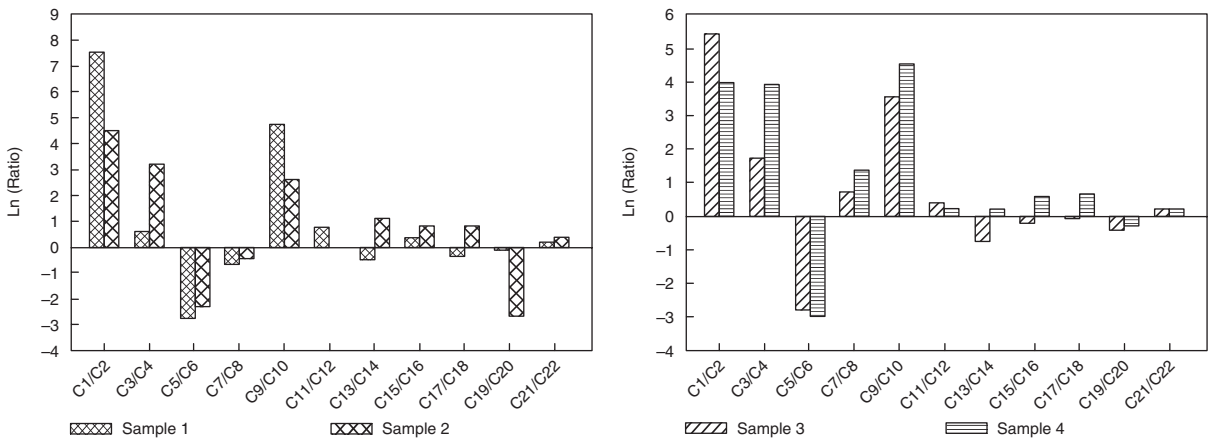


Figure 6.9

Bar chart of log ratios for 11 pairs for Samples 1 and 2, and 3 and 4, respectively.

naturally to the idea of a two-sample *t*-test for paired data (Sokol and Rohlf, 1995). One can calculate the mean and variance of the differences between LR pair values (log-ratio difference; LRDs) from two samples and calculate a confidence interval for the LRDs. The distribution used to calculate the confidence interval is a *t* distribution on K − 1 degrees of freedom. If the mean LRD confidence interval does not include zero (which is necessary to assume equal ratios) the ratios differ between samples.

However, a confidence interval that includes zero does not necessarily mean that the profiles are the same. If some differences are strongly negative but others are strongly positive, the average difference may be close to zero but the variability is very high. In this case the LRD confidence interval may

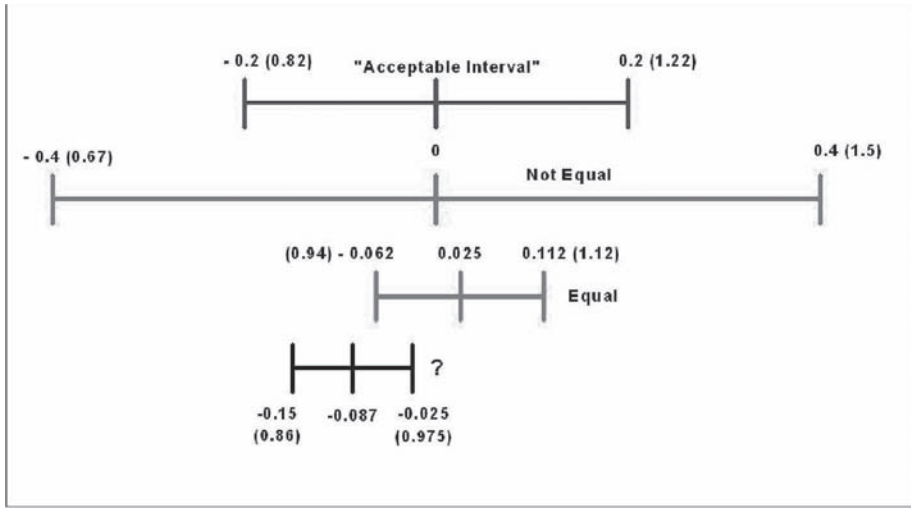


Figure 6.10
A graphical view of acceptable intervals.

include zero but may also include values that would suggest ratios in the two samples are, in fact, not equal. One way to address this problem is to define an "acceptable interval," which includes the range of LRD values one might expect in samples that have a common source.

As discussed next, determination of an acceptable interval is specific to a given situation—that is, how different can similar samples be?

If the probability that the mean LRD value from a given comparison is outside the acceptable interval is substantial (say, >0.10 or 0.20), the comparison does not provide strong support for the two samples having a common source. This approach is illustrated graphically in Figure 6.10. Here we have defined an acceptable interval of −0.2 to 2 in the log metric, which corresponds to a range of ratios from 0.82 to 1.22 (remember, differences in a log metric are equivalent to ratios in a linear metric). We also show three 95% confidence intervals for hypothetical sets of LRD values: −0.4 to 0.4, −0.62 to 0.112, and −0.15 to −0.025. The first is substantially outside the acceptable interval. In fact, if we assume that the interval is based on a *t*-distribution with 10 degrees of freedom, the probability of the true LRD being within the acceptable interval is only 0.35. In the second example, the 95% confidence interval is entirely within the acceptable interval and includes zero, which provides good support for the hypothesis that the two samples have similar composition. The third interval is slightly problematic; it is entirely within the acceptable interval, but does not include zero. In this case we have reason to reject the null hypothesis that the mean LRD is zero but still have evidence that it is within the interval one might expect for samples with similar composition.

6.4.3 COMPUTING PROBABILITIES FOR ACCEPTABLE INTERVALS

We first assume an acceptable interval with limits L1 and L2 (−0.2 to 0.2 or 0.82 to 1.22 in original units in our example). We then calculate the mean (m) and standard (s) deviation for LRD values for a pair of samples. The number of degrees of freedom (ν) is the number of ratios minus 1. That is, if we have 11 ratios as in Table 6.3, $\nu = 10$. We can then use a t distribution with ν degrees of freedom to calculate two probabilities. The first, P1, is the probability that the true mean LRD (μ) is less than L1. That is, P1 = P (μ < L1 | m, s); the probability that the true mean, μ, LRD is less than L1 given the observed sample mean M and standard deviation s of the mean LRD. The second probability, P2, is defined as P2 = P (μ > L2 | m, s); the probability that the true mean, μ, LRD is greater than than L2 given the observed sample mean M and standard deviation S of the mean LRD. Note that S is standard deviation of the LRD values divided by the square root of the sample size. To calculate these probabilities we first calculate t statistics as:

$$t_1 = (\text{L1} - \text{M})/\text{S} \quad \text{and} \quad t_2 = (\text{L2} - \text{M})/\text{S}$$

P1 is the probability that the standard t deviate is less than t_1 and P2 is the probability that the standard t deviate is greater than t_2. The probability, PA, that the true mean, μ, is in the acceptable interval is PA = 1 − (P1 + P2). Both P1 and P2 can be found using the probability calculation functions provided in many statistics packages. If PA is less than 0.8 to 0.9, evidence for similar composition is not compelling.

As an example, we will compare samples 1 and 2 (see Table 6.3). We find that the mean LRD for 1 versus 2 is 0.1657. The standard deviation of mean LRD12 is 0.537. Since we have 11 ratios we use a t-distribution with 10 degrees of freedom. Given an acceptable interval (AI) in logs of −0.2 to 0.2 we get $t_1 = -0.681$ and $t_2 = 0.064$. The probability of the true mean being in the acceptable interval is about 0.269 (27%) but if the acceptable interval were −0.8 to 0.8 (0.45 to 2.23), the probability of the true mean being in the acceptable interval would be about 0.816 (82%). Thus, the evidence for similar composition is equivocal. If the acceptable interval is narrow the two samples are arguably different but if the acceptable interval is wide, they may be similar.

For a second example, we compare samples 3 and 4 (see Table 6.3). The mean LRD for 3 versus 4 is −0.4192 and the standard deviation of the mean LRD34 is 0.2784. Again, we have 11 ratios so we use a t-distribution with 10 degrees of freedom. Given an AI of −0.2 to 0.2, the probability of the true mean being in the interval is about 0.20 (20%) but if the acceptable interval

were −0.8 to 0.8 (0.45 to 2.23), the probability of the true mean being in the acceptable interval is about 0.90 (49%). Again, the evidence for similar composition depends in the definition of similarity. The wider of the two intervals considered admits quite large differences so we would say that the evidence for similar composition is equivocal.

6.4.4 TWO-SAMPLE LOG-RATIO DIFFERENCE TESTS: A BOOTSTRAP ALTERNATIVE

LRD statistics allow us to get a formal statistical evaluation of whether two samples differ in composition, but if the data do not follow log-normal distributions, the hypothesis tests based on the *t*-distribution will be incorrect. In such cases, if the number of pairs is fairly large (say greater than 10) one can do bootstrap resampling (Efron and Tibshirani, 1993; Ginevan and Splitstone, 2003) to obtain a nonparametric estimate of the distribution of the log ratio difference between the two samples. A flowchart for bootstrap resampling is shown in Figure 6.11. Recall that the mean is given by the sum of the observations divided by the sample size. To obtain bootstrap bounds or a bootstrap distribution for the mean log ratio difference, we resample the data, with replacement, to obtain a large number (say, 5000) of bootstrap samples, and then calculate the sample mean for each of these samples. The result would be 5000 sample means based on these bootstrap samples. From this set of 5000 sample means we can determine a nonparametric 95% upper bound on the mean as the 250th largest of the sample means generated, or a 90% upper bound as the 500th largest of the sample means generated. Similarly if we have an acceptable interval with limits L1 and L2, we can count all values of the bootstrap distribution (the 5000 means generated) between L1 and L2, and divide this count by 5000, which will give an approximation of the probability that the true mean is between L1 and L2.

We did 5000 bootstrap samples for the 11 pairs of bootstrap samples shown in Table 6.3 and obtained probabilities of the mean LRD being in the −0.2 to 0.2 interval of 0.285 for LRD12 and 0.196 for LRD34. For the wider intervals we estimated probabilities of 0.86 and 0.93 for LRD12 and LRD34, respectively. The point here is that one can get reasonable approximations to the probabilities calculated using the *t*-distribution using bootstrap resampling even when the sample size is quite small. Note that 11 observations are definitely on the small side for bootstrap methods (Ginevan and Splitstone, 2003). It is also true that the significance tests for the log-log regression methods discussed earlier can be approximated using bootstrap methods that do not depend on the assumption of log-normality. This is beyond the scope of the present discussion but is treated in detail in Efron and Tibshirani (1993).

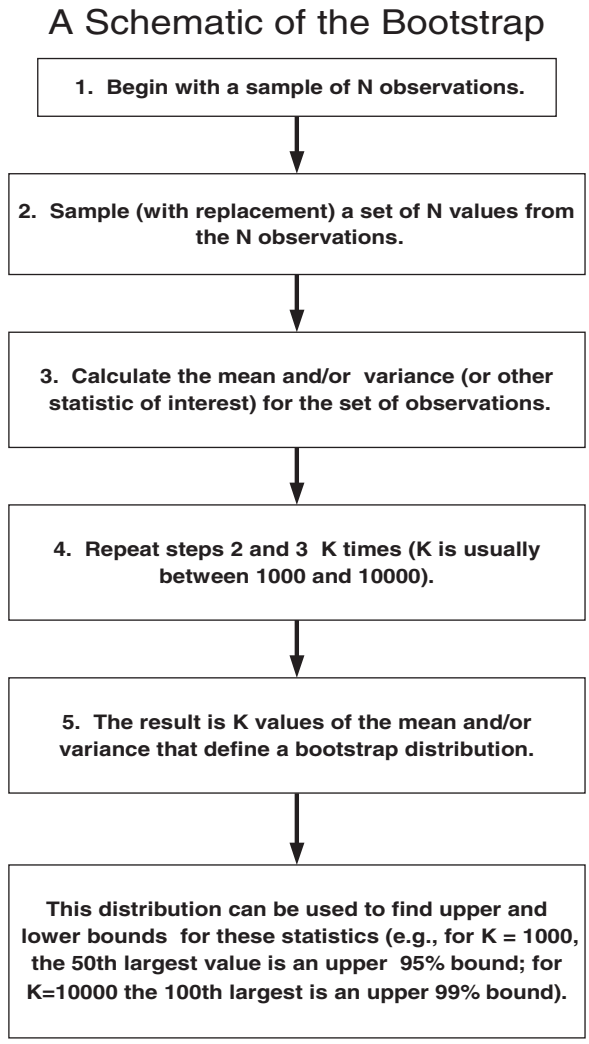


Figure 6.11
A flowchart for bootstrap resampling.

6.4.5 MANY SAMPLES

If we have many samples and a number of interesting pairs to form log ratios, several options are possible. First, we could do a number of pairwise tests and use something like the Bonferroni correction (Ginevan and Splitstone, 2003) to account for the multiple comparisons problem. If we have several areas and want to know if there are differences in ratios among areas, analysis of variance procedures could be used to test whether the average ratios in the areas are different. One could also perform principal components analysis (PCA) to get a look at the multivariate patterns in the log(ratio) metric. An extensive account of methods for dealing with multisample sets of log-ratio data can be found in Aitchison (2003).

6.4.6 LOG(RATIO) DATA: CONCLUSIONS

The question is, of course, is the log(ratio) metric as good as the more commonly employed scaled concentration approaches? The log(ratio) method is certainly useful in some applications. The methods discussed in detail in this section are useful in getting a clear look at how relative composition changes across samples and for putting the alleged similarity of single samples into perspective. More importantly, log(ratio) data are expected to be normal if the chemicals involved in the ratios follow log-normal distributions, thus one can apply a variety of normal statistical theory methods (e.g., *t*-tests) with some confidence. In the final analysis the procedures discussed here, though not necessarily better, provide a different perspective, and environmental forensics investigations will always benefit from examining the data from a variety of valid perspectives.

6.5 FINAL CONCLUSIONS

The unifying principal of the methods presented here is that relative composition is an informative metric to consider in environmental forensics investigations and that the logarithmic transformation is useful in evaluations of relative composition. One might be concerned that under logarithmic transformation small changes are given as much weight as large ones. That is, the change from 1 to 10 is, under log transformation, the same as the change between 100 and 1000. However one must remember that the focus is relative composition, not absolute concentration changes. Moreover environmental data usually is characterized by many relatively low concentration samples and only a few high concentration samples. Thus a logarithmic transform normalizes the data and makes it easier to see low-level patterns. One might also ask how one can choose appropriate pairs for a log-ratio analysis. Such selection is a decision like many others in an environmental forensics study. Where should we sample? Which chemical species should we measure the concentrations of? Similarly we want to determine which pairs might have interesting ratios. All such decisions are based on professional knowledge and expert judgment.

REFERENCES

Aitchison, J. (2003) *The Statistical Analysis of Compositional Data.* The Blackburn Press, Caldwell, NJ.

Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap.* Chapman and Hall, London.

Fuller, W.A. (1987) *Measurement Error Models.* New York, Wiley.

Ginevan, M.E. and Splitstone, D.E. (2003) *Statistical Tools for Environmental Quality Measurement*. Chapman & Hall/CRC, Boca Raton, FL.

Hogg, R.V. and Craig, A.T. (1995) *Introduction to Mathematical Statistics*. Prentice Hall, Englewood Cliffs, NJ.

LaKind, J.S., Graves, C.G., Ginevan, M.E., Jenkins, R.A., Naiman, D.Q., and Tardiff, R.G. (1999a) Exposure to environmental tobacco smoke in the workplace and the impact of away-from-work exposure, *Risk Analysis* 19(3), 343–352.

LaKind, J.S., Jenkins, R.A., Naiman, D.Q., Ginevan, M.E., Graves, C.G., and Tardiff, R.G. (1999b) Use of environmental tobacco smoke constituents as markers for exposure, *Risk Analysis* 19(3), 353–367.

LaKind, J.S., Ginevan, M.E., Naiman, D.Q., James, A.C., Jenkins, R.A., Dourson, M.L. *et al.* (1999c) Distribution of exposure concentrations and doses for constituents of environmental tobacco smoke, *Risk Analysis* 19(3), 369–384.

Ott, W.R. (1994) *Environmental Statistics and Data Analysis.* CRC Press, Boca Raton, FL.

Rawlings, J.O. *et al.* (1998) *Applied Regression Analysis: A Research Tool*. Springer-Verlag, New York, Chapter 10; 334–338.

Sokol, R.R. and Rohlf, F.J. (1995) *Biometry.* W. H. Freeman, NY.

CHAPTER 7

PRINCIPAL COMPONENTS ANALYSIS AND RECEPTOR MODELS IN ENVIRONMENTAL FORENSICS

Glenn W. Johnson, Robert Ehrlich, William Full, and Scott Ramos

7.1 Introduction 208
- 7.1.1 Philosophy and Approach: A Case for Exploratory Data Analysis 208
- 7.1.2 Formal Description of the Receptor Modeling Problem 210
- 7.1.3 Demonstration Data Sets 212

7.2 Principal Components Analysis 214
- 7.2.1 PCA Overview 214
- 7.2.2 Data Transformations 217
- 7.2.3 Eigenvector Decomposition 220
- 7.2.4 Determining the Number of Significant Principal Components 222
  - 7.2.4.1 Single Index Methods 223
  - 7.2.4.2 Variable-by-Variable Goodness of Fit 226
- 7.2.5 PCA Output 231

7.3 Self-Training Receptor Modeling Methods 233
- 7.3.1 Polytopic Vector Analysis (PVA) 235
  - 7.3.1.1 Scaling Functions: Back-Calculation to Original Metric 236
  - 7.3.1.2 Eigenvector Decomposition and Determining the Number of End-Members 237
  - 7.3.1.3 Determining End-Member Compositions and Mixing Proportions 238
  - 7.3.1.4 Results of PVA Applied to Data Set 2 243
- 7.3.2 Alternating Least Squares 244
  - 7.3.2.1 Initial estimates 245
  - 7.3.2.2 Least squares 245
  - 7.3.2.3 Constraints 245
  - 7.3.2.4 Convergence 246
- 7.3.3 Target Transformation Factor Analysis 246

7.3.4 Extended Self-Modeling Curve Resolution 248
7.3.5 Positive Matrix Factorization 250
7.4 The Influence of Alteration Processes on Mixing Models 251
7.5 Summary 252
Acknowledgments 254
References 254

7.1 INTRODUCTION

The identification of chemical contaminant sources is a common problem in environmental forensic investigations. Successful inference of sources depends on sampling plan design, sample collection procedures, chemical analysis methods, and knowledge of historical industrial processes in the study area. However, in complex situations where multiple sources contribute similar types of contaminants, even careful project planning and design may not be enough. If sources cannot be linked to a unique chemical species (i.e., a tracer chemical), then mapping the distributions of individual contaminant concentrations is insufficient to infer source. If, however, a source exhibits a characteristic chemical fingerprint defined by diagnostic proportions of a large number of analytes, source inference may be accomplished through analysis of multiple variables; that is, through use of multivariate statistical methods. The objective of a multivariate approach to chemical fingerprinting is to determine (1) the number of fingerprints present in the system, (2) the multivariate chemical composition of each fingerprint, and (3) the relative contribution of each fingerprint in each collected sample.

Development of numerical methods to determine these parameters has been a major goal in environmental chemometrics and receptor modeling for more than 25 years. The result has been development of a series of procedures designed to accomplish this. Procedures developed early in this history are most useful in solving relatively simple problems. Later procedures are designed to solve more general problems, which take into account complications such as bad data, commingled plumes (i.e., mixing of source fingerprints), and the presence of sources not assumed or anticipated at the start of an investigation. The objective of this chapter is to discuss this family of procedures in terms of their strengths and limitations, in order to guide the working environmental scientist in the use of appropriate procedures.

7.1.1 PHILOSOPHY AND APPROACH: A CASE FOR EXPLORATORY DATA ANALYSIS

In terms of experimental design, the source apportionment problem in environmental forensic investigations falls between two extremes. At one extreme,

all potential sources are known in terms of their chemical composition, location, history, and duration of activity. At the other extreme, none of these are known with any certainty. Chemicals at the receptor (e.g., estuary sediments, groundwater at a supply well) may be the result of activities long absent or distant from the vicinity of the site.

In the first case (*a priori* knowledge of all sources) the problem is a relatively simple one. Appropriate sampling locations can be determined using a conventional experimental design, which is part of conventional experimental statistics. Determination of contribution of each source can be extracted using a variety of linear methods, such as chemical mass balance receptor models (see Chapter 8). However, even when the contributing sources are known, environmental forensic investigations often prove to be more complex than initially anticipated. Chemicals in the environment may not retain their original composition. That is, chemical compositions may change over time by processes such as biodegradation and volatilization. This is true even for relatively recalcitrant contaminants such as polychlorinated biphenyls and dioxins (Bedard and Quensen, 1995; Chiarenezelli *et al.*, 1997; Johnson *et al.*, 2006). The result of degradation will be resolution of one or more fingerprints that were not originally anticipated.

At the other extreme, where nothing is known with certainty, potential sources may be suspected, but samples of the sources (i.e., fingerprint reference standards) may not have been collected, and may not exist in the literature. The industrial history of a region may be imperfectly known. A small, low profile operation may be a major but completely overlooked source of contamination. For cases toward this end of the spectrum, we must take leave of the elegance of conventional experimental statistics, and move into the realm of exploratory data analysis (EDA). The fundamental difference between these two approaches (experimental statistics and EDA) is that the former is associated with creation of explicit hypotheses, and evaluation of data in terms of well-defined tests and strong probabilistic arguments. In contrast, the objective of EDA is to find patterns, correlations, and relationships in the data itself, with few assumptions or hypotheses (Tukey, 1977). If the fruits of an EDA result in a map where the concentrations of a multivariate fingerprint increase monotonically toward an effluent pipe, and the fingerprint composition is consistent with the process associated with that source, the obvious inference is that the potential source is the actual source. We recognize that we are not working in the realm of classical statistics or formal hypothesis testing, and that EDA is based on less rigorous probabilistic statements. However, such an approach should not be construed as second best. In environmental forensics, an EDA approach may be the only valid option. The cost of highly rigorous probabilistic methods is often a set of narrow assumptions regarding the

structure of the data. Such methods cannot be supported if those critical assumptions cannot be safely made. Moreover, in environmental forensic investigations, we often lack sufficient information to know what hypotheses to test. Thus, at least at the beginning of a project, an EDA approach is usually most prudent.

Regardless of the data analysis strategy chosen, another important consideration is the presence of bad or questionable data. Common problems with environmental chemical data include the following: (1) chemical analyses performed by different laboratories or by different methods, which may introduce a systematic bias; (2) the presence of data at concentrations at or below method detection limits; (3) the presence of coelution (nontarget analytes that elute at the same time as a target analyte); and (4) the all-too-common problem of error in data entry, data transcription, or peak integration.

Unfortunately such errors rarely manifest themselves as random noise. More often, they contribute strong systematic variability. If unrecognized, the result may be derivation of fingerprints that have little to do with the true sources. Therefore, a necessary adjunct to any data analysis in environmental forensics is identification of outliers. As we proceed with a discussion of fingerprinting methods, a major consideration in that regard must be inclusion of vigilant outlier identification and data cleaning procedures. If such an effort results in deletion or modification of data, the data must be clearly identified, and justification for the action must be provided in the narrative that accompanies the analyses.

7.1.2 FORMAL DESCRIPTION OF THE RECEPTOR MODELING PROBLEM

Receptor modeling in environmental forensics involves the inference of sources and their contributions through analysis of chemical data from the ambient environment (Gordon, 1988; Hopke, 1991). The objectives are to determine (1) the number of chemical fingerprints in the system; (2) the chemical composition of each fingerprint; and (3) the contribution of each fingerprint in each sample. The starting point is a data-table of chemical measurements in samples collected from the receptor (e.g., estuarine sediments, ambient air in a residential area). These data usually are provided in spreadsheet form where rows represent samples and columns represent chemical analytes. To the multivariate data analyst this table is a matrix. We will refer to the original data table as the m row by n column matrix $\mathbf{X}$, where m is the number of samples and n is the number of analytes. We wish to know the number of fingerprints present (k) and chemical composition of each (objectives 1 and 2). This can be expressed as a matrix $\mathbf{F}$, which has k rows

and n columns. We also wish to know a third matrix (**A**), which has m rows and k columns, and represents the contribution of each chemical fingerprint in each sample (objective 3). Thus the following linear algebraic equation formally expresses the receptor modeling problem.

Matrix dimensions

$$\underset{(m\times n)}{\mathbf{X}} = \underset{(m\times k)}{\mathbf{A}}\underset{(k\times n)}{\mathbf{F}} + \underset{(m\times n)}{\boldsymbol{\varepsilon}} \tag{7.1}$$

where **A** is a matrix of source contributions or mixing proportions (with m rows and k columns), **F** is a matrix of source profiles or fingerprint chemical compositions (with k rows and n columns), and matrix $\boldsymbol{\varepsilon}$ is a matrix of residuals or measurement error. For $i = 1$ to m, $j = 1$ to n, and given k sources, each element in **X** can be found from:

$$x_{ij} = \sum_{h=1}^{k} a_{ih} f_{hj} + e_{ij} \tag{7.2}$$

Given our table of raw data, we have three knowns (**X**, m, and n), and three unknowns: (1) the number of fingerprints (k); (2) the fingerprint compositions (**F**); and (3) the contributions of each fingerprint in each sample (**A**). In the rare case where fingerprints are known *a priori*, both k and **F** are known, and only one unknown (**A**) remains. If this is the case, source compositions are contained within a "training data set" consisting of the compositions of all potential sources, and solving Equation 7.1 for **A** is straightforward. The problem can be solved by regression techniques, which are the basis of chemical mass balance (CMB) approaches (see Chapter 8). Unfortunately, in environmental forensics investigations, we seldom have the luxury of *a priori* knowledge of contributing sources. The use of a training data set constitutes, if not a formal hypothesis test, at least an implicit hypothesis test. If at all possible, we would like to derive source patterns directly from analysis of ambient data. That is, we want to employ chemometric methods that are self-training.

The number of chemometric methods available to us is large, however, the nature of environmental forensics investigations dictates the use of methods with very special features: they must allow for the conceptual model of mixtures of multiple sources; and they should allow resolution of contributing chemical fingerprints without *a priori* assumption of the number, chemical composition, or geographic/temporal distribution. Finally, the results must be interpretable in a scientific context. Several commonly used chemometric methods, which satisfy these considerations, are presented next.

7.1.3 DEMONSTRATION DATA SETS

The methods demonstrated in this chapter are illustrated using polychlorinated biphenyls (PCB) data. PCBs are a group of chlorinated organic compounds that are commonly the focus of environmental forensic investigations. PCBs were widely used in commercial and industrial settings for much of the twentieth century. Commercial PCB products were marketed under the trade name Aroclor by Monsanto (the former U.S. manufacturer). Commercial applications of PCBs included their use in fluorescent light ballasts, carbonless copy paper, and as dielectric fluids in electrical transformers and capacitors. PCBs are of concern in environmental studies because they are persistent, toxic, and tend to bioaccumulate in tissues of higher predators (Tanabe *et al.*, 1987). Johnson *et al.* (2006) provides a comprehensive review of PCB chemistry, sources, fate, and transport, all presented in context of environmental forensics investigations. PCBs are used for demonstrations in this chapter because they are a group of contaminants that often require a multivariate approach to data analysis. Commercial PCB formulations (e.g., Aroclors) have unique chemical compositions composed of multiple congeners, but no single congener is a diagnostic tracer for one and only one source.

Two synthetic PCB data sets will be used for demonstration purposes. Each of these represents a three-source system. The congener patterns for the three-source fingerprints were taken from Aroclor standard compositions reported by Frame *et al.* (1996). The source compositions are Aroclor 1248 (Frame sample G3.5) and two variants of Aroclor 1254 (Frame samples A4 and G4). Frame *et al.* (1996) first reported markedly different congener patterns for different lots of Aroclor 1254. Subsequent investigations by Frame indicated that the atypical Aroclor 1254 was the result of a late production change in the Aroclor 1254 manufacturing process that occurred in the early 1970s (Frame, 1999). Subsequent work narrows the time frame of late-production Aroclor 1254 between 1971 and 1972 (Larry Hansen, personal communication; Johnson *et al.*, 2006). The congener compositions of these three-source fingerprints are shown as bar graphs in Figure 7.1.

The two variants of Aroclor 1254 were used for this demonstration because it provides a typical example of the surprises often encountered in environmental forensic investigations. For example, given a situation where (1) the production history of an area was well established; (2) it was known with great certainty that only two PCB formulations were ever used at a site: Aroclor 1248 and Aroclor 1254; and (3) the data analyst was not aware of the two Aroclor 1254 variants; he or she might reasonably make an *a priori* assumption of a two-source system. Clearly, in this case, that assumption would be incorrect.

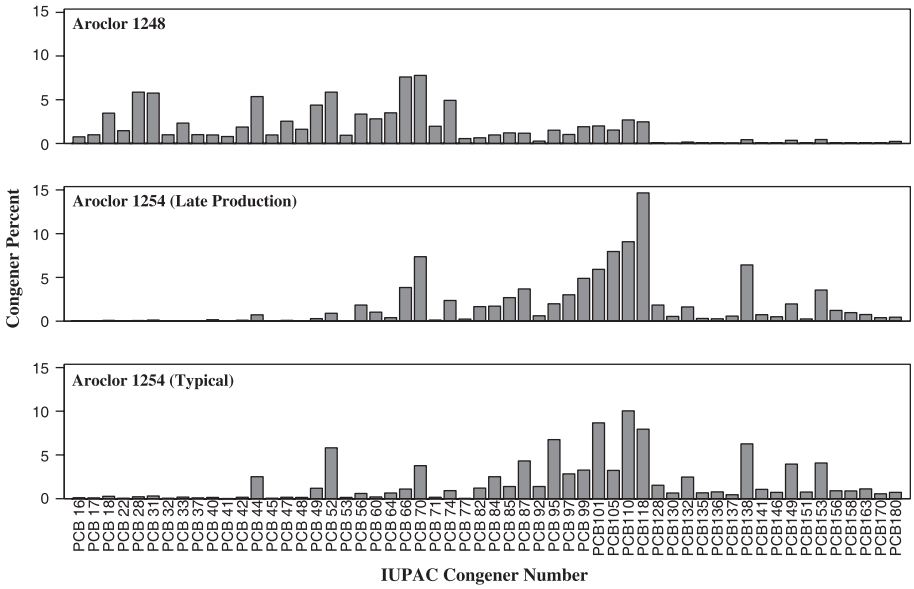


Figure 7.1
Congener compositions of three PCB product formulations (Aroclors) used to create the artificial data sets used for demonstration in this chapter (data from Frame et al.*, 1996).*

In environmental forensics investigations, even the simplest, well-understood systems can yield surprises.

These three Aroclor compositions were used to create two synthetic data sets. The first of the two (Data Set 1) is quite simple, almost to the point of being unrealistic as an analogue for environmental forensics investigations. However, it is instructive in that it provides a good intuitive understanding of principal components analysis (PCA), which is the mathematical basis of all the methods presented here. Data Set 1 is a 24-sample, 56-congener matrix. The data are simple in that each sample represents a contribution from one and only one source. That is, it is a strongly clustered data set, with no samples representing mixtures of more than one Aroclor. To create some intersample variability within each of the three source categories, random Gaussian noise was added to each sample in the data set. Data Set 1 is included as Table 7A.1 of the Chapter 7 Appendix.

The second data set (Data Set 2) is considerably more complex, and is more representative of data encountered in environmental forensics investigations. Data Set 2 is shown in Table 7A.3, and was created such that:

(1) All samples are *mixtures* of the three Aroclor compositions. That is, no sample in the data set represents a 100% contribution from a single source. Varying contributions of multiple sources impacts every sample. The two matrices **A** and **F** (Equation 7.1) used to calculate the original noise-free matrix **X** are shown in Table 7.2. For ease of presentation, the transposed 56×3 matrix $\mathbf{F}^t$ (rather than the original matrix **F**) is shown in Table 7A.2.

(2) Ten percent Gaussian noise was added to simulate random error. For each matrix value $\mathbf{X} = x_{ij}$ a random value was drawn from a normal distribution with mean zero and standard deviation 1. This random number was multiplied by 0.10, and that result was multiplied by the matrix value x_{ij} to create the noise term for that matrix element. This noise value was then added to the original matrix value x_{ij} and the sum substituted for the original x_{ij}.

(3) The data are represented in units of concentration (ng/g), and a method detection limit was established for each sample. As such, many low concentration matrix elements are censored as a function of the detection limits. Nondetects are indicated in Table 7.3 with a U qualifier. For subsequent numerical analyses, we adopt the common practice of replacing nondetect values with half the detection limit.

(4) Data transcription errors were added to one sample (Sample 22). The error simulated was a data translation mistake. The reported concentrations of some congeners were offset by one, thus assigning incorrect concentrations to 14 of the 56 spreadsheet columns.

(5) For the congener PCB 141, a coelution problem was introduced in 35 of the 50 samples. Coelution of non-PCB peaks with PCB congeners during gas chromatographic analysis is a common problem in PCB chemistry. The coelution simulated here represents one such example. The pesticide *p,p′*-DDT elutes very close to PCB 141 on a Chrompack CP-SIL5-C18 GC column (R. Wagner, personal communication). Therefore, if *p,p′*-DDT is present in a sample undergoing congener specific PCB analysis, the *p,p′*-DDT could erroneously be reported as PCB congener IUPAC 141.

7.2 PRINCIPAL COMPONENTS ANALYSIS

7.2.1 PCA OVERVIEW

Principal components analysis (PCA) is a widely used method in environmental chemometrics, as it is in many scientific disciplines. PCA is used on its own, and as an intermediate step in receptor modeling methods. Before presenting the computational steps in PCA, it is useful to provide a more intuitive discussion, using Data Set 1 as an example. The objective of PCA is to reduce the dimensionality of a data set in which there are a large number of interrelated (i.e., correlated) variables. This reduction in dimension is achieved by transforming the data to a new set of uncorrelated reference variables (principal components or PCs). The PCs are sorted such that each in turn accounts for a progressively smaller percentage of variance within the data set. If nearly all variability between samples can be accounted for by a small number of PCs, then relationships between multivariate samples may be assessed by simple

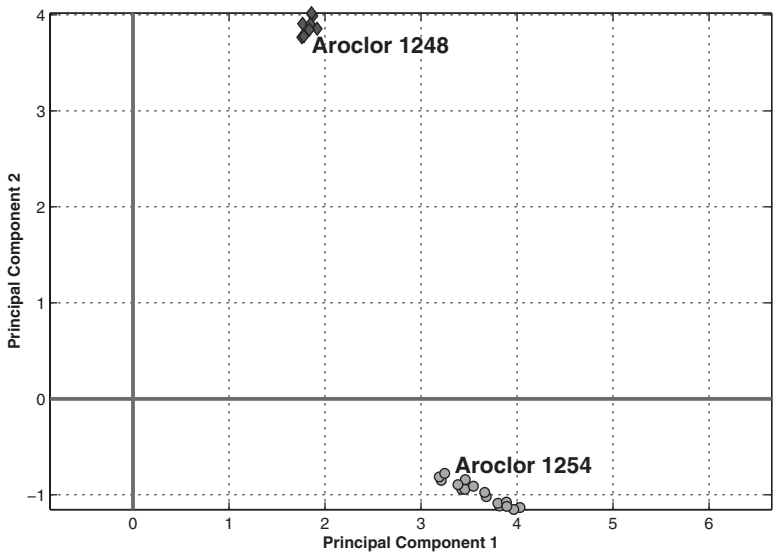


Figure 7.2
Two principal component score plots of Data Set 1. Two PCs account for >92% of the variance of Data Set 1, but are insufficient to allow distinction of the two Aroclor 1254 variants.

inspection of a two- or three-dimensional plot: a principal components scores plot. Figure 7.2 shows a two-PC scores plot for Data Set 1. Two principal components account for more than 92% of the variance in Data Set 1, and the scores plot clearly divides the samples into two clusters: Aroclor 1248 and Aroclor 1254.

There is an obvious problem: although this is indeed a two Aroclor system (as Figure 7.2 clearly suggests) it is not a two-source system. The first two PCs do not differentiate between the two Aroclor 1254 variants. This illustrates a common problem in the application of PCA to environmental chemical data. All too often, investigators will present a two-PC scores plot like Figure 7.2, accompanied by a statement of justification indicating that two principal components account for 92.5% of the variance. Such a statement leaves the tacit implication that the residual 7.5% of the variance is random noise, which in this case is clearly not the case.

A three-component scores plot for Data Set 1 is shown in Figure 7.3. Three PCs account for 97.5% of the variance; an incremental increase over the percentage accounted for by two PCs. However, that small percentage of total variance is *not* random. The three-PC scores plot clearly distinguishes three clusters rather than two, and effectively allows the analyst to infer the presence of three sources.

This example highlights several important precautionary notes.

(1) For better or for worse, the use of mathematical techniques such as PCA carries with it the aura of precision and exactitude. In the case of the two-PC scores plot

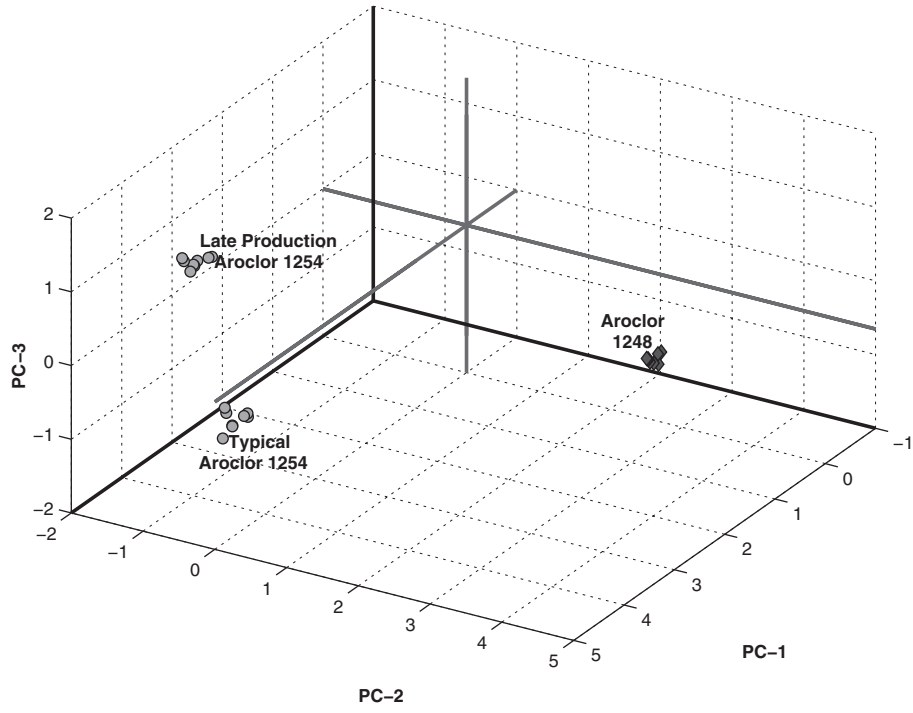


Figure 7.3
Three principal component score plots of Data Set 1. Three PCs account for 97.5% of the variance, and allow clear distinction of the three PCB sources.

(see Figure 7.2) the strong clustering into two Aroclor groups, coupled with the statement that two PCs account for 92.5% of the variance, may be sufficiently intimidating to impress skeptics. Moreover, it may even provide a false sense of security to the naïve analyst.

(2) In an environmental forensics setting, we seldom have prior knowledge of the true data structure, so we should not be arrogant in our application of rules-of-thumb regarding what percentage of variance should be considered significant.

(3) Scientific and legal significance is clearly *not* a function of variance. If a party involved in environmental litigation had used Aroclor 1254, but ceased all operations in the mid-1960s (prior to Monsanto's 1971–1972 change in the Aroclor 1254 production process), the small difference in percentage of variance accounted for between two and three PCs would have enormous implications.

(4) Practitioners of PCA-based methods in environmental forensics must employ more sophisticated goodness-of-fit diagnostics than percentage variance, and they must have a reasonable understanding of how such methods work.

Because PCA is a powerful exploratory data analysis tool on its own, as well as an intermediate step in receptor modeling methods, we will discuss this method in considerable detail later. The key steps in PCA include (1) data transformations; (2) singular value decomposition (eigenvector decomposi-

tion); (3) determination of the number of significant eigenvectors; and (4) visual display of scores and loadings plots. Each of these steps is discussed in turn, next.

7.2.2 DATA TRANSFORMATIONS

To the data analyst, the laboratory results received from the chemist are raw data. In environmental chemistry these raw data usually are transmitted in units of concentration. Data Set 2 (see Table 7A.3: presented in units of ng/g) is a good example. However, seldom is a statistical analysis performed on a matrix in this form. Rather, the matrix is transformed in some manner. A data transformation is the application of some mathematical function to each measurement in a matrix. Taking the square root of every value in a matrix is an example. In analysis of environmental chemical data, data transformations are done either for (1) reasons related to the environmental chemistry of the system; or (2) mathematical reasons, to optimize the analysis.

A common transformation done for chemical or physical reasons is sample normalization. In an environmental system, concentrations can vary widely due to dilution away from a source. For example, in the case of contaminated sediment investigations, concentrations may decrease exponentially away from an effluent pipe. However, if the relative proportions of individual analytes remain relatively constant then we might infer a single source scenario, coupled with dilution away from the source. Inference of source by pattern recognition techniques is concerned more with the relative proportions between analytes than with absolute concentrations. Thus, a transformation is necessary to normalize out concentration/dilution effects. One that is commonly employed is a transformation to a percent metric, where each value is divided by the total concentration of the sample. This transformation is very commonly done, and depending on the discipline in which it is used, it may be referred to as a constant row-sum transformation or an area percent normalization. Stated mathematically, given an m sample by n variable nonconstant row-sum matrix $\mathbf{X}_{(\text{conc})}$, and an $m \times m$ diagonal matrix $\mathbf{C}$ (with the m row-sums—total concentrations of samples—of $\mathbf{X}_{(\text{conc})}$ along the diagonal and zeros on the off-diagonals) the constant row-sum matrix $\mathbf{X}$ may be calculated as:

$$\mathbf{X} = 100 * \mathbf{C}^{-1} * \mathbf{X}_{(\text{conc})} \tag{7.3}$$

An alternative to the constant row-sum transformation is to normalize the data with respect to a single species or compound (a so-called normalization variable). This transformation involves setting the value of the normalization variable to 1.0, and the values of all other variables to some proportion of 1.0,

such that their ratios with respect to the normalization variable remain the same as in the original metric.

The second type of transformation is done more for mathematical/statistical purposes. In any chemical data set, there is usually a strong relationship between the mean value of an analyte and its variance (variance is square of the standard deviation). Therefore, chemicals measured in trace concentrations almost always exhibit smaller variance than those measured at much higher concentrations. Multivariate procedures such as PCA are variance driven, so in the absence of some transformation across variables, the analytes with highest mean and variance usually have the greatest influence on the analysis. Polychlorinated dibenzo-*p*-dioxins provide a particularly instructive example. In most environmental systems, 2,3,7,8-TCDD (dioxin) typically is measured at orders of magnitude at lower concentrations than the octachlorinated congener OCDD. Thus the mean and variance of 2,3,7,8-TCDD is typically orders of magnitude lower than OCDD. However, this does not imply less precision or accuracy in the chemical measurement of 2,3,7,8-TCDD. Nor does it imply that 2,3,7,8-TCDD is of secondary environmental importance to OCDD. In fact, the opposite is usually the case because 2,3,7,8-TCDD has much higher toxicity than does OCDD (van den Berg *et al.*, 1998). As such, analysis of environmental chemical data often require some sort of homogeneity of variance transformation. A number of transformations may be applied to produce homogeneity of variance. One of the most commonly used transformations in PCA is the autoscale transform (also known as the Z transform): given a matrix **X**, with calculated means and standard deviations (s_j) in each column ($j = 1, 2 \ldots n$) the autoscaled matrix **Z** is calculated:

$$\mathbf{Z}_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \tag{7.4}$$

The autoscale transformation guarantees absolutely equal variance in that it sets the mean of each column to 0.0 and the standard deviation to 1.0. A related scaling choice is Pareto scaling (Wold *et al.*, 1993), which divides the numerator of Equation 7.4 by the square root of the standard deviation s_j.

Another common homogeneity of variance transformation is the range-transform (also known as the minimum/maximum transformation). Following the convention of Miesch (1976a), where the original matrix is denoted **X**, the range is denoted as *X*-prime ($\mathbf{X}' = \{X'_{ij}\}$). The transformation is performed as follows:

$$x'_{ij} = (x_{ij} - x_{\min j})/(x_{\max j} - x_{\min j}) \tag{7.5}$$

This results in a matrix where the minimum value in each column equals 0.0 and the maximum equals 1.0. The range transformation produces variances that are approximately homogeneous (unlike the Z transform, which results in absolute homogeneity of variance). However, the range transform has two advantages: (1) it carries no implication of a standard normal distribution; and (2) because transformed values are within the 0.0 to 1.0 range, all values are nonnegative. The latter feature is particularly useful in receptor modeling, where explicit constraints of nonnegativity are used to derive source fingerprints. Malinowski (1977, 1991) noted that a disadvantage of the range transform is that it is extremely sensitive to outliers. As we will see, however, in an environmental forensics investigation, detection and evaluation of outliers is a crucial part of the process, so this feature of the range transform is usually desirable.

A third transformation (the equal vector length transform) is often, but not always, employed in multivariate analysis of chemical data. The equal vector length transform is performed to force each of the sample vectors to have equal Euclidean length. If all vectors have equal length, then the differences between samples are a function only of the angles between samples. Thus the similarities and differences between samples can be expressed as a similarity matrix of cosines (Davis, 1986). The cosine between two identical samples is 1.0. In contrast, the cosine between two completely dissimilar samples (i.e., vectors at 90°) is 0.0. By Miesch's convention (Miesch, 1976a), the constant-sum input matrix is indicated as **X**, the range transform matrix as **X′**, and the equal vector length transform as **X″**. For matrix algebra-based programming languages, Hopke (1989) presented a computationally efficient way to calculate **X″** by first defining matrix **Y**, a $m \times m$ diagonal matrix where the off-diagonal elements are zero and diagonal elements y_i equal the inverse of the square root of the sum-of-squares along rows of **X′**:

$$y_i = \frac{1}{\left(\sum_{j=1}^{n} x_{ij}'^2 \right)^{1/2}} \tag{7.6}$$

The transformed matrix **X″** may then be calculated as follows:

$$\mathbf{X''} = \mathbf{YX'} \tag{7.7}$$

This transformation has an added advantage for receptor modeling. By definition, if all samples have equal vector lengths, each sample must lie on an $n - 1$ dimensional surface, unit length from the origin. To demonstrate this, a simple three-chemical example is shown in Figure 7.4. In this case

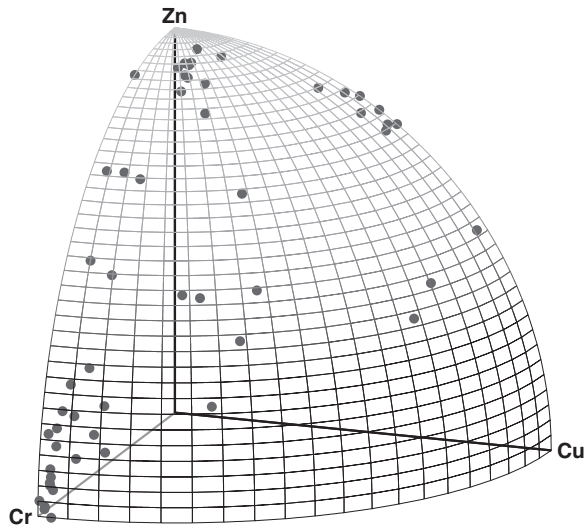


Figure 7.4
*A simple three chemical system (*n = *3: chromium, copper, and zinc) illustrates the two-dimensional (*n − *1) curved plane geometry of a matrix that has undergone the equal vector length transformation. All sample vectors are unit length from the origin, and lie on the* n − *1 dimensional spherical surface.*

n = 3, because only three variables (chromium, copper, and zinc) are present. These data were transformed by Equations 7.6 and 7.7; thus all samples lie on a two-dimensional surface, which is unit length from the origin ($n - 1 = 2$). As we will see in Section 7.3, receptor/mixture modeling involves resolution of a $k - 1$ dimensional geometric figure within k dimensional principal component space. Thus, this transformation has particularly attractive features in that regard.

In summary, normalization across rows (constant row sum or normalization to a marker chemical) is almost always done in environmental chemometric analyses. Some homogeneity of variance transformation (e.g., range transform or autoscale transform) is also typically performed in order to keep high concentration variables from dominating an analysis. The constant vector length transform is sometimes done, usually when (1) there is some advantage to being able to express relationships between samples simply in terms of angles between samples; or (2) when the PCA is an intermediate step in receptor modeling, which involves resolution of a $k - 1$ dimensional simplex within k dimensional space. The PCA that resulted in the plots shown in Figures 7.2 and 7.3 involved transformation by the constant row sum transformation (Equation 7.3) followed by the range transform (Equation 7.5). The constant vector length transformation was not used.

7.2.3 EIGENVECTOR DECOMPOSITION

Eigenvector decomposition is a simple mathematical procedure that allows a reduction in dimensionality of a data set. This is the core mathematical opera-

tion involved in principal components analysis. It is most often accomplished through singular value decomposition (SVD) of the transformed matrix $\mathbf{X}'$ or $\mathbf{X}''$. As shown in Figure 7.4, the transformed $m \times n$ matrix may be thought of as m vectors plotted in n dimensional space. Each variable is represented as one of n orthogonal axes of a cartesian coordinate system. There are, however, an infinite number of sets of n mutually orthogonal basis vectors that may equivalently be used to plot the sample vectors without loss of information. The set of eigenvectors extracted from a similarity matrix of the original data is one such alternative reference space. The number of eigenvectors (i.e., the number of principal components) will equal m or n, whichever is smaller. However, there are usually correlations between analytes due to common or similar sources. Thus, a relatively small subset of eigenvectors is typically sufficient to capture the variability in the system, and the interrelationships between samples can be observed in a reduced dimensional reference space without loss of information.

Eigenvector decomposition is a well-established part of the core knowledge of mathematics and is frequently used in the physical and natural sciences. The calculation of eigenvectors and eigenvalues is relatively straightforward, but lengthy and cumbersome. As such, a conceptual discussion of the topic is presented next, and you are referred to any number of linear algebra texts for a complete mathematical discussion. Davis (1986) provides a detailed, lucid but less rigorous treatment, using examples from the earth sciences.

Given an error-free, noise-free matrix of $m \geq k$ samples and $n \geq k$ variables, resulting from k sources, only k nonzero eigenvectors and eigenvalues will be extracted. If $k = 3$, the first eigenvector will account for a high percentage of the total variance in the data set. The second eigenvector is constrained in that it must be mutually orthogonal to the first, and accounts for the highest percentage of residual variance (that variance not accounted for by the first eigenvector). The third eigenvector is mutually orthogonal to the first two, and accounts for the remainder of the variance. The data set may equivalently be expressed in this three-dimensional reference space, without loss of information. Given a transformed matrix $\mathbf{X}'$ composed of m samples along the rows, and n variables (chemical analytes) along columns, PCA is accomplished through SVD of $\mathbf{X}'$:

$$\mathbf{X}' = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{V}' \tag{7.8}$$

where $\mathbf{U}$ equals the eigenvector matrix ($m \times k$) resulting from a decomposition of $[\mathbf{X}']\ [\mathbf{X}']^t$. $\mathbf{V}$ is the eigenvector matrix ($n \times k$) resulting from a decomposition of $[\mathbf{X}']^t[\mathbf{X}']$. $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues. Alternatively, the diagonal matrix of the square root of the eigenvalues ($\mathbf{\Lambda}^{1/2}$) may be referred

to as matrix **S**, the singular values. For principal components analysis, Equation 7.8 is typically reexpressed in terms of principal components scores and loadings as follows:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}^{1/2} \tag{7.9}$$

$$\mathbf{F} = \mathbf{V}^{t} \tag{7.10}$$

Although the mathematics of eigenvector decomposition and PCA is consistent across the literature, the jargon unfortunately is not. In the mathematical geology literature, **A** is termed *loadings* and **F** is termed *scores* (Miesch, 1976; Zhou *et al.*, 1983; Ehrlich and Full, 1987; Ehrlich *et al.*, 1994; Johnson *et al.*, 2000). Exactly the *opposite* terminology is used in the chemometrics literature: **A** is termed *scores* and **F** is termed *loadings* (Malinowski, 1980, 1991; Brereton, 1992; Massart *et al.*, 1997; Tauler *et al.*, 2000). It is not clear when or why this divergence in jargon occurred, but both nomenclatures are used in recent applied environmental chemistry source apportionment studies, and often within the same journals (e.g., Johnson *et al.*, 2000; Tauler *et al.*, 2000). Given that this chapter covers PCA and receptor modeling methods from both sets of literature, as these terms are used through the remainder of this chapter, the jargon will be clarified in the context of each particular discussion. The previous discussion of score-plots in Section 7.2.1 (see Figures 7.2 and 7.3) follows the chemometrics convention.

The total number of calculated PCs equals m or n, whichever is smaller. A typical objective of PCA, however, is to reduce the dimensionality of the data set. This is done either to facilitate visualization of scores and loadings, or to use the reduced dimensional PCA space as an alternative cartesian coordinate system for receptor modeling. A model involving a reduced number of principal components (k) may be represented as follows:

$$\underset{(m\times n)}{\mathbf{X}'} = \underset{(m\times k)}{\mathbf{A}'}\,\underset{(k\times n)}{\mathbf{F}'} + \varepsilon \tag{7.11}$$

where k equals the number of PCs retained for the model, and ε represents error or residuals. Criteria for determining k are presented in Section 7.2.4.

7.2.4 DETERMINING THE NUMBER OF SIGNIFICANT PRINCIPAL COMPONENTS

The choice of the number k is equivalent to making a decision on the number of significant principal components. Of the many aspects of PCA-based

methods, no topic has created more argument or controversy than the criteria used to determine the correct number of eigenvectors (i.e., k, the number of factors, sources, subpopulations, or end-members).

Numerous methods have been proposed for determination of k (Cattell, 1966; Exner, 1966; Miesch, 1976a; Malinowski, 1977; Wold, 1978; Ehrlich and Full, 1987; Henry *et al.*, 1999). The spirit and intent of these methods are similar: the estimated data set, as back-calculated from reduced dimensional space (i.e., $\mathbf{X}_{\text{hat}}$ or $\hat{\mathbf{X}}$), should reproduce the measured data ($\mathbf{X}$) with reasonable fidelity.

7.2.4.1 Single Index Methods

Six criteria commonly used in chemometrics and mathematical geology were applied to Data Sets 1 and 2, and the results are shown in Tables 7.1 and 7.2, respectively. The PCA of both data sets involved the constant sum and range

Table 7.1

Reproduction Indicies for Data Set 1 (see Table 7A.1).

PC No.	Eigenvalue	(1) Cumulative Percentage Variance	(2) Screen Test	(3) Normalized Varimax Loadings	(4) Malinowski Indicator Function	(5) Cross Validation PRESS(i)/ RSS($i-1$)	(6) Signal-to-Noise Ratio
1	220.430	57.36	42.64	16	0.000674		39.32
2	135.150	92.54	7.46	8	0.000315	0.20	35.04
3	19.212	**97.53**	**2.47**	**16**	**0.000204**	**0.43**	**12.35**
4	1.198	97.85	2.15	1	0.000215	1.26	1.59
5	0.990	98.10	1.90	1	0.000229	1.41	1.35
6	0.861	98.33	1.67	0	0.000246	1.57	1.37
7	0.781	98.53	1.47	1	0.000266	1.76	1.41
8	0.719	98.72	1.28	1	0.000290	1.96	1.29
9	0.634	98.88	1.12	1	0.000318	2.19	1.34
10	0.556	99.03	0.97	0	0.000352	2.50	1.18
11	0.535	99.17	0.83	1	0.000392	2.79	1.27
12	0.475	99.29	0.71	1	0.000442	3.22	1.17
13	0.438	99.41	0.59	1	0.000503	3.67	1.21
14	0.382	99.50	0.50	1	0.000583	4.35	1.03
15	0.371	99.60	0.40	1	0.000681	5.00	1.15
16	0.317	99.68	0.32	1	0.000814	6.01	1.04
17	0.272	99.75	0.25	0	0.001001	7.11	1.09
18	0.213	99.81	0.19	1	0.001296	8.86	0.81
19	0.184	99.86	0.14	0	0.001769	11.31	0.81
20	0.169	99.90	0.10	0	0.002569	14.73	0.83
21	0.143	99.94	0.06	0	0.004157	21.04	0.84
22	0.126	99.97	0.03	1	0.007798	30.60	0.84
23	0.066	99.99	0.01	0	0.027834	308.49	0.64
24	0.043	100.00	0.00	0			0.00

Table 7.2

Reproduction Indicies for Data Set 2 (see Table 7A.3).

PC No.	Eigenvalue	(1) Cumulative Percentage Variance	(2) Screen Test	(3) Normalized Varimax Loadings	(4) Malinowski Indicator Function	(5) Cross Validation PRESS(*i*)/ RSS(*i*−1)	(6) Signal-to-Noise Ratio
1	585.600	77.92	22.08	50	0.000102	–	73.29
2	115.450	93.28	6.72	47	0.000059	0.33	31.89
3	17.807	**95.65**	4.35	**18**	0.000050	0.73	10.79
4	8.494	96.78	3.22	2	0.000046	**0.97**	6.54
5	4.751	97.41	2.59	1	0.000043	1.21	4.38
6	3.073	97.82	2.18	2	**0.0000421**	1.38	**3.30**
7	1.706	98.05	1.95	2	0.0000422	1.62	1.88
8	1.507	98.25	1.75	1	0.0000424	1.75	1.88
9	1.279	98.42	1.58	1	0.000043	1.88	1.85
10	1.055	98.56	1.44	0	0.000043	2.04	1.49
11	0.973	98.69	1.31	1	0.000044	2.19	1.60
12	0.898	98.81	1.19	1	0.000045	2.33	1.53
13	0.806	98.92	1.08	1	0.000046	2.54	1.51
14	0.782	99.02	0.98	1	0.000047	2.66	1.45
15	0.693	99.11	0.89	1	0.000048	2.81	1.31
16	0.596	99.19	0.81	1	0.000049	3.06	1.25
17	0.570	99.27	0.73	0	0.000050	3.29	1.28
18	0.532	99.34	0.66	1	0.000051	3.57	1.21
19	0.483	99.40	0.60	1	0.000053	3.83	1.16
20	0.434	99.46	0.54	1	0.000055	4.21	1.10
21	0.421	99.52	0.48	1	0.000056	4.52	1.14
22	0.371	99.57	0.43	0	0.000058	5.03	1.10
23	0.364	99.61	0.39	0	0.000060	5.25	1.10
24	0.313	99.66	0.34	0	0.000062	5.72	0.95
25	0.287	99.69	0.31	0	0.000065	6.33	0.94
26	0.274	99.73	0.27	0	0.000067	6.82	1.00
27	0.246	99.76	0.24	0	0.000070	7.74	0.93
28	0.245	99.80	0.20	0	0.000073	8.24	0.94
29	0.215	99.82	0.18	1	0.000076	8.89	0.88
30	0.178	99.85	0.15	0	0.000080	10.12	0.76
31	0.160	99.87	0.13	0	0.000084	11.25	0.72
32	0.142	99.89	0.11	1	0.000089	12.08	0.67
33	0.122	99.90	0.10	1	0.000095	13.78	0.62
34	0.111	99.92	0.08	0	0.000102	16.02	0.61
35	0.105	99.93	0.07	1	0.000109	18.24	0.58
36	0.093	99.95	0.05	0	0.000117	21.61	0.60
37	0.084	99.96	0.04	0	0.000125	22.87	0.54
38	0.060	99.96	0.04	0	0.000138	28.54	0.45
39	0.057	99.97	0.03	0	0.000152	32.49	0.44
40	0.047	99.98	0.02	0	0.000170	39.18	0.41
41	0.040	99.98	0.02	1	0.000191	46.38	0.40
42	0.031	99.99	0.01	0	0.000221	58.76	0.35
43	0.026	99.99	0.01	0	0.000260	73.10	0.32
44	0.020	99.99	0.01	0	0.000317	99.11	0.28
45	0.016	100.00	0.00	0	0.000396	140.67	0.26
46	0.014	100.00	0.00	0	0.000479	172.73	0.26
47	0.007	100.00	0.00	0	0.000697	317.28	0.19
48	0.003	100.00	0.00	0	0.001372	555.71	0.13
49	0.002	100.00	0.00	0	0.005275	–	0.12
50	0.002	100.00	0.00	0	–	–	0.00

transformations. Each of the six indices, and the rationale for their use, is discussed here.

(1) *Cumulative percentage variance.* The rationale for this criterion is simple: a reduced dimensional model should account for a large percentage of the variance in the original matrix. However, as discussed in Section 7.2.1, any *a priori* choice regarding what percentage of variance one should consider to be significant is problematic. Workers in multivariate statistics, chemometrics, and mathematical geology generally acknowledge that any proposed cutoff criterion is arbitrary (Malinowski, 1991; Deane, 1992; Reyment and Jöreskog, 1993). The lack of a clear criterion makes the cumulative variance method dubious. Nonetheless, in Tables 7.1 and 7.2, we note that the commonly used cutoff of 95% suggests retention of three principal components for both Data Sets 1 and 2.

(2) *Screen test.* The screen test of Cattell (1966) is based on the supposition that the residual variance, not accounted for by a k principal component model, should level off at the point where the principal components begin accounting for random error. When residual variance is plotted versus principal component number, the point where the curve begins to level off should show a noticeable inflection point, or *knee*. The problem with this criterion is that there is often no unambiguous inflection point, and when such is the case, the decision as to the number of significant principal components is arbitrary.

(3) *Normalized varimax loadings.* The use of this index was described Ehrlich and Full (1987). If an eigenvector carries systematic information, then typically, a large number of samples will have high loadings (mathematical geology jargon) on that varimax factor. High numbered factors that account for noise and little variance typically have loadings that are small for all samples (i.e., <0.1). Factors with many samples >0.1 indicate principal components that should be retained for a model. This criterion calls for rotation of principal components using the varimax procedure of Kaiser (1958), normalization of the varimax loadings matrix to sum to 1.0 across all sample rows, and tabulation of the number of samples that exceed 0.1 for each factor. The analyst looks for a sharp drop in the index as an indication of the appropriate number of eigenvectors. Again, if there is no clear, sharp decline in this index, the appropriate decision can be ambiguous.

(4) *Malinowski indicator function.* Malinowski (1977) presented an indicator function, which is calculated as a function of the residual standard deviation (Malinowski, 1977; Hopke, 1989). The function reaches a minimum when the correct number of principal components is retained. The index has worked well with relatively simple data structures but Hopke (1989) reported that it has not proven as successful with complex environmental chemical data.

(5) *Cross-validation*. Cross-validation is a commonly used method for determination of number of significant principal components. It involves successive deletion of data points, followed by prediction of the deleted data with increasing numbers of eigenvectors. The PRESS statistic (Predicted Residual Error Sum of Squares) is then calculated for each number of potential eigenvectors. Many criteria have been proposed based on calculation of some function of PRESS (Wold, 1978; Eastman and Krzanowski, 1982; Deane, 1992; Grung and Kvalheim, 1994). The PRESS values and criteria in Tables 7.1 and 7.2 are those presented by Deane (1992).

(6) *Signal-to-noise ratio*. This method was presented by Henry *et al.* (1999). Henry's NUMFACT criterion involves calculation of a signal-to-noise (S/N) ratio. Henry found that given random data, an S/N ratio as high as 2.0 could be obtained. Based on that, the rule-of-thumb criterion recommended by Henry is that principal components with S/N ratios greater than 2.0 should be retained for a model.

As is evident in Table 7.1, for Data Set 1 (a relatively simple data set with random Gaussian noise) each of these indices provides an accurate estimate of the true number of sources. All six indices correctly indicate a three-component system. Table 7.2 reports the values for the same indices, as applied to the more complicated, error-laden Data Set 2. Here, the reproduction indices suggest anywhere between three sources and six sources. Clearly, the complications present in Data Set 2 (which are quite reasonable in terms of common environmental chemistry scenarios) are sufficient to introduce ambiguity between these various indices.

This ambiguity is due in part to the fact that all of these are single-index methods. Each involves calculation of a single numerical value or statistic, which represents the data set as a whole as a function of the number of principal components retained. The data analyst typically compares the behavior of the index as additional PCs are retained, relative to some rule-of-thumb cutoff criterion. The idea of a rule-of-thumb decision criterion (i.e., a minimum, a change in slope, a threshold) is troublesome in exploratory data analysis, because we have very little information to evaluate the efficacy of these rules. In such situations, we need other tools to gain deeper insight into the chemical system.

7.2.4.2 Variable-by-Variable Goodness of Fit

Single index methods (in particular, criteria based on percentage of variance) are misleading because they carry the tacit assumption that variability *not* accounted for by a reduced dimensional model is spread evenly across all originally measured variables. Miesch (1976a) presented a method that addresses this problem. Miesch's method evaluated goodness on a variable-

by-variable basis, rather than for the data set as a whole. The variance accounted for by *each* of the originally measured variables is evaluated for each potential number of principal components. Given an m sample by n variable data matrix $\mathbf{X}$ of rank m or n (whichever is smaller) the index used by Miesch (1976a) was the coefficient of determination (CD) between each variable in the original data matrix ($\mathbf{X}$), and its back-calculated reduced dimensional equivalent ($\hat{\mathbf{X}}$). For each number of potential eigenvectors, 1, 2 . . . *rank*, Miesch calculated an $n \times 1$ vector:

$$r_j^2 \cong \frac{s(x)_j^2 - s(d_j)^2}{s(x)_j^2} \tag{7.12}$$

where $s(x)_j^2$ is the variance of values in the jth column of $\mathbf{X}$, and $s(d_j)^2$ is the variance of residuals between column j of $\mathbf{X}$ and column j of $\hat{\mathbf{X}}$. Miesch used the $\cong$ in this equation presumably because he recognized that this was not a conventional r^2 or coefficient of determination (CD) as defined by least squares linear regression. It is not the variance accounted for by the least squares regression line of $\mathbf{X}_j$ and $\hat{\mathbf{X}}_j$. Rather, the Miesch CD is the r^2 with respect to a line of one-to-one back-calculation between $\mathbf{X}_j$ and $\hat{\mathbf{X}}_j$. For CDs less than 1.0, the analyst must make some decision as to what value may be accepted. That decision is made in context of the analyst's experience, knowledge of measurement error (if available), and scientific context. As an example, if a certain PCB congener is known to be less accurate and/or less precise using a certain gas chromatography (GC) column, the data analyst may justifiably accept a lower CD for that congener, than for other congeners.

A graphical extension of Miesch's method (the CD scatter plot) has been implemented by Johnson (Johnson, 1997; Johnson *et al.*, 2000). The appropriate graphic to illustrate the fit of the Miesch CD is a series of n scatter plots that show the measured value for each variable $\mathbf{X}_j$ plotted against the back-calculated values from the k proposed principal components ($\hat{\mathbf{X}}_j$).

A scatter plot array for Data Set 2 is presented in Figure 7.5, and shows 56 plots (one for each PCB congener) as back-calculated from a three-PC model. The Miesch CD for each congener is calculated and reported at the top left corner of each scatter plot. When an insufficient number of principal components are retained there should be evident nonrandom deviations from the 1:1 fit line. For three principal components a good fit is observed for most congeners, but there is a systematic lack of fit observed for (1) nondetect censored data points (indicated as squares); (2) Sample 22, which had a data transcription error (triangle); and (3) PCB congener 141, which coelutes with DDT. In particular, note that on many graphs, there are two nondetect samples at high measured concentration. These are Samples 9 and 17 from Table 7A.3.

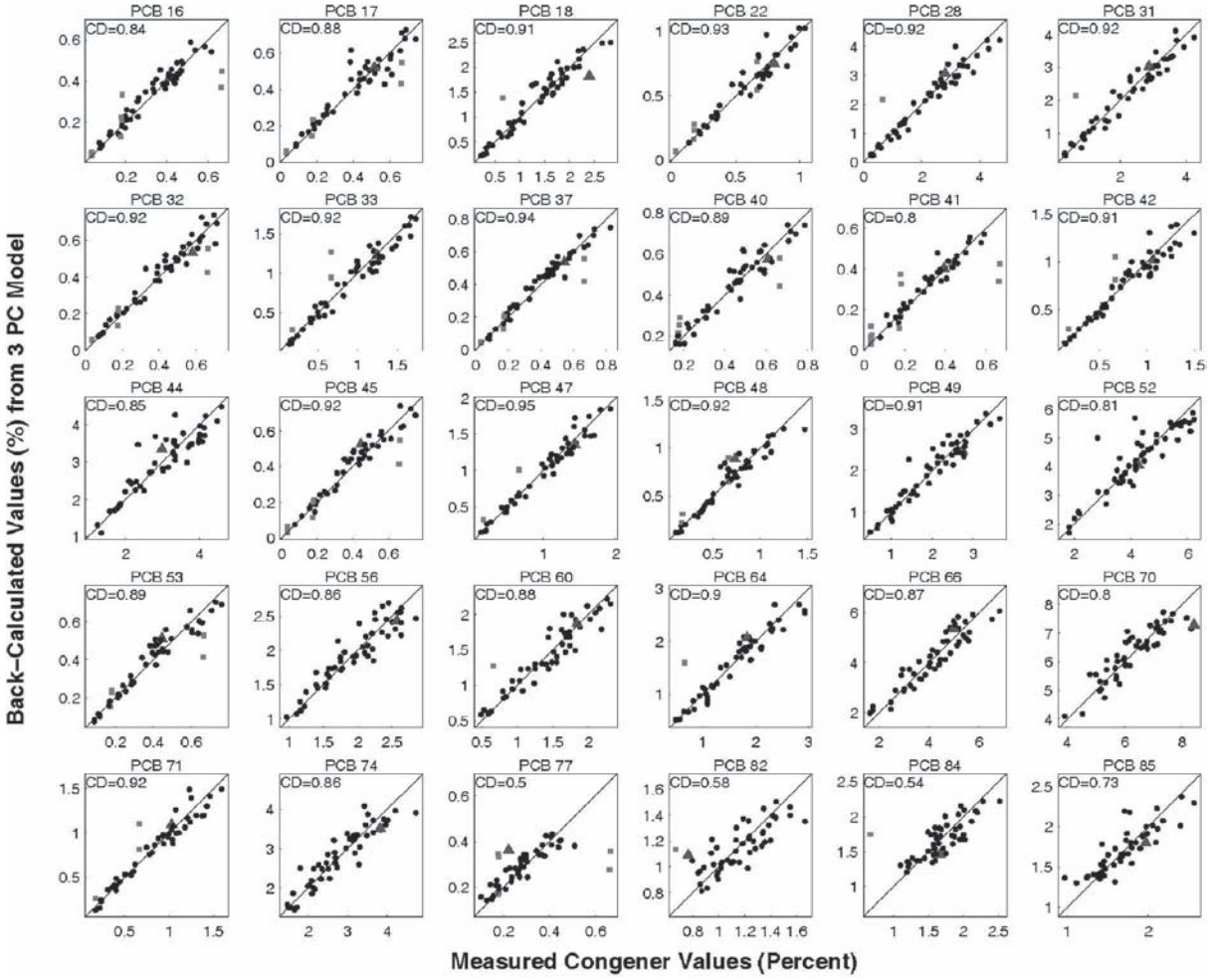


Figure 7.5A

Goodness-of-fit scatter plot array and Miesch coefficients of determination (CDs) for first 30 variables in Data Set 2. The x *axis is measured concentration. The* y *axis is the value back-calculated from a three principal component model. Nondetects (censored data points) are indicated as squares* (□). *Sample 22, which had a data transcription error, is indicated as a triangle* (△). *A full color version of this figure is available at books.elsevier.com/companions/9780123695222.*

Both had low total PCB concentrations, and yielded nondetects for more than half of the reported PCB congeners.

The important point with regard to determining the number of significant principal components is that these errors, although not related to chemical sources fingerprints, are not random. Rather, they represent systematic signal within the data set, and thus they greatly influence indices shown in Tables 7.1 through 7.3.

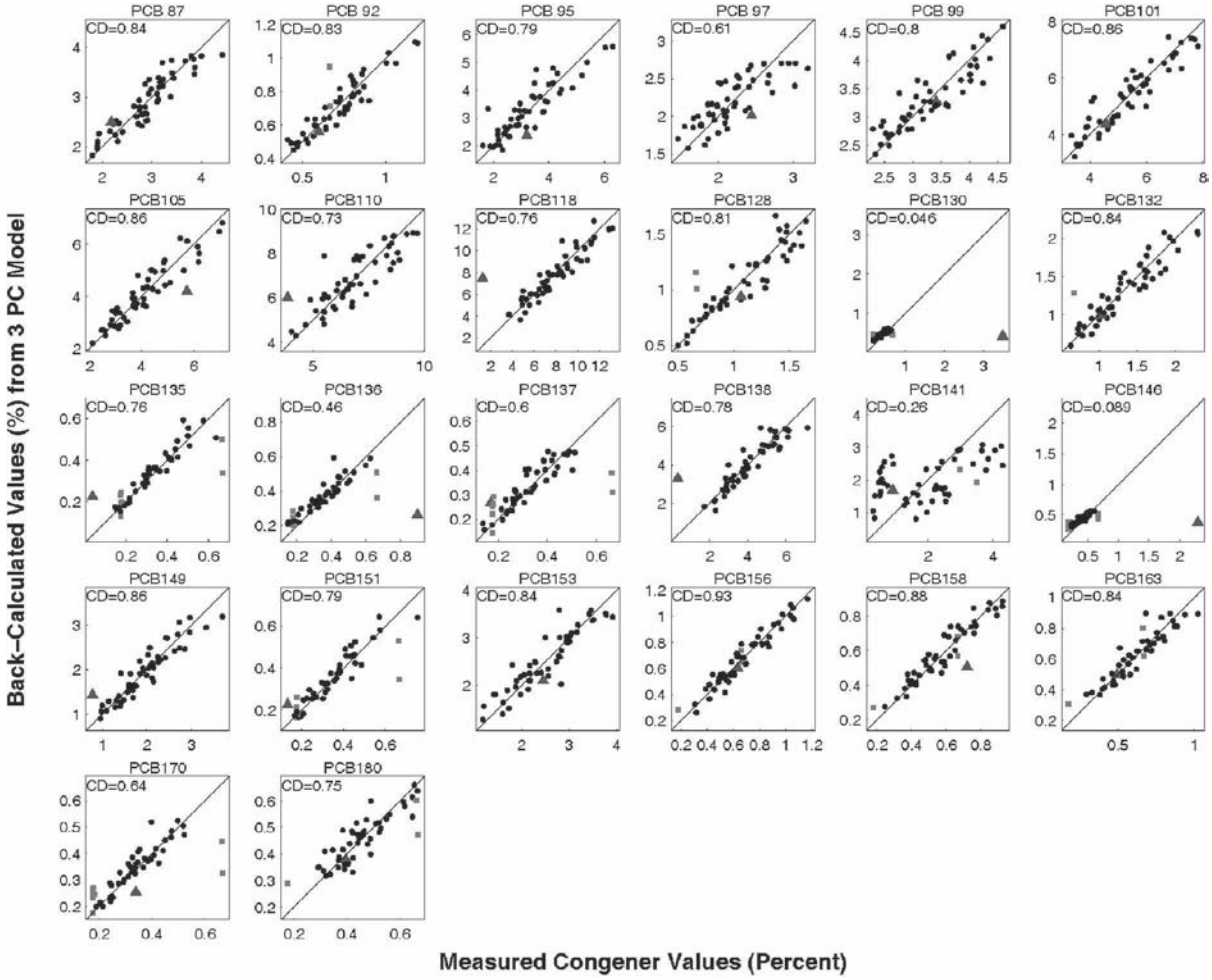


Figure 7.5B

Goodness-of-fit scatter plot array and Miesch coefficients of determination for last 26 variables in Data Set 2. The x *axis is measured concentration. The* y *axis is the value back-calculated from a three principal component model. Nondetects (censored data points) are indicated as squares* (□). *Sample 22, which had a data transcription error, is indicated as a triangle* (△). *Note lack of fit for PCB 141 due to DDT coelution. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.*

The strength of CD scatter plots is that they allow rapid evaluation of sample-by-sample goodness of fit and variable-by-variable goodness of fit. In addition, they facilitate outlier detection. These are evaluated simultaneously, as a function of the number of principal components retained. That combination quickly leads the analyst to more insightful and direct questions than are possible based solely on what threshold numerical index one judges

acceptable. Having identified these outliers, the questions that we must now ask are:

(1) What is the cause of deviation from good fit? Is it random error or systematic variability not accounted for by k PCs?
(2) If the cause of the observed deviation is systematic, is it due to (1) data entry error, (2) analytical error, or (3) presence of an additional source of variability in the field?
(3) How does one evaluate the number of significant PCs in the presence of such deviations?

In most cases, these questions cannot be answered in the realm of chemometrics and numerical data analysis. Rather the decision must be made in a larger context: geographic/temporal distribution of samples, analytical error, data entry error, method detection limits, as well as the possibility of an additional source, present in only one or a few samples. The analyst must switch hats and play the role of forensic chemist rather than numerical data analyst. As discussed earlier, decisions of significance are often best made in the scientific context of the investigation, rather than through use of a rule-of-thumb threshold criteria. In the case of Data Set 2, the appropriate decision is different for each type of outlier:

(1) In the case of PCB 141 coeluting with DDT, the appropriate decision is to have the chemist go back and reanalyze the chromatograms to ensure that PCB 141 (a shoulder on the DDT peak) is correctly quantified.
(2) In the case of the data transcription error, the appropriate decision is to correct the errors in the spreadsheet, and rerun the analysis.
(3) In the case of the two low concentration samples with multiple censored data points, usually the only realistic solution is to remove those two samples from the data set.
(4) In the case of the remaining censored data points, the nondetects are generally at the low end of the measured range (as is evident in the scatter plots) and thus do not adversely affect the accuracy of back-calculation. We usually wish to retain as many samples as possible in the analysis. Therefore, in this case, typically we would leave these remaining nondetect samples in the matrix.

These changes were made to Data Set 2. The PCA was rerun, and the revised reproduction indices are shown in Table 7.3. As for the much simpler Data Set 1, these indices are now in general agreement with each other, correctly indicating the presence of a three-component system.

Although all-too-seldom used in evaluation of principal components models, the use of scatter plots in context of multivariate goodness of fit evaluation is not at all new. They are standard in evaluation of linear regression models (Draper and Smith, 1981).

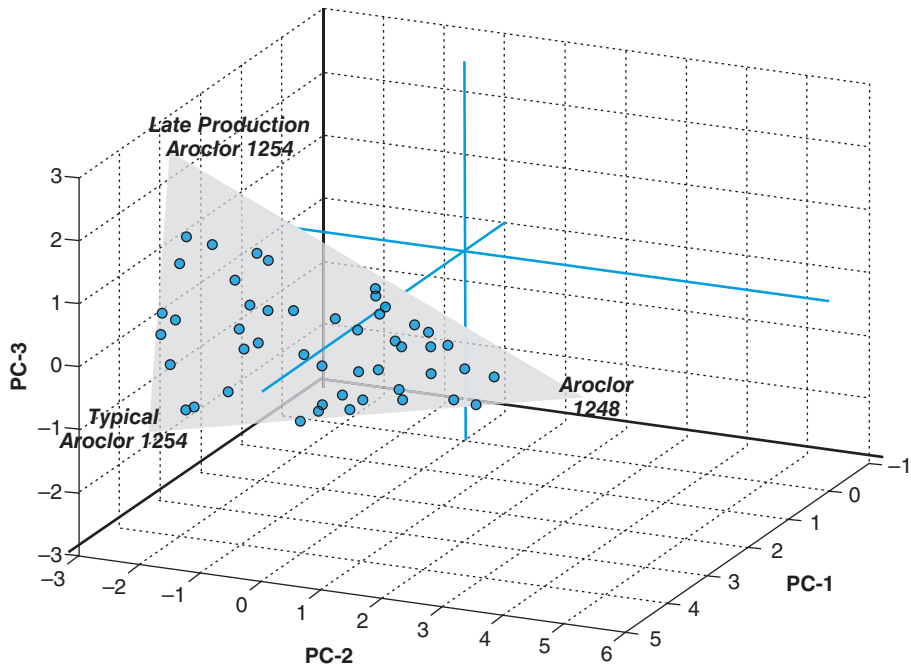


Figure 7.6
Three PC score plots for Data Set 2, after errors have been corrected. No samples in Data Set 2 (dots) represent pure Aroclor sources. All are mixtures of the three sources. The approximate locations of pure sources are indicated at the vertices (corners) of the gray-shaded triangle.

7.2.5 PCA OUTPUT

As discussed in Section 7.2.1, the most common way of presenting PCA results is in terms of PCA score plots (chemometrics jargon), where the analyst can evaluate relationships between samples on a two- or three-dimensional graphic. Scores plots for Data Set 1 were presented as Figures 7.2 and 7.3. For simple, clustered data sets such as this, the score-plot visualization method is very effective. The problem, however, is that regardless of the complexity of the data or the results of goodness-of-fit diagnostics (see Section 7.2.4), graphical limitations dictate visualization of a maximum of three principal components at a time. This puts subtle pressure on the analyst to choose three or fewer components. Therefore, the widespread occurrence of two and three component plots in the literature may be due more to this bias than to the inherent simplicity of environmental chemical systems.

Figure 7.6 shows the three PC score plot (chemometrics jargon) for cleaned Data Set 2. For reference, the corners of the gray shaded triangle represent the approximate locations of the three Aroclor sources. All Data Set 2 samples plot within a triangle defined by these three vertices. This data cloud geometry is commonly observed when samples are mixtures of multiple sources.

Numerous software packages, including most general-purpose packages, perform PCA and allow the user numerous data transformation and visualization options. A few of these include Statistical Analysis System (SAS: Cary,

Table 7.3

Reproduction indices for modified Data Set 2. For correctable errors (DDT coelution with PCB 141 and data transcription error in Sample 22) the data were modified accordingly. For uncorrectable problems (low concentration samples with many nondetects in two samples 9 and 17) the samples were removed from the matrix.

PC#	Eigenvalue	(1) Cumulative Percent Variance	(2) Screen Test	(3) Normalized Varimax Loadings	(4) Malinowski Indicator Function	(5) Cross Val. Press(i)/ RSS(i–1)	(6) Signal-to-Noise Ratio
1	561.330	76.26	23.74	46	0.000117	–	38.27
2	135.180	94.63	5.37	44	0.000059	0.25	14.83
3	19.919	**97.34**	**2.66**	**22**	0.000044	**0.57**	**2.55**
4	2.191	97.63	2.37	2	**0.0000434**	1.09	1.87
5	1.642	97.86	2.14	1	0.0000438	1.18	1.90
6	1.416	98.05	1.95	1	0.0000443	1.27	2.14
7	1.321	98.23	1.77	0	0.0000448	1.31	1.98
8	1.147	98.38	1.62	0	0.000046	1.40	1.76
9	1.025	98.52	1.48	1	0.000046	1.50	1.85
10	0.905	98.65	1.35	1	0.000047	1.61	1.59
11	0.839	98.76	1.24	1	0.000048	1.75	1.67
12	0.819	98.87	1.13	0	0.000050	1.77	1.58
13	0.712	98.97	1.03	0	0.000051	1.89	1.54
14	0.656	99.06	0.94	0	0.000052	2.02	1.54
15	0.612	99.14	0.86	1	0.000054	2.16	1.35
16	0.561	99.22	0.78	1	0.000055	2.31	1.37
17	0.511	99.29	0.71	1	0.000057	2.50	1.39
18	0.489	99.35	0.65	1	0.000059	2.74	1.30
19	0.482	99.42	0.58	1	0.000061	2.92	1.30
20	0.439	99.48	0.52	1	0.000063	3.13	1.31
21	0.412	99.53	0.47	0	0.000065	3.26	1.22
22	0.357	99.58	0.42	0	0.000068	3.60	1.16
23	0.339	99.63	0.37	1	0.000071	3.90	1.18
24	0.315	99.67	0.33	0	0.000074	4.15	1.07
25	0.276	99.71	0.29	0	0.000077	4.67	1.10
26	0.269	99.75	0.25	0	0.000081	4.83	0.94
27	0.227	99.78	0.22	0	0.000085	5.47	0.96
28	0.218	99.81	0.19	0	0.000089	6.03	0.89
29	0.199	99.83	0.17	1	0.000094	6.26	0.81
30	0.162	99.86	0.14	0	0.000101	7.05	0.77
31	0.148	99.88	0.12	1	0.000108	7.67	0.71
32	0.129	99.89	0.11	1	0.000116	8.79	0.71
33	0.119	99.91	0.09	1	0.000126	10.20	0.73
34	0.117	99.93	0.07	0	0.000136	11.61	0.68
35	0.106	99.94	0.06	0	0.000147	12.95	0.64
36	0.086	99.95	0.05	0	0.000162	15.57	0.59
37	0.079	99.96	0.04	0	0.000178	15.72	0.45
38	0.053	99.97	0.03	0	0.000204	19.88	0.46
39	0.050	99.98	0.02	0	0.000236	22.65	0.41
40	0.041	99.98	0.02	0	0.000279	28.45	0.40
41	0.037	99.99	0.01	0	0.000335	36.13	0.40
42	0.035	99.99	0.01	1	0.000403	40.03	0.31
43	0.020	99.99	0.01	0	0.000538	59.53	0.30
44	0.019	100.00	0.00	0	0.000740	74.99	0.26
45	0.013	100.00	0.00	0	0.001170	118.92	0.26
46	0.011	100.00	0.00	0	0.002074	175.26	0.19
47	0.006	100.00	0.00	0	0.006138	–	0.13
48	0.002	100.00	0.00	0	–	–	0.00

NC), Number Cruncher Statistical System (NCSS: Kaysville, Utah), and Pirouette (Infometrix, Bothell, WA).

As is evident in Figure 7.6, PCA can be effective in inferring the presence of a two- or three-component mixed system, but by itself PCA is not capable of determining the chemical compositions of the sources, or their relative contributions. In the case of analysis of mixtures, we need another technique in our numerical toolbox. Such tools are discussed in Section 7.3.

7.3 SELF-TRAINING RECEPTOR MODELING METHODS

An increasingly common approach in environmental forensic investigations involves the use of receptor models or mixing models. These models are designed to resolve three parameters of concern in a multivariate, mixed chemical system: (1) the number of components in the mixture, (2) the identity (i.e., chemical composition or fingerprints) of each component, and (3) the relative proportions of each component in each sample. These objectives are stated mathematically as the determination of k, **A**, and **F** in Equation 7.1. The source apportionment equation is similar to the scores and loadings expressions given in Equations 7.9, 7.10, and 7.11. However, principal component scores and loadings are abstract, orthogonal matrices, and do not represent feasible chemical compositions or source contributions. In particular, they are abstract in that when expressed in terms of the original chemical variables, principal component scores and loadings typically contain negative elements.

Receptor modeling methods therefore involve rotation of matrices **A** and **F** to an oblique solution derived within reduced (k-dimensional) principal component space. The rotation is performed per explicit nonnegativity constraints imposed on matrices **A** and **F**. The reason for explicit nonnegativity constraints are simple. A chemical fingerprint cannot have a chemical composition with, for example, –10% PCB 138; thus matrix **F** should not contain negative elements. Similarly, a sample cannot have a –35% contribution from a given source; thus matrix **A** cannot have negative elements (Miesch, 1976a; Full *et al.*, 1981, 1982; Gemperline, 1984; Hopke, 1989; Henry and Kim, 1990; Tauler, 1995; Windig and Heckler, 1995; Kim and Henry, 1999). Nonnegativity constraints are typical requirements for all receptor models. In addition, other explicit restrictions may be imposed, if one has additional knowledge of physical/chemical constraints on the system (Henry and Kim, 1990; Tauler, 1995; Johnson, 1997; Kim and Henry, 1999).

Multivariate receptor modeling methods typically use PCA as an intermediate step to determine the number of contributing sources, and sometimes also

to provide a reduced dimensional reference space for resolution of the model. Thus, in all the receptor modeling methods discussed here, determining the number of sources (k) essentially reduces to the problem of choosing the correct number of principal components. As such, discussions regarding determination of significant principal components (see Section 7.2.4) are equally relevant to receptor modeling.

An assumption of the conceptual mixing model/receptor model is that the system must be overdetermined. That is, the data set must contain more variables or samples (whichever number is smaller) than there are sources. If we measure only four chemicals in each sample, and six sources contribute to the contamination, we cannot completely or realistically resolve the model. Another assumption is that of linear mixing. We assume that the relative proportions of variables in each source are fixed, and that source contributions are linearly additive. That is, as we increase the proportion of a source fingerprint in a sample, the variables that are characteristic of that fingerprint will increase proportionally (linearly) in that sample.

After the choice of k (see Section 7.2.4), the receptor model then resolves the chemical compositions of sources (**F**) and the contributions of the sources in each of the samples (**A**). Recall, however, that in environmental forensics investigations, we rarely have *a priori* knowledge of all sources. If possible, we would like to derive source patterns directly from analysis of ambient data. Several such methods have been used in environmental source apportionment investigations. Five of them will be discussed here: (1) the DENEG algorithm used in polytopic vector analysis (Full *et al.*, 1981, 1982); (2) the combined use of error estimates and nonnegative constraints in positive matrix factorization (PMF: Paatero and Tapper, 1994); (3) the unique vector iteration method used in target transformation factor analysis (TTFA: Roscoe and Hopke, 1981; Gemperline, 1984; Hopke, 1989; Malinowski, 1991); (4) the N-dimensional edge detection method used in extended self-modeling curve resolution (ESMCR: Henry and Kim, 1990; Henry *et al.*, 1994; Kim and Henry, 1999; Henry, 2003); and (5) Alternating Least Squares (ALS: Tauler, 1995; Tauler *et al.*, 1995).

These methods are analogous in that (1) they do not require a training data set; (2) they use PCA as a means for determining dimensionality (i.e., number of sources) of the data set; (3) they involve solution of quantitative source apportionment equations by development of oblique vector models; and (4) each involves the use of explicit nonnegative constraints.

The PVA algorithm has not been set out in any single paper, but rather in series of papers that spanned more than 20 years. This chapter provides the opportunity to outline the entire PVA algorithm. As such, the mathematics of PVA will be discussed in greater detail than the other methods. ALS, TTFA, ESMCR, and PMF are presented in less mathematical detail, with the objective

of providing an intuitive understanding of how those algorithms operate. You are referred to the original cited papers for more detail on how those algorithms operate. Each of these five methods was applied to Data Set 2 (see Table 7A.3), modified as described in Section 7.4.2.2, and yielded source compositions that closely matched the compositions of the Aroclor sources in Table 7A.2. We will focus on these five methods; however, it should be noted that there are yet other methods with similar objectives that have been described in the literature (Ozeki *et al.*, 1995; Rachdawong and Christensen, 1997; Xie *et al.*, 1998; Imamoglu *et al.*, 2002, 2004).

7.3.1 POLYTOPIC VECTOR ANALYSIS (PVA)

PVA was developed for analysis of mixtures in the geological sciences. It has evolved over a period of 40 years, with different aspects of the algorithm presented in a series of publications, by a number of different authors. The roots of PVA are in principal components analysis, pattern recognition, linear algebra, and mathematical geology. Its development in mathematical geology can be traced back to the early 1960s and John Imbrie (Imbrie, 1963). Following Imbrie's work, a series of FORTRAN-IV programs were published (Manson and Imbrie, 1964; Klovan, 1968; Klovan and Imbrie, 1971). The resulting software developed by Imbrie (at Brown University) and Ed Klovan (at the University of Calgary) was called CABFAC (Calgary and Brown FACtor Analysis) and quickly became the most commonly used multivariate analysis algorithm in the geosciences. Subsequent investigators that proved crucial in development of the PVA algorithm included A.T. Miesch and William Full. Miesch, a geochemist with the U.S. Geological survey in the 1970s, was one of the first to take full advantage of Imbrie's oblique vector rotation methods (Miesch, 1976a,b). Miesch also developed the variable-by-variable goodness-of-fit criteria (Miesch CDs) discussed in Section 7.2.4.2. William Full developed the DENEG algorithm, which allows end-members (sources) to be resolved without *a priori* knowledge of their composition, and without use of a training data set (Full *et al.*, 1981, 1982).

The name polytopic vector analysis follows directly from the jargon of Imbrie (1963) and Full *et al.* (1981, 1982). Imbrie (and others) referred to eigenvector decomposition models, resolved in terms of orthogonal axes as *factor analysis*. Solutions resolved in terms of oblique vectors he termed *vector analysis*. Although Imbrie's is not the definition of true factor analysis as defined by Harman (1960), his terminology has held within mathematical geology and a number of other subdisciplines. PVA involves resolution of oblique vectors as source compositions, thus the term vector analysis was used

rather than factor analysis. The term *polytopic* derives from the concept that PVA involves resolution of a $k - 1$ dimensional solid, a simplex or polytope, within k dimensional principal component space. As illustrated in Figure 7.6, the polytope or simplex in three-dimensional space is a two-dimensional triangle. In two space it is a straight line. In four space it is a tetrahedron. Because the objective of PVA is to resolve a $k - 1$ dimensional simplex, the constant vector length transform typically is employed because it forces all sample vectors to have unit length, and thus all data are constrained to a $k - 1$ dimensional space within k space (see Figure 7.4).

This section presents the full PVA algorithm running under default conditions; that is, (1) the EXTENDED CABFAC algorithm (Klovan and Imbrie, 1971) and (2) the iterative oblique vector rotation algorithm originally presented as EXTENDED QMODEL (Full *et al.*, 1981, 1982). Because PVA evolved within the mathematical geology literature, the discussions that follow will use the terms *scores* and *loadings* as per the mathematical geology convention (see Section 7.2.3). Although a number of scaling, data transformation, and calculation options may be implemented in PVA, the following algorithms represent the core of PVA as it is presently implemented under default options by the commercial version of the PVA software package (PVA: Tramontane Inc., Salt Lake City, Utah). The software is composed of two modules: VSPACE and PVA. The VSPACE module performs the PCA/dimensionality analysis of the Klovan and Miesch's EXTENDED CABFAC, along with the CD scatter plot goodness-of-fit diagnostics of Johnson (1997; Johnson *et al.*, 2000). The PVA module is essentially the EXTENDED QMODEL and FUZZY QMODEL algorithms of Full *et al.* (1981, 1982). Readers not interested in the formalism of the PVA algorithm can skip ahead to Section 7.3.1.4.

7.3.1.1 Scaling Functions: Back-Calculation to Original Metric

The transformations presented in Section 7.2.2 serve to optimize the eigenvector decomposition, but interpretation and evaluation of matrices $\mathbf{A}''$ and $\mathbf{F}''$ in scientific context is difficult. Calculations are best performed in transform metric, but evaluation/interpretation must be done in measurement metric. Thus the results must be back-calculated. Where double prime (e.g., $\mathbf{X}''$) indicates a matrix that has been transformed in turn by the range transform and the constant vector length transform, the mapping functions presented by Miesch (1976a) allow us to translate the equations $\hat{\mathbf{X}}'' = \mathbf{A}''\,\mathbf{F}''$ back to $\hat{\mathbf{X}} = \mathbf{AF}$ (percent or constant row-sum metric). The mathematics are discussed as follows. As scaling is called upon in various portions of PVA, refer back to this section for a description. The first step in back-calculation to measurement metric is the definition of what Miesch termed scale factors: s_k. Given k

retained eigenvectors, Miesch defines a $1 \times k$ row vector of scale factors $s = \{s_k\}$ where each element s_k is defined as:

$$s_k = \frac{\mathrm{K} - \sum_{j=1}^{n} x_{\min j}}{\sum_{j=1}^{n} \left(f''_{kj} \left(x_{\max j} - x_{\min j} \right) \right)} \tag{7.13}$$

K is a constant (the sum of each row, usually 100), f''_{kj} is the element of the scores matrix (mathematical geology jargon), and $x_{\max j}$ and $x_{\min j}$ are the maximum and minimum values in the jth column of the original data matrix **X**.

The elements of the back-calculated scores matrices **F′** and **F** are then calculated, in turn, as follows:

$$f'_{kj} = s_k f''_{kj} \tag{7.14}$$

$$f_{kj} = f'_{kj} \left(x_{\max j} - x_{\min j} \right) + x_{\min j} \tag{7.15}$$

Similarly, the elements of the back-calculated loadings matrices **A′** and **A** (mathematical geology jargon) are calculated as:

$$a'_{ik} = \frac{a''_{ik}}{s_k} \tag{7.16}$$

If **r** is then defined as a column vector of the m row-sums of **A′**, the elements of **A** are:

$$a_{ik} = \frac{a'_{ik}}{r_i} \tag{7.17}$$

7.3.1.2 Eigenvector Decomposition and Determining the Number of End-Members

A singular value decomposition is carried out on transformed matrix **X′** or **X″** as presented in Equation 7.8. The results are translated into scores and loadings (**A″** and **F″**) again, as per Equations 7.9 and 7.10. The scores and loadings matrices are then translated back to constant sum metric (**A** and **F**) using the scaling functions presented in the preceding section (Equations 7.13 through 7.17). The task then is determination of k, the appropriate number of sources. As discussed in Section 7.2.4, this reduces the problem of choosing the number of significant principal components, and all discussions

presented in that section are equally relevant here. The methods used most often in PVA are (1) the normalized loadings criteria (Ehrlich and Full, 1987—Section 7.2.4.1) and (2) Miesch CDs and scatter plots as described in Section 7.2.4.2. For each number of potential principal components ($k = 1, 2, \ldots n$, if n is smaller; or $k = 1, 2, \ldots m$, if m is smaller), $\hat{\mathbf{X}}$ is calculated as:

Matrix dimensions
$$\underset{(m\times n)}{\hat{\mathbf{X}}} = \underset{(m\times k)}{\mathbf{A}}\,\underset{(k\times n)}{\mathbf{F}} \tag{7.18}$$

The Miesch CDs are calculated and scatter plots constructed using the reduced dimensional scores and loadings (mathematical geology jargon), as expressed in percentage metric.

7.3.1.3 Determining End-Member Compositions and Mixing Proportions

Following determination of k (the number of end-members) a set of mathematical procedures are used to resolve the second and third objectives of the receptor modeling problem: (2) determine the composition of the end-members, and (3) determine the relative proportions of each end-member in each sample. This process is termed *polytope resolution*. The polytope resolution phase typically is conducted within k dimensional varimax space, but can be equivalently performed in unrotated principal component space. The first k columns of $\mathbf{A}''$ are taken, yielding a reduced matrix $\mathbf{A}''$ of size $m \times k$. Because $\hat{\mathbf{X}} = \mathbf{A}''\,\mathbf{F}''$, we may determine matrix $\mathbf{F}''$ by the following matrix regression equation:

$$\mathbf{F}'' = (\mathbf{A}''^{T}\mathbf{A}'')^{-1}\mathbf{A}''^{T}\mathbf{X}'' \tag{7.19}$$

This equation is required because matrix $\mathbf{A}''$ may not be inevitable. Using the scaling equations in Section 7.3.1.1, matrices $\mathbf{A}''$ and $\mathbf{F}''$ are then transformed back to the original constant row-sum metric, and the estimate of $\mathbf{X}$ ($\hat{\mathbf{X}}$) is calculated as:

Matrix dimensions
$$\underset{(m\times n)}{\hat{\mathbf{X}}} = \underset{(m\times k)}{\mathbf{A}}\,\underset{(k\times n)}{\mathbf{F}} \tag{7.20}$$

Selection of Initial Polytope

The first task of the polytope phase is an initial estimate of a polytope. In PVA, this task is analagous to the Initial Estimates task in ALS (see Section 7.3.2.1). In PVA, the most commonly employed method is the extended method, so named because it was used by Full *et al.* (1981) in the EXTENDED QMODEL Fortran IV algorithm. The formalism of the extended method is described later. Alternative methods, their advantages and disadvantages, are subsequently discussed.

The extended method establishes the initial polytope by taking the k most mutually extreme samples in the data set as vertices (the EXRAWC subroutine of Klovan and Miesch, 1976). EXRAWC first picks a good candidate set for these k samples: those with maximum loadings (mathematical geology jargon) on each of the first k factors of $\mathbf{A}''$. The columns of varimax rotated loadings matrix $\mathbf{A}''$ are scanned, and the maximum absolute value loadings in each column are identified. The rows (samples) of $\mathbf{A}''$ corresponding to the maximum loadings are then put into a new $k \times k$ matrix $\mathbf{O}_0$. This operation is done without duplication (i.e., no two rows of $\mathbf{O}_0$ are the same). Clearly, the samples that make up $\mathbf{O}_0$ are candidates for the k most mutually extreme samples. The PVA algorithm uses these k vectors as oblique reference axes for all samples. The resultant oblique loadings and scores matrices $\mathbf{A}_0''$ and $\mathbf{F}_0''$ are calculated as:

$$\mathbf{F}_0'' = \mathbf{O}_0 \mathbf{F}'' \tag{7.21}$$

$$\mathbf{A}_0'' = \mathbf{A}'' \mathbf{O}_0^{-1} \tag{7.22}$$

Matrices $\mathbf{F}_0''$ and $\mathbf{A}_0''$ are then scaled back to measurement space using the scaling functions described in Section 7.3.1.1, yielding $\mathbf{F}_0$ and $\mathbf{A}_0$.

Matrix $\mathbf{A}_0$ is then inspected to determine if the k samples in $\mathbf{O}_0$ are indeed the k most mutually extreme. Following the method of Imbrie (1963) and Miesch (1976a), if the maximum loadings in matrix $\mathbf{A}_0$ equal 1.0 and correspond to the samples taken for matrix $\mathbf{O}_0$, then the k most mutually extreme samples have been taken and matrix $\mathbf{O}_0$ then is used as initial oblique reference axis for iterative model resolution. If loadings greater than 1.0 are identified in $\mathbf{A}_0$, the sample(s) corresponding to that maximum loading replaces the original sample in $\mathbf{O}_0$, and the process is repeated until a suitable set of k samples is obtained. If the algorithm finds no set of samples with all elements less than 1.0, then the original matrix $\mathbf{O}_0$ is taken as the initial set of oblique vectors. The algorithm for determining extreme samples is termed the EXRAWC procedure.

Outlier samples related to analytical problems or human error should be corrected or omitted from the analysis, as discussed in Section 7.2.4. However, if such errors are not corrected, the extended method can perform poorly. Outliers are extreme samples, and the extended method defines the initial polytope using extreme samples as vertices. In such cases, another method based on the fuzzy c means clustering algorithm of Bezdek (Bezdek, 1981; Bezdek *et al.*, 1984) will often produce better results. The mathematics of so-called FUZZY QMODEL are presented by Full *et al.* (1982). Full demonstrated that the k fuzzy cluster centers as defined in k dimensional eigenspace were

always well within the convex hull defined by the sample data cloud, and were minimally affected by the presence of outliers. The vectors that define the k fuzzy cluster centers are taken as the row vectors of $\mathbf{O_0}$, and matrices $\mathbf{A_0}$ and $\mathbf{F_0}$ are defined relative to these vectors, as discussed earlier. The main disadvantage of the fuzzy method was that it is more computationally expensive to run. However, in these days of fast and inexpensive desktop computers, this is much less of a problem than was the case in the 1980s. In the absence of outliers due to error (as we hope would result from diligent outlier identification in the PCA step) no advantage is gained by choosing fuzzy over extended.

A third option is to use the samples with maximum loadings in $\mathbf{A''}$. In instances where the EXRAWC subroutine does not converge, the EXTENDED option will default to the set of samples with maximum loadings.

A final option is external input of end-member compositions. If end-member compositions are known or suspected prior to the analysis, those suspected sample compositions may be plugged into the model as potential end-members. Suppose for instance that Aroclors 1248 and the two 1254 variants *were* suspected as the contributing sources for Data Set 2. If those source compositions were used as external end-members, and the model converged without iteration, the tested end-members would be considered feasible. This method is essentially the same as target testing as described by Hopke (1989) and Malinowski (1991) and is also similar to methods that require use of a training data set (such as chemical mass balance approaches). Although this option is included in PVA software, in practice it is seldom used, because it constitutes a tacit hypothesis test, which is contrary to an exploratory data analysis philosophy. Regardless of the method used, the result is ultimately a set of k vectors as the rows of matrix $\mathbf{O_0}$. Matrices $\mathbf{A_0}$ and $\mathbf{F_0}$ are then defined relative to these vectors.

Testing Matrices $\boldsymbol{A_0}$ and $\boldsymbol{F_0}$ for Negative Values

Once an initial polytope is defined, the algorithm scans matrices $\mathbf{A_0}$ and $\mathbf{F_0}$ for negative values. The DENEG subroutine described next distinguishes between adjustable negative values and nonadjustable negative values, based on three user-defined numerical criteria. The first, t_1, is the mixing proportions cutoff criterion. The default t_1 is −0.05. In other words, the algorithm will allow up to a −5% mixing proportion in any sample. The purpose of t_1 is to allow for some noise in the model.

If $\mathbf{A_0}$ does contain negative matrix elements less than t_1, the matrix is scanned again using a second criterion, t_2, referred to as the DENEG value. The default DENEG value is −0.25, but may be modified as the user sees fit. The DENEG subroutine recognizes adjustable mixing proportions only if they

fall in the range between t_1 and t_2. At first glance, the need for the parameter t_2 is not obvious, but in effect, it serves as our final line of defense against outliers. In early development of the algorithm, Full observed that in the presence of outliers, a model would often converge (i.e., no negatives less than −0.05) with the lone exception of an outlier with an extremely high negative value (less than −25%) in matrix **A**. By using the t_2 criterion, the DENEG subroutine would not iterate in an attempt to fit that single sample into the model.

Finally, the algorithm scans $\mathbf{F_0}$ for negative values using a third criteria, t_3, referred to as the end-member composition cutoff criterion. Again, the default t_3 value is −5%, and serves the purpose of allowing some noise in the model. If there are no adjustable negative elements in matrices $\mathbf{A_0}$ and $\mathbf{F_0}$, the algorithm stops. $\mathbf{A_0}$ is taken as the mixing proportions matrix and $\mathbf{F_0}$ as the end-member compositions matrix.

The DENEG Algorithm

If $\mathbf{A_0}$ and $\mathbf{F_0}$ contain adjustable negatives, the DENEG algorithm is implemented. DENEG is an iterative process that alternates between solving for **A** (given nonnegative constraints on **F**) followed by solving for **F** (given nonnegative constraints on **A**). Thus DENEG is analogous to the least squares step in ALS (see Section 7.3.2.2). This procedure is continued until mixing proportions and end-member chemical compostions have no adjustable negative values, or until user-defined iteration cutoff criteria are met.

Geometrically the DENEG procedure is a process of alternate polytope expansion and rotation. The results of DENEG applied to the synthetic three-source PCB data set are shown graphically in Figure 7.7. Recall that data transformations are performed such that each sample vector has unit length (be that in measurement space or reduced dimensional principal components space). Thus, in $k = 3$ dimensional space, all Data Set 2 data points plot on sphere, unit length from the origin (see Figure 7.7).

DENEG begins by constructing an initial simplex in principal component space (Iteration 0: see Figure 7.7). In this case, the EXTENDED method (see Section 7.3.1.3) was used to define the initial polytope (Iteration 0), so the vertices of the Iteration 0 triangle are located at the three most mutually extreme samples in the data set. Had pure end-members been contained within the data set, the nonnegativity criteria would have been met at Iteration 0, and the algorithm would have converged without iteration. However, negative values are present in $\mathbf{A_0}$ (i.e., samples are located outside the Iteration 0 simplex; see Figure 7.7). Thus DENEG begins a series of iterations that alternately adjusts **A** and **F** until neither matrix contains negative values. The DENEG algorithm converged after one iteration, and the Iteration 1 simplex

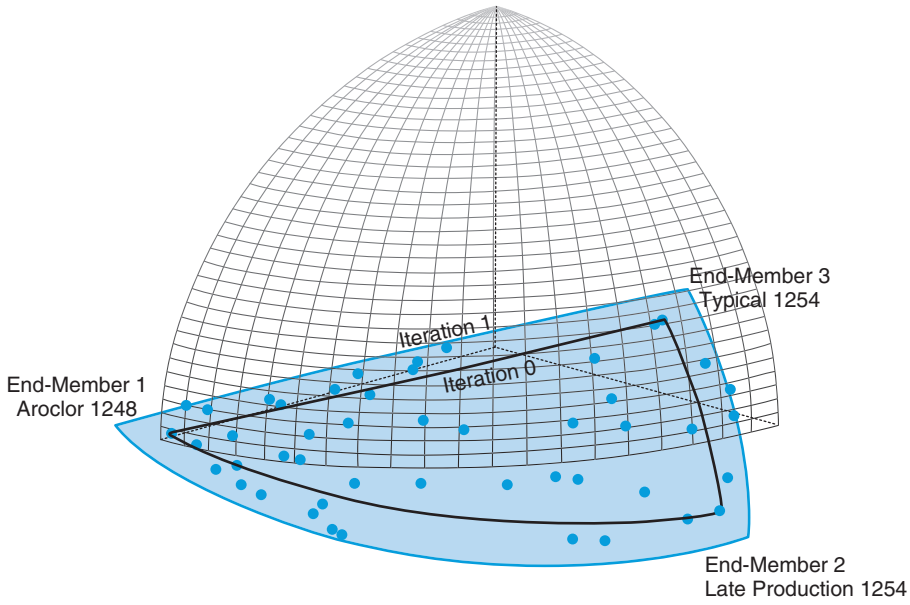


Figure 7.7
Graphical representation of PVA's DENEG algorithm applied to Data Set 2, after errors have been corrected. In three-PC space, the algorithm establishes an initial simplex with vertices at the three most mutually extreme samples in the data set (Iteration 0), then alternately expands and rotates the simplex until nonnegativity constraints are met. The algorithm converged at Iteration 1. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

is shown in Figure 7.7 as the shaded triangle. The associated matrix operations are described as follows.

DENEG begins by defining a $1 \times k$ row vector $\mathbf{D}$ whose elements d_i ($i = 1, 2, \ldots k$) are the lowest adjustable mixing proportions in columns 1, 2, . . . k of matrix $\mathbf{A_0}$. We then define a scalar z:

$$z = \frac{1}{(1 + d_1 + d_2 + \ldots + d_n)} \tag{7.23}$$

A new matrix $\mathbf{A_1} = \{a_{1_{ij}}\}$ may then be calculated as:

$$a_{1_{ij}} = (a_{0_{ij}} + d_j) \times z \tag{7.24}$$

The resultant matrix $\mathbf{A_1}$ represents an end-member mixing proportions matrix, with no samples composed of negative mixing proportions. Geometrically, Equation 7.24 has the effect of moving the edges of the polytope out in an edge parallel direction, and stopping at the outermost sample from that edge: an edge parallel expansion. The corresponding matrix $\mathbf{F_1}$ is then calculated by matrix regression as

$$\mathbf{F_1} = (\mathbf{A_1}^T \mathbf{A_1})^{-1} \mathbf{A_1}^T \hat{\mathbf{X}} \tag{7.25}$$

Matrix $\mathbf{F_1}$ is the end-member compositions matrix. The algorithm scans $\mathbf{F_1}$ for adjustable negatives (less than t_3). If none are encountered, the algorithm

ceases iterating, and $\mathbf{F_1}$ and $\mathbf{A_1}$ are the final mixing proportions and end-member composition and mixing proportions matrices. If adjustable negatives are encountered in $\mathbf{F_1}$, the algorithm continues to the polytope rotation part of the algorithm.

If $\mathbf{F_1}$ contains adjustable negatives, the algorithm continues by substituting all adjustable negative values of $\mathbf{F_1}$ with zeros. A new matrix $\mathbf{F_2}$ is then defined by renormalizing the rows of the nonnegative $\mathbf{F_1}$ to sum to 100%. The algorithm scales $\mathbf{F_2}$ using the transforms described in a previous section. Matrix $\mathbf{F''_2}$ is then taken as a prospective end-member compositions matrix. $\mathbf{F_2}$ is scaled to the transform metric and k dimensional PCA space, and defines a new $k \times k$ oblique matrix $\mathbf{O_2}$. Oblique scores and loadings matrices are then calculated as:

$$\mathbf{F''_2} = \mathbf{O_2}\,\mathbf{F''} \tag{7.26}$$

$$\mathbf{A''_2} = \mathbf{A''}\mathbf{O_2^{-1}} \tag{7.27}$$

Matrix $\mathbf{A''_2}$ is scaled back to measurement space using the scaling functions described in Section 7.3.1.1, yielding $\mathbf{A_2}$. Matrix $\mathbf{A_2}$ then is inspected for adjustable negatives. If none are encountered, the program ceases iterating. If adjustable negatives are encountered in $\mathbf{A_2}$, the iterations continue. The algorithm redefines matrices $\mathbf{A_2}$ and $\mathbf{F_2}$ as $\mathbf{A_0}$ and $\mathbf{F_0}$ and DENEG loops back to the beginning. In the event that iterations do not converge (i.e., they do not result in nonnegative matrices), two additional cut-off criteria are defined. Criterion t_4 is a measure of how similar one iteration is to the next. Criterion t_5 is the user-defined maximum number of iterations. Iterations will continue until no adjustable negatives are encountered, or until one of these two additional cutoff criteria are met.

7.3.1.4 Results of PVA Applied to Data Set 2

PVA was run on Data Set 2, modified as indicated in Section 7.2.4.2. Using Miesch's scaling functions (see Section 7.3.1.1) the vertices of the Iteration 1 triangle (see Figure 7.7) may be back-calculated to percentage metric, to yield estimates of source compositions (matrix **F**). The mixing proportions matrix (**A**) is also determined by back-calculation to the original percentage metric. The resolved end-member compositions are illustrated in Figure 7.8. The three patterns resolved by PVA are clearly quite similar to the true Aroclor compositions. This similarity is quantified using the cosine theta similarity metric for each end-member as compared to the true source profile (Johnson *et al.*, 2006). All three PVA end-members exhibit a cosine theta in excess of 0.99, when compared to the true source profiles used to create Data Set 2.

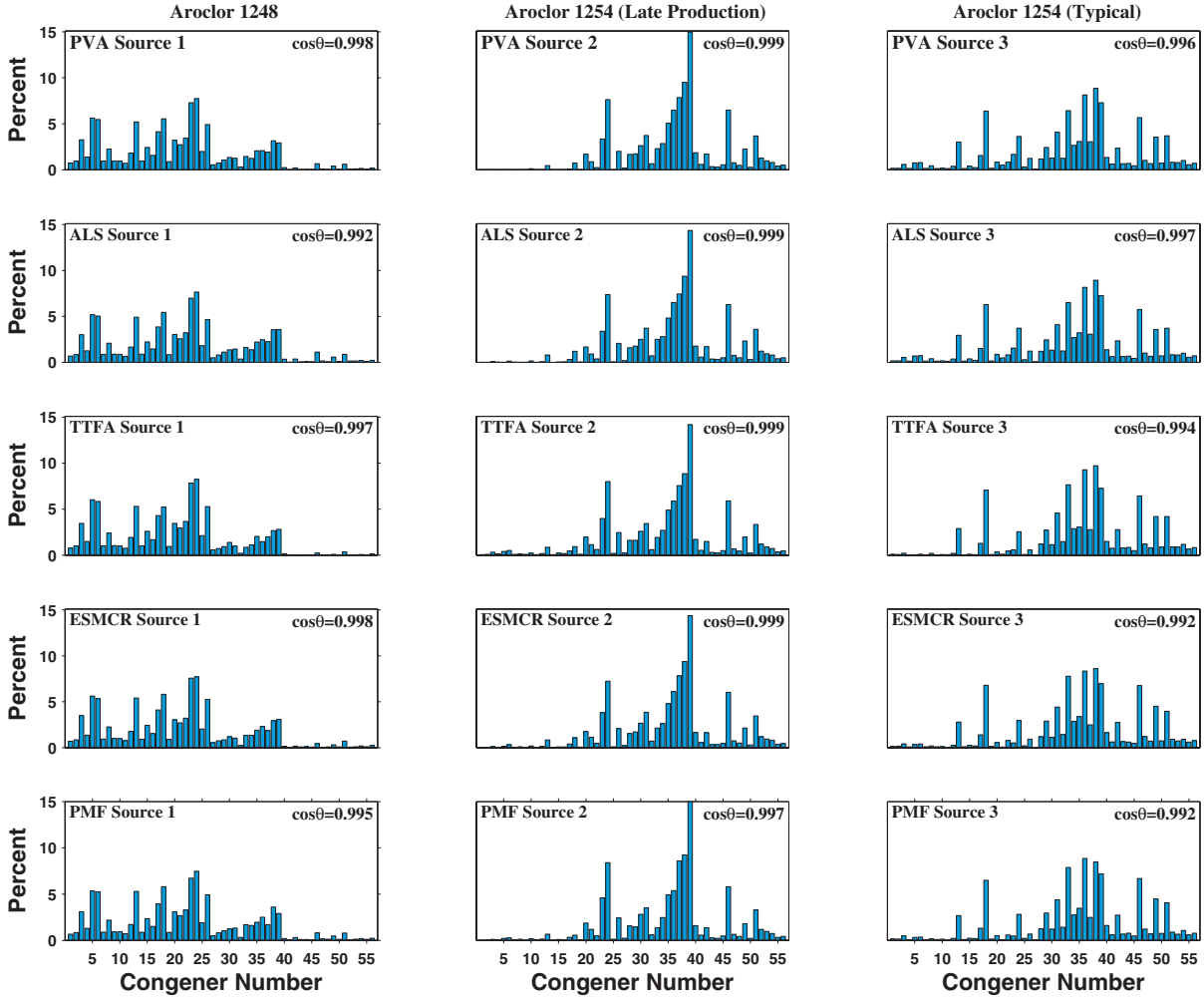


Figure 7.8

Comparison of source compositions derived from PVA (row 1), ALS (row 2), TTFA (row 3), ESMCR (row 4), and PMF (row 5). Congener numbers on the x *axis correspond to the that presented in Table 7A.2. Cos* θ *value at the top right of each bar graph indicate the similarity of each profile to the true source profile, as calculated with the cosine theta similarity metric.*

PVA has been applied in a number of multivariate chemical source apportionment investigations, in sediment, water, and biota studies (Ehrlich *et al.*, 1994; Doré *et al.*, 1996; Jarman *et al.*, 1997; Bright *et al.*, 1999; Johnson *et al.*, 2000; DeCaprio *et al.*, 2005; Magar *et al.*, 2005).

7.3.2 ALTERNATING LEAST SQUARES

Alternating Least Squares (ALS) is another self-training method that allows source profiles and source contributions to be derived in absence of

a priori knowledge of the chemical compositions of the sources (Tauler *et al.*, 1993; Tauler, 1995). Like PVA, ALS seeks to find a solution to Equation 7.1 ($\mathbf{X} = \mathbf{AF} + \varepsilon$). ALS is also similar to PVA in that PCA is used to determine the dimensionality of the data set (i.e., the number of sources), and the reduced principal components space serves as a set reference coordinates for resolution of an oblique vector solution with explicit nonnegative constraints. ALS is implemented in four steps, as outlined next.

7.3.2.1 Initial Estimates

The Initial Estimates step in ALS is analagous to PVA's Selection of Initial Polytope. Several approaches have been described in the context of ALS, based on finding the purest samples in the data matrix. Methods proposed include the needle search method (Gemperline, 1986), evolving factor analysis (Maeder, 1987), or convexity (Grande and Manne, 2000). Alternatively, an initial estimate of the purest variables can be attempted (Windig and Heckler, 1992).

7.3.2.2 Least Squares

Depending on which matrix is estimated initially, the companion matrix can be determined via the matrix pseudo inverse, denoted by the $^+$ symbol. Given an initial or intermediate estimate of **A**, we wish to find **F** that minimizes $\|\mathbf{X} - \mathbf{AF}\|$, subject to the constraints (see Section 7.3.2.3) imposed on **F**.

$$\mathbf{F} = \mathbf{A}^+\mathbf{X} \tag{7.28}$$

Alternatively, given some initial or intermediate estimate of **F**, we wish to find **A** that minimizes $\|\mathbf{X} - \mathbf{AF}\|$, subject to the constraints imposed on **A**.

$$\mathbf{A} = \mathbf{X}\mathbf{F}^+ \tag{7.29}$$

7.3.2.3 Constraints

Estimates of **A** and **F** are refined by applying constraints after the least squares step presented earlier. Different constraints are incorporated depending on the nature of the application; one or all constraints can be applied during each interation. Typical constraints imposed in ALS include the following.

Nonnegativity. As with almost all receptor modeling methods, nonnegativity constraints are imposed on both the source fingerprints and the compositions.

Closure. Mass balance can be invoked such that the sum of the compositions equals a constant, typically 1.0% or 100%.

Equality. If one or more source fingerprints are known *a priori*, this information can be included as a constraint.

Inequality. If a particular source is known never to exceed a certain proportion in any sample, a constraint can be invoked to prevent the source contribution from exceeding that limit.

7.3.2.4 Convergence

Following application of a new least squares estimate with constrained matrices, a better estimate of the **A** and **F** matrices is obtained. These steps are repeated until no further improvement of the matrices is observed or until a maximum number of iterations is reached.

In many applications, applying ALS with these hard physical constraints will provide quite reasonable solutions. ALS has been applied to environmental chemical data in a number of investigations (Tauler *et al.*, 1995; 2000). However, resolved fingerprints may sometimes exhibit noise in the profiles resulting in distortions in the results. Soft constraints have been proposed to minimize these effects by applying penalty functions to constraints (Gemperline and Cash, 2003).

Software implementations of ALS are available in the Matlab environment via the World Wide Web (de Juan; http://www.ub.es/gesq/mcr/mcr.htm) and in commercial chemometrics packages such as Pirouette (Infometrix, Bothell, WA). When ALS was applied to Data Set 2, the resultant source profiles were very similar to the true source compositions (cos θ similarity metrics exceed 0.99) and to those derived by PVA (a method that shares many traits). This consistency of results between ALS and PVA results has also been observed in analysis of GC data from crude oil mixtures (Ramos *et al.*, 2005).

7.3.3 TARGET TRANSFORMATION FACTOR ANALYSIS

Historically, one of the early receptor modeling methods to see application in air receptor modeling was target transformation factor analysis (TTFA), which developed within analytical chemistry/chemometrics rather than mathematical geology/geochemistry (Roscoe and Hopke, 1981; Gemperline, 1984; Hopke, 1989; Malinowski, 1991).

In TTFA, a subroutine that allowed estimates of source composition to be determined in the absence of known sources was the unique vector rotation method (Roscoe and Hopke, 1981; Gemperline, 1984; Hopke, 1989; Malinowski, 1991). This method begins by establishing a $n \times n$ matrix where each row vector is 100% of a single analyte. These are referred to as unique vectors. Each of these vectors is iteratively rotated, in turn, within principal component space. For Data Set 2, Figure 7.9 shows the rotation trajectories for each of

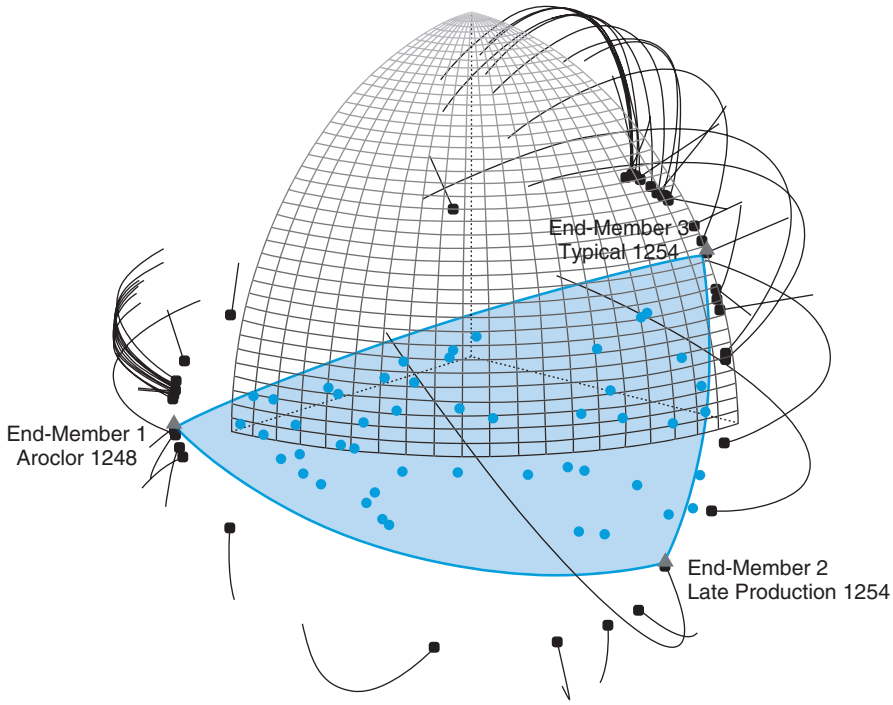


Figure 7.9
Graphical representation of the unique vector iteration method as applied to Data Set 2, after errors have been corrected. This method is used in target transformation factor analysis and involves iterative rotation of a series of n unique vectors. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

the 56 unique vectors. Like DENEG, the method often involves transformation of sample vectors to unit length (Hopke, 1989). As such, Figure 7.9 illustrates the iterative rotations within the constant vector length metric. Each of the iteration trajectories is shown as a curved line on Figure 7.9, which terminates at a location marked (■). The 56 resultant oblique vectors represent a matrix of candidate source profiles for the receptor model. Three of the candidate profiles (▲) were chosen as source compositions in this example, and the shaded triangle shown is the ternary diagram (simplex) constructed based on the chosen profiles.

The unique vector iteration method has been applied in a number of source apportionment investigations, particularly in air receptor modeling (Chang *et al.*, 1988; Hopke, 1989; Borbély-Kiss *et al.*, 1993; Moro *et al.*, 1997). However, the method can be cumbersome, especially in a situation such as Data Set 2, where a large number of variables ($n = 56$) necessitates evaluation of the feasibility of many permutations of candidate source profiles. A TTFA software package FANTASIA (Factor Analysis to Apportion Sources in Aerosols) is available through Dr. Philip Hopke, Clarkson University, Potsdam, New York. However, TTFA is now not as commonly seen in the literature, and the Hopke group's more recent research has utilized positive matrix factorization (see Section 7.3.5).

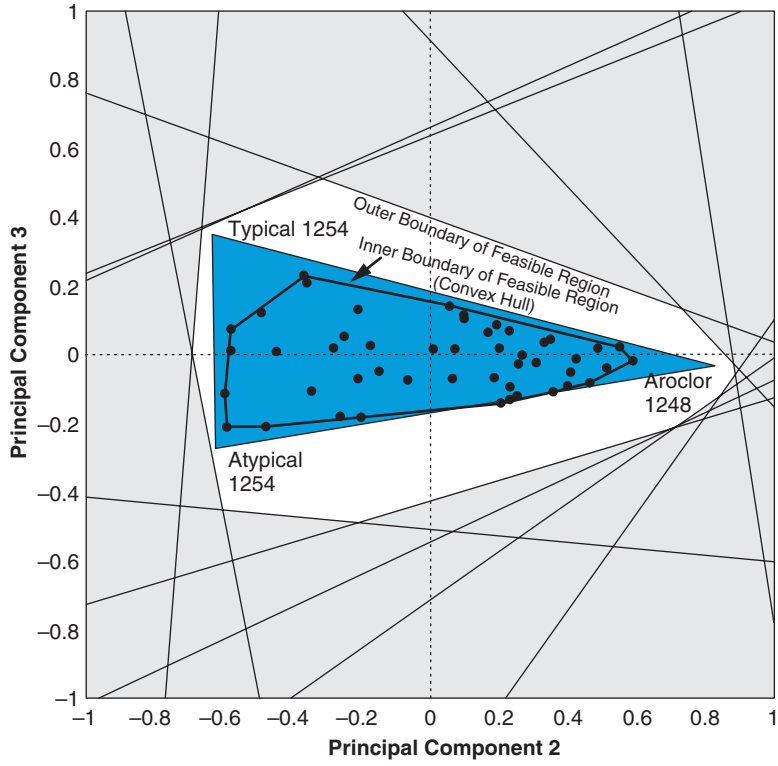


Figure 7.10
Graphical representation of SAFER3D as applied to cleaned up Data Set 2, after errors have been corrected. This method is used in extended self-modeling curve resolution, and resolves source compositions between the inner and outer boundaries of the feasible region. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

7.3.4 EXTENDED SELF-MODELING CURVE RESOLUTION

Another receptor modeling method is extended self-modeling curve resolution (ESMCR: Henry and Kim, 1990; Kim and Henry, 1999; Henry, 2003). Unlike PVA and TTFA, ESMCR typically does not involve transformation to unit vector length. As such, Figures 7.10 and 7.11 show the data projected onto a flat plane perpendicular to principal component 1. Henry and colleagues have presented two primary variations of ESMCR. The first was the SAFER method (Source Apportionment by Factors with Explicit Restrictions). The SAFER method begins by defining the feasible region where the simplex vertices and edges may reside. The inner boundary of the feasible region is defined by the convex hull of the data cloud (see Figure 7.10). The nonnegativity constraints on the analytes define the outer boundary of the feasible region. Each of the lines through the gray shaded region in Figure 7.10 is a zero line for a particular PCB congener, projected into three-PC space. Any potential source compositions that plots in the gray shaded region will have a negative composition for at least one analyte. For a three-component system such as this, a feasible mixing model may be defined by direct inspection of the data plotted in principal component space, and manually located by point-

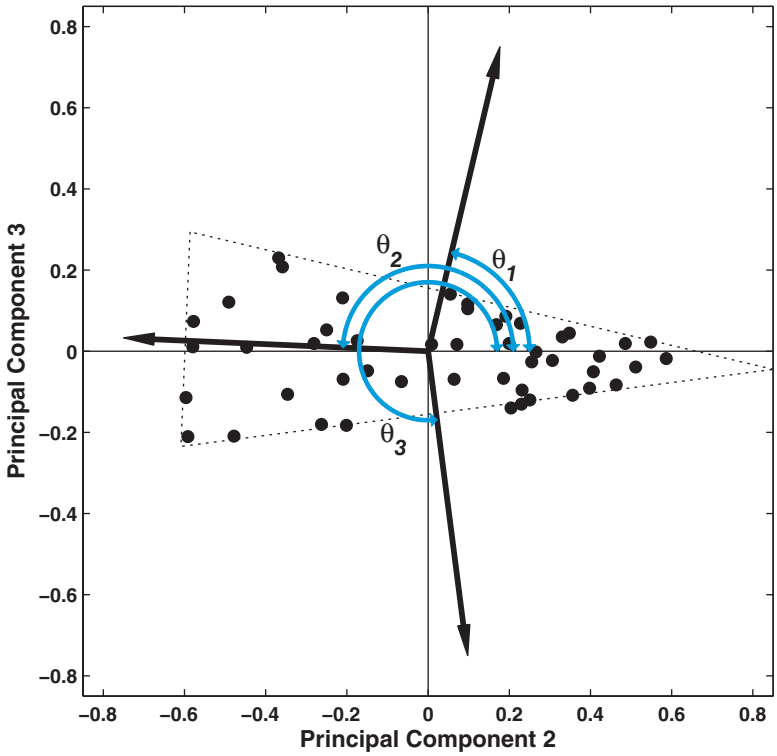


Figure 7.11
Graphical representation of the UNMIX edge-detection algorithm, as applied to cleaned up Data Set 2, after errors have been corrected. As vectors are rotated through PCA space at various angles, a function called the figure of merit is defined, which will exhibit local maxima in directions perpendicular to the edges. The three vectors shown here are normal to the edges of the Data Set 2 data cloud, as determined by the figure of merit.

ing and clicking within the feasible region (this method is termed SAFER3D). An extension of the SAFER method allowed higher dimensional models to be developed through use of additional linear constraints (Kim and Henry, 1999). Applications of SAFER were primarily within the realm of air contaminant source apportionment (Henry *et al.*, 1994; Henry *et al.*, 1997).

In a subsequent publication, Henry (2003) presented an alternative method of resolving higher dimensional mixing models: a method termed N-dimensional edge detection. Henry recognized that finding the edges, faces, or hyperplanes of a mixing diagram (simplex) simultaneously allows identification of simplex vertices. In other words, if you find all edges of a ternary diagram, then by definition, you also have found the corners of the ternary diagram (aka end-members). Henry also recognized that if samples in a data set have zero or very low mixing proportions of one source, that those samples will define a relatively flat edge of the data cloud in PCA space. That flat edge is parallel to (if not equivalent to) the true edge of the simplex. Henry's edge detection algorithm works by defining a series of vectors through the origin of reduced dimensional principal components space (see Figure 7.11). As these vectors are rotated through PCA space at various angles, a function called the figure of merit is defined, which will exhibit local maxima

in directions perpendicular to the edges (Henry, 2003). Figure 7.11 shows three vectors normal to the edges of the Data Set 2 data cloud (as would be determined by the figure of merit). This figure illustrates the three source case, but the edge detection method has been generalized to work in higher dimensions. Once the directions of the normals to the edges has been found, the final location of the edge is done by a statistical fit to the data points on the edge.

Applications of Henry's N-dimensional edge detection method primarily have been within the realm of air contaminant source apportionment (Lewis *et al.*, 2003; Mukerjee *et al.*, 2004). Software to implement this method is available as a package called EPA UNMIX for Windows, and is available through the U.S. Environmental Protection Agency (EPA) on their air quality/receptor modeling Web page (http://www.epa.gov/scram001/receptorindex.htm). The Data Set 2 source compositions resolved by UNMIX are shown in Figure 7.8, and again, source profiles obtained through Henry's ESMCR are in good agreement with the true source profiles.

7.3.5 POSITIVE MATRIX FACTORIZATION

Positive Matrix Factorization (PMF) was developed by Paatero and colleagues (Paatero and Tapper, 1994). As with the other methods discussed here, PMF involves solution of Equation 7.1 through an iterative method of rotation, such that matrices **A** and **F** contain no negative values (once again, explicit nonnegativity constraints). However, in addition to these nonnegativity constraints, PMF iterations also involve minimization of the Frobenius norm of the residuals of the reduced dimensional model (ε in Equation 7.1) divided element by element, by the error estimates of **X**. Error estimates of **X** are expressed in a matrix of standard deviations. Given the matrix of residuals, divide by standard deviation error estimates (σ_{ij}^2),

$$B_{ij}^2 = \frac{\varepsilon_{ij}^2}{\sigma_{ij}^2} \tag{7.30}$$

The Frobenius norm (a scalar q) is calculated as

$$q = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} B_{ij}^2} \tag{7.31}$$

As is evident in these two equations, the scalar q is a function of the residuals of the model (ε), scaled by the expected error in the original measurement of **X**. A key distinction between PMF and other methods discussed herein is

that the PMF iteration process combines both nonnegativity constraints on **A** and **F** (source contributions and source compositions, respectively) with minimization of the scalar *q*.

As with the other methods described in this chapter, PMF was applied to Data Set 2, and the resultant source composition estimates are shown graphically in Figure 7.8. Again, the source patterns resolved are quite similar to the true Aroclor compositions used to construct this data set. PMF has been applied primarily to receptor modeling studies in air (Juntto and Paatero, 1994; Xie *et al.*, 1998; Chueinta *et al.*, 2000; Ramadan *et al.*, 2000; Begum *et al.*, 2003), but recently has seen application in sediment studies as well (Bzdusek *et al.* 2006a,b). The PMF software as discussed here (as well as extended PMF software for three-way analysis) is available through Dr. Pentii Paatero, University of Helsinki, Finland. A version of the PMF software is also available through EPA via their air quality/receptor modeling Web page (http://www.epa.gov/scram001/receptorindex.htm).

7.4 THE INFLUENCE OF ALTERATION PROCESSES ON MIXING MODELS

Self-training receptor models are valuable tools in the analysis of large data sets from field areas where sources cannot be predicted ahead of time. However, environmental processes have the potential to confound receptor model results. For example, PCB patterns commonly are altered in the environment by processes such as microbial dechlorination and volatilization (Bedard and Quensen, 1995; Chiarenzelli *et al.*, 1997). The impact of weathering and degradation on source patterns is always a concern in environmental forensics, but it often is not addressed in the context of multivariate chemometric methods. Johnson and Quensen (2000) used PVA to analyze data from laboratory microbial dechlorination studies. The data used in that study had been reported by Quensen *et al.* (1990), an investigation that involved bench-scale dechlorination experiments using Aroclors in sediment in the presence of anaerobic microorganisms from Hudson River, New York and Silver Lake, Massachusetts. Johnson & Quensen (2000) analyzed data from one of those experiments (Aroclor 1248). The experimental data represented a single source and a single degradation process. If the single source-degradation system behaved as linearly mixing end-members, one would expect resolution of a two end-member system, where one end-member would be unaltered Aroclor 1248; the other would be the extreme end product of the dechlorination process. All intermediate dechlorination products would plot along a continuum between those two extremes. This is not what was observed.

Goodness-of-fit diagnostics (CD scatter plots) indicated the presence of three (not two) principal components, and a three end-member model was resolved. One of the three patterns matched unaltered Aroclor 1248. The second matched the expected extreme dechlorination product. The third was an intermediate dechlorination product. The resolution of a third end-member suggests that dechlorination does not manifest itself mathematically in the same way that a linearly mixing source does. However, given these results and no prior knowledge of the system under study, the fingerprint compositions were readily interpretable in terms of source and alteration processes (assuming of course that the analyst was aware of the various dechlorination processes described in the literature).

This has important implications for those doing chemical pattern recognition in environmental forensics. Given a situation of multiple sources, and patterns modified by one or more alteration processes, source apportionment is more difficult. However, even given this more complex scenario, receptor models are still valuable tools because, with minimal *a priori* assumptions, they allow identification of reasonable estimates of both source profiles and alteration patterns. This has been shown in several field applications where receptor modeling methods have indeed been effective in identifying both source and dechlorination patterns in the field (Imamoglu *et al.*, 2002, 2004; Barabas *et al.*, 2004; Magar *et al.*, 2005; Bzdusek *et al.*, 2006a,b).

7.5 SUMMARY

Environmental forensics, by its very nature, involves analysis of complicated chemical data sets. These data sets typically contain many samples and with many measured chemicals. Thus, by definition, we are working in a multivariate, multidimensional world, and we must bring multivariate statistical methods to bear on these problems. However, the nature of environmental forensics investigations is such that we usually do not fully understand the systems under study. It may well be that classical statistical methods related to formal hypothesis testing are not the wisest choice. Rather, it is often more prudent to employ multivariate exploratory data analysis methods. These methods must have special features: (1) they must be able to handle mixtures; (2) the results must be interpretable in a scientific context; (3) they should minimize *a priori* assumptions of data distributions and source chemistry; and (4) they must be able to handle systems more complex than two or three sources. These conditions greatly limit our choices, but the methods presented in this chapter satisfy these criteria.

All the methods discussed can produce spurious results when faced with bad data. A single highly aberrant measurement can disrupt the variance

structure of the data. This is the Achilles heel of multivariate procedures that depend on variance. Thus one of the most crucial steps in the data analysis process is vigilant outlier detection and data cleaning. PCA and receptor modeling must include systematic and objective procedures for evaluation of the quality of the data, and this is often best accomplished through the use of goodness-of-fit scatter plots. Once the data are cleaned, determination of the number of sources is relatively straightforward. If this step is effectively done, then the various methods used to determine the number of sources are usually consistent.

A hierarchy of procedures can be used to analyze the cleaned data. PCA, the earliest of the procedures discussed here, works best in simple cases, where there are few sources contributing to the system, and there is limited mixing between sources. If an initial PCA indicates the presence of mixtures, it is usually best to move to a method capable of resolving the nature of that mixture. Five methods have been presented here: PVA, ALS, TTFA, SAFER, and PMF. Each of these methods is effective. TTFA's unique vector iteration method is the least attractive because it requires the user to evaluate a large number of candidate source profiles for feasibility.

A common question asked of practitioners of these methods is: how many samples are required for these methods to be effective? Alternatively, one may come across the overly broad statement that receptor modeling is effective only when one has copious data. The accurate, if less-than-satisfying response is that it is sampling-plan dependent. Given a three source system and a very small data set (e.g., ≤10 samples) with one or several samples close to the composition of pure sources, any of the methods discussed herein should work well. Further, if those 10 samples do indeed span the range of chemical variability observed in the field, augmenting the data set with some arbitrary number of samples, all of which are intermediate within the range of variability of the first 10, there will little if any improvement in the accuracy of estimation of the fingeprints. Alternatively, given a three source system and a 1000 sample data set where all samples have very similar relative contributions of each of the sources (i.e., there is little variability in the data set) all these methods will struggle. The key is not some magic number of samples, but rather development and implementation of a sampling plan that reasonably spans the area under study (both temporally and geographically). A rigid statement that many samples are required is only true to the extent that more samples often (but not always) means that the modeled data set will have a better chance of spanning the range of variability observed in the field.

Finally, in our eagerness to apply such methods to environmental chemical data, and our striving to develop more quantitative and rigorous methods, we must not lose sight of the chemical and scientific context of our project. Every

time we find a surprise in a data set (which will be quite often, if we are doing our job correctly) we must relate that information back to the full context of study. For example, if an outlier is indicated on a scatter plot, we must go back to the chemist, or to the data entry technician, or to the field-team leader, and ask what might have caused that sample to be unique. The appropriate data analysis decision (e.g., deletion, modification, collection of new samples) will vary, depending on what we learn in those discussions. The data analysis may bring it to our attention, but questions of cause, source, and scientific significance can rarely be answered in the context of numerical data analysis. More often, we are better served to consider such questions in the context of industrial history, chemistry, geology, and biology. We may borrow methods from mathematics and statistics, but we must remain principally environmental scientists.

ACKNOWLEDGMENTS

Thank you to Ron Henry for providing the software to run SAFER and UNMIX and for reviewing portions of this manuscript. Pentii Paatero provided software to run PMF. Chris Evans, Barry Lavine, Jason Ritter, and Noemi Barabas provided critical review and comment on early iterations of the PVA formalism section. We also express our gratitude to Robert Wagner and Jeff Chiarenzelli for discussions regarding plausible PCB coelution and data error scenarios. Support for portions of this work was provided by GeoChem Metrix, Inc.

REFERENCES

Barabas, N., Goovaerts, P., and Adriaens, P. (2004) Modified polytopic vector analysis to identify and quantify a dioxin dechlorination signature in sediments. 2. Application to the Passaic River. *Environmental Science and Technology* 38, 1821–1827.

Bedard, D.L. and Quensen, J.F. (1995) Microbial reductive dechlorination of polychlorinated biphenyls. In *Microbial Transformation and Degradation of Toxic Organic Chemicals*, L.Y. Young and C.E. Cerniglia, eds. Wiley-Liss, New York, 127–216.

Begum, B.A., Hopke, P.K., and Zhao, W. (2005) Source identification of fine particles in Washington, DC, by expanded factor analysis modeling. *Environmental Science and Technology* 39, 1129–1137.

Bezdek, J.C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

Bezdek, J.C., Ehrlich, R., and Full, W.E. (1984) FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences* 10, 191–203.

Borbély-Kiss, I., Koltay, E., and Szabo, G. (1993) Apportionment of atmospheric aerosols collected over Hungary to sources by target transformation factor-analysis. *Nuclear Instruments and Methods in Physics Research* B75, 287–291.

Brereton, R.G. (1992) *Multivariate Pattern Recognition in Chemometrics, Illustrated by Case Studies*. Elsevier, Amsterdam, 325.

Bright, D.A., Cretney, W.J., MacDonald, R.W., Ikonomou, M.G., and Grundy, S.L. (1999) Differentiation of polychlorinated dibenzo-*p*-dioxin and dibenzofuran sources in coastal British-Columbia, Canada. *Environmental Toxicology and Chemistry* 18, 1097–1108.

Bzdusek, P.A., Lu, J., and Christensen, E.R. (2006a) PCB congeners and dechlorination in sediments of Sheboygan River, Wisconsin, determined by matrix factorization. *Environmental Science and Technology* 120–129.

Bzdusek, P.A., Christensen, E.R., Lee, C.M., Pakdeesusuk, U., and Freedman, D.L. (2006b) PCB congeners and dechlorination in sediments of Lake Hartwell, South Carolina, determined from cores collected in 1987 and 1998. *Environmental Science and Technology* 2006, 109–119.

Cattell, R.B. (1966) The screen test for the number of factors. *Multivariate Behavioral Research* 1, 245–276.

Chang, S.N., Hopke, P.K, Gordon, G.E., and Rheingrover, S.W. (1988) Target transformation factor analysis of airborne particulate samples selected by wind-trajectory analysis. *Aerosol Science and Technology* 8, 63–80.

Chiarenezelli, J.R., Scrudato, R.J., and Wunderlich, M.L. (1997) Volatile loss of PCB Aroclors from subaqueous sand. *Environmental Science and Technology* 31, 587–602.

Davis, J.C. (1986) *Statistics and Data Analysis in Geology*. John Wiley & Sons, New York.

Deane, J.M. (1992) Data reduction using principal components analysis. In *Multivariate Pattern Recognition in Chemometrics*, R.G. Brereton, ed. Elsevier, New York, 125–177.

DeCaprio, A.P., Johnson, G.W., Tarbell, A.M., Carpenter, D.O., Chiarenzelli, J.R., Morse, G.S. *et al.* (2005) PCB exposure assessment by multivariate statistical analysis of serum congener profiles in an adult Native American population. *Environmental Research* 98, 284–302.

Doré, T.J., Bailey, A.M., McCoy, J.W., and Johnson, G.W. (1996) An examination of organic/carbonate-bound metals in bottom sediments of Bayou Trepagnier, Louisiana. *Transactions of the Gulf Coast Association Geology Society* 46, 109–116.

Draper, N.R. and Smith, H. (1981) *Applied Regression Analysis*. John Wiley & Sons, New York, 709.

Eastman, H.T. and Krzanowski, W.J. (1982) Cross-validatory choice of the number of components from a principal components analysis. *Technometrics* 24, 73–77.

Ehrlich, R. and Full W.E. (1987) Sorting out geology, unmixing mixtures. In *Use and Abuse of Statistical Methods in Earth Sciences*, W. Size, ed. Oxford University Press, 34–46.

Ehrlich, R., Wenning, R.J., Johnson, G.W., Su, S.H., and Paustenbach, D.J. (1994) A mixing model for polychlorinated dibenzo-*p*-dioxins and dibenzofurans in surface sediments from Newark Bay, New Jersey using polytopic vector analysis. *Archives of Environmental Contamination and Toxicology* 27, 486–500.

Exner, O. (1966) Additive physical properties. *Collection of Czechoslovak Chemical Communications* 31, 3222–3253.

Frame, G.M. (1999) Improved procedure for single DB-XLB column GC-MS-SIM quantitation of PCB congener distributions and characterization of two different preparations sold as 'Aroclor 1254'. *Journal of High Resolution Chromatography* 22, 533–540.

Frame, G.M., Cochran, J.W., and Bøwadt, S.S. (1996) Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener-specific analysis. *Journal of High Resolution Chromatography* 19, 657–668.

Full, W.E., Ehrlich, R., and Klovan, J.E. (1981) Extended Q Model—Objective definition of external end members in the analysis of mixtures. *Journal of Mathematical Geology* 13, 331–344.

Full, W.E., Ehrlich, R., and Bezdek, J.C. (1982) Fuzzy Q Model—A new approach for linear unmixing. *Journal of Mathematical Geology* 14, 259–270.

Gemperline, P.J. (1984) *A priori* estimates of the elution profiles of pure components in overlapped liquid chromatography peaks using target transformation factor analysis. *Journal of Chemical Information and Computer Sciences* 24, 206–212.

Gordon, G.F. (1988) Receptor models. *Environmental Science and Technology* 22, 1132–1142.

Grung, B. and Kvalheim, O.M. (1994) Rank determination of spectroscopic profiles by means of cross validation: The effect of replicate measurements on the effective degrees of freedom. *Chemometrics and Intelligent Laboratory Systems* 22, 115–125.

Harman, H.F. (1960) *Modern Factor Analysis*. University of Chicago Press.

Henry, R.C. and Kim, B.M. (1990) Extension of self-modeling curve resolution to mixtures of more than three components. Part 1: Finding the basic feasible region. *Chemometrics and Intelligent Laboratory Systems* 8, 205–216.

Henry, R.C., Lewis, C.W., and Collins, J.F. (1994) Vehicle related hydrocarbon source compositions from ambient data: The GRACE/SAFER method. *Environmental Science and Technology* 28, 823–832.

Henry, R.C., Spiegelman, C.H., Collins, J.F., and Park, J.F. (1997) Reported emissions of organic gases are not consistent with observations. *Proceedings of the National Academy of Sciences*, USA 94, 6596–6599.

Henry, R.C., Park, E.S., and Spiegelman, C.H. (1999) Comparing a new algorithm with the classic methods for estimating the number of factors. *Chemometrics and Intelligent Laboratory Systems* 48, 91–97.

Henry, R.C. (2003) Multivariate receptor modeling by N-dimensional edge detection. *Chemometrics and Intelligent Laboratory Systems* 65, 179–189.

Hopke, P.K. (1989) Target transformation factor analysis. *Chemometrics and Intelligent Laboratory Systems* 6, 7–19.

Hopke, P.K. (1991) An introduction to receptor modeling. *Chemometrics and Intelligent Laboratory Systems* 10, 21–43.

Imamoglu, I., Li, K., and Christensen, E.R. (2002) Modeling polychlorinated biphenyl congener patterns and dechlorination in dated sediments from the Ashtabula River, Ohio, USA. *Environment Toxicology Chemistry* 21, 2283–2291.

Imamoglu, I., Li, K., Christensen, E.R., and McMullin, J.K. (2004) Sources and dechlorination of polychlorinated biphenyl congeners in the sediments of Fox River, Wisconsin. *Environmental Science and Technology* 38, 2574–2583.

Imbrie, J. (1963) *Factor and Vector Analysis Programs for Analyzing Geologic Data*. Office of Naval Research. Tech Report No. 6, 83.

Jarman, W.M., Johnson, G.W., Bacon, C.E., Davis, J., Risebrough, R.W., and Ramer, R. (1997) Levels and patterns of polychlorinated biphenyls in water collected from the San Francisco Bay and Esturary, 1993–1995. *Fresnius' Journal of Analytical Chemistry* 359, 254–260.

Johnson, G.W. (1997) Application of polytopic vector analysis to environmental geochemistry problems. PhD Dissertation. University of South Carolina. Columbia, SC.

Johnson, G.W., Jarman, W.J., Bacon, C.E., Davis, J., Ehrlich, R., and Risebrough, R.W. (2000) Resolving polychlorinated biphenyl source fingerprints in suspended particulate matter of San Francisco Bay. *Environmental Science and Technology* 34, 552–559.

Johnson, G.W., Quensen, J.F., III, Chiarenzelli, J.R., and Hamilton, C. (2006) Polychlorinated Biphenyls. In *Environmental Forensics: Contaminant Specific Guide*, R. Morrison and B. Murphy, eds. Elsevier.

Kaiser, H.F. (1958) The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 187–200.

Kim, B.M. and Henry, R.C. (1999) Extension of self-modeling curve resolution to mixtures of more than 3 components Part 2—finding the complete solution. *Chemometrics and Intelligent Laboratory Systems* 49, 67–77.

Klovan, J.E. (1968) Q-mode factor analysis program in FORTRAN-IV for small computers. Kansas Geological Survey Computer Contribution 20, 39–51.

Klovan, J.E. and Imbrie, J. (1971) An algorithm and fortran-IV program for large-scale Q-mode factor analysis and calculation of factor scores. *Journal of Mathematical Geology* 3, 61–77.

Klovan, J.E. and Miesch, A.T. (1976) EXTENDED CABFAC and QMODEL, computer programs for Q-mode factor analysis of compositional data. *Computers and Geosciences* 1, 161–178.

Lewis, C.W., Norris, G.A., Henry, R.C., and Conner, T.L. (2003) Source apportionment of Phoenix PM2.5 aerosol with the Unmix receptor model. *Journal of Air Waste Management Association* 53, 325–338.

Magar, V.S., Johnson, G.W., Brenner, R., Durell, G., Quensen, J.F., III, Foote, E. *et al.* (2005) Long-term recovery of PCB-contaminated sediments at the Lake Hartwell Superfund Site: PCB Dechlorination I—End-Member Characterization. *Environmental Science and Technology* 39, 3538–3547.

Malinowski, E.R. (1977) Determination of the number of factors and the experimental error in a data matrix. *Analytical Chemistry* 49, 612–617.

Malinowski, E.R. and Howery, D.G. (1980) *Factor Analysis in Chemistry*. John Wiley & Sons, New York, 251.

Malinowski, E.R. (1991) *Factor Analysis in Chemistry*, 2e. John Wiley & Sons, New York, 350.

Manson, V. and Imbrie, J. (1964) FORTRAN program for factor and vector analysis of geological data using an IBM 7090 or 7094 computer system. Kansas Geological Survey Special Distribution Publication 13.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C., deJong, S., Lewi, P.J., and Smeyers-Verbeke, J. (1997) Handbook of Chemometrics and Qualimetrics: Part A Elsevier, Amsterdam.

Miesch, A.T. (1976a) Q-mode factor analysis of geochemical and petrologic data matrices with constant row sums. *Geological Survey Prof. Paper*, 574-g, 1–47.

Miesch, A.T. (1976b) Interactive computer programs for petrologic modeling with extended Q-mode factor analysis. *Computers and Geosciences* 2, 439–492.

Moro, G., Lasagni, M., Rigamonti, N., Cosentino, U., and Pitea, D. (1997) Critical review of the receptor model based on target transformation factor analysis. *Chemosphere* 35, 1847–1865.

Mukerjee, S., Norris, G.A., Smith, L.A., Noble, C.A., Neas, L. M, Özkaynak, H.A., and Gonzales, M. (2004) Receptor model comparisons and wind direction analyses of volatile organic compounds and submicrometer particles in an arid, binational, urban air shed. *Environmental Science and Technology* 38, 2317–2327.

Ozeki, T., Koide, K., and Kimoto, T. (1995) Evaluation of sources of acidity in rainwater using a constrained oblique rotational factor analysis. *Environmental Science and Technology* 29, 1638–1645.

Ramos, S., Rohrback, B., Johnson, G., and Kaufman, R. (2005) Using gas chromatography and curve resolution to quantify contributions to mixed crude oils. Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy. Orlando, FL. Feb. 27–Mar. 4, 2005.

Reyment, R.A. and Jöreskog, K.G. (1993) *Applied Factor Analysis in the Natural Sciences*. Cambridge University Press, Cambridge.

Roscoe, B.A. and Hopke, P.K. (1981) Comparison of weighted and unweighted target transformation rotations in factor analysis. *Computers and Chemistry* 5, 1–7.

Tauler, R., Kowalski, B., and Fleming, S. (1993) Multivariate curve resolution applied to spectral data from multiple runs of an industrial process. *Analytical Chemistry* 65, 2040–2047.

Tauler, R. (1995) Multivariate curve resolution applied to second order data. *Chemom. Intell. Lab. Sys*. 30, 133–146.

Tauler, R., Smilde, A.K., and Kowalski, B.R. (1995) Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution. *Journal of Chemometrics* 9, 31–58.

Tauler, R., Barcelo, D., and Thurman, E.M. (2000) Multivariate correlation between concentrations of selected herbicides and derivatives in outflows from selected U.S. midwestern reservoirs. *Environmental Science and Technology* 34, 3307–3314.

Tanabe, S., Kannan, N., Subramanian, A., Watanabe, S., and Tatsukawa, R. (1987) Highly toxic coplanar PCBs: Occurrence, source, persistence and toxic implication to wildlife and humans. *Environmental Pollution* 47, 147–163.

Tukey, J.W. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

Van den Berg, M., Birnbaum, L., Bosveld, A.T.C., Brunstrom, B., Cook, P., Feeley, M. *et al.* (1998) Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environment Health Perspectives* 106, 775–792.

Windig, W. and Heckler, C.E. (1992) Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the simplisma approach. *Chemometrics Intelligent Laboratory Systems* 14, 195–207.

Wold, S. (1978) Cross-validatory estimation of the number of components in factor and principal components analysis models. *Technometrics* 20, 397–406.

Wold, S., Johansson, E., and Cocchi, M. (1993) PLS: Partial least squares projections to latent structures. In *3D-QSAR in Drug Design: Theory, Methods and Applications*, H. Kubinyi, ed. ESCOM Science, Ledien, 523–550.

Xie, Y.L., Hopke, P.K., and Paatero, P. (1998) Positive matrix factorization applied to a curve resolution problem. *Journal of Chemometrics* 12, 357–364.

Zhou, D., Chang, T., and Davis, J.C. (1983) Dual extraction of *R*-mode and *Q*-mode factor solutions. *Journal of Mathematical Geology* 15, 581–606.

APPENDIX

Table 7A.1

Data Set 1 (24 samples, 56 congeners) created by addition of Gaussian noise to three Aroclor compositions reported by Frame et al. (1996). Units in percent.

Sample	PCB 16	PCB 17	PCB 18	PCB 22	PCB 28	PCB 31	PCB 32	PCB 33	PCB 37	PCB 40	PCB 41	PCB 42	PCB 44
Aroclor 1248	0.76	0.75	2.70	1.16	6.79	4.98	1.05	1.50	1.04	0.74	0.64	1.90	6.34
Aroclor 1248	0.87	1.06	2.61	1.46	5.76	6.11	0.61	2.65	0.89	1.08	0.66	1.80	5.82
Aroclor 1248	0.66	0.99	3.72	1.65	6.46	6.68	0.97	2.64	1.22	0.85	0.63	1.95	5.89
Aroclor 1248	0.66	0.96	3.67	1.72	5.39	5.76	0.93	2.08	0.89	1.02	0.82	1.78	4.77
Aroclor 1248	0.65	1.01	3.95	2.03	6.17	3.78	1.14	2.46	1.05	0.92	0.87	1.85	6.89
Aroclor 1248	0.58	1.18	3.84	1.56	5.05	4.63	0.98	2.26	0.78	1.07	0.88	2.16	5.13
Aroclor 1248	0.70	0.96	3.61	1.30	3.80	5.65	1.07	2.96	1.02	1.07	0.89	1.50	5.93
Aroclor 1248	0.84	1.03	2.72	1.42	4.89	6.62	0.95	2.19	1.20	1.06	0.82	1.99	5.22
Aroclor 1254 (late production)	0.02	0.02	0.09	0.02	0.06	0.11	0.01	0.06	0.01	0.10	0.02	0.08	0.75
Aroclor 1254 (late production)	0.02	0.03	0.09	0.02	0.07	0.14	0.01	0.07	0.01	0.17	0.02	0.06	0.99
Aroclor 1254 (late production)	0.02	0.02	0.08	0.01	0.07	0.11	0.01	0.05	0.01	0.16	0.02	0.10	0.75
Aroclor 1254 (late production)	0.02	0.02	0.09	0.03	0.07	0.16	0.01	0.05	0.01	0.17	0.03	0.12	0.86
Aroclor 1254 (late production)	0.02	0.02	0.07	0.02	0.08	0.10	0.01	0.04	0.01	0.17	0.02	0.08	0.82
Aroclor 1254 (late production)	0.02	0.03	0.08	0.02	0.10	0.14	0.01	0.06	0.01	0.19	0.02	0.10	0.72
Aroclor 1254 (late production)	0.02	0.02	0.04	0.02	0.07	0.14	0.01	0.04	0.01	0.19	0.03	0.14	0.83
Aroclor 1254 (late production)	0.02	0.02	0.09	0.03	0.06	0.15	0.01	0.05	0.01	0.16	0.02	0.06	0.59
Aroclor 1254 (typical)	0.07	0.10	0.24	0.04	0.20	0.32	0.07	0.20	0.07	0.14	0.01	0.14	2.64
Aroclor 1254 (typical)	0.11	0.09	0.30	0.04	0.21	0.30	0.05	0.18	0.07	0.14	0.01	0.14	1.94
Aroclor 1254 (typical)	0.11	0.08	0.21	0.03	0.14	0.22	0.05	0.16	0.08	0.10	0.01	0.16	2.35
Aroclor 1254 (typical)	0.08	0.09	0.27	0.05	0.22	0.31	0.03	0.11	0.08	0.13	0.01	0.24	2.29
Aroclor 1254 (typical)	0.10	0.09	0.35	0.05	0.16	0.33	0.04	0.15	0.08	0.12	0.01	0.12	2.82
Aroclor 1254 (typical)	0.08	0.07	0.23	0.04	0.17	0.30	0.04	0.20	0.08	0.09	0.01	0.13	3.02
Aroclor 1254 (typical)	0.11	0.08	0.33	0.04	0.13	0.24	0.05	0.17	0.07	0.11	0.01	0.14	2.72
Aroclor 1254 (typical)	0.08	0.08	0.31	0.04	0.19	0.27	0.05	0.21	0.08	0.15	0.01	0.18	3.00

Table 7A.1
Continued

PCB 45	PCB 47	PCB 48	PCB 49	PCB 52	PCB 53	PCB 56	PCB 60	PCB 64	PCB 66	PCB 70	PCB 71	PCB 74	PCB 77	PCB 82
0.83	2.83	1.70	4.28	5.45	0.73	2.27	3.76	3.52	6.81	9.70	2.31	5.66	0.70	0.55
1.08	2.70	1.23	4.30	6.28	0.77	2.88	2.49	3.28	9.21	9.01	2.04	4.60	0.49	0.70
1.19	1.87	1.87	4.35	5.01	0.88	3.50	2.41	3.50	7.20	7.40	2.15	4.43	0.57	0.70
1.09	3.14	1.93	4.50	5.08	1.16	3.73	2.71	3.24	6.84	7.48	2.76	5.41	0.66	0.55
1.21	2.80	1.50	4.37	6.43	0.79	4.03	3.16	3.18	7.54	5.85	2.02	4.68	0.67	0.55
0.97	2.80	1.44	4.52	5.77	1.02	2.84	3.33	3.41	8.81	7.60	2.06	4.84	0.61	0.77
0.82	2.43	1.37	4.57	5.54	1.14	3.81	2.72	4.08	7.71	7.67	2.18	5.17	0.46	0.77
1.10	1.77	1.57	4.25	5.71	0.69	4.27	2.80	3.46	8.06	8.19	1.98	4.29	0.51	0.77
0.02	0.06	0.05	0.26	0.69	0.04	1.62	0.95	0.36	3.58	7.92	0.12	2.08	0.23	1.52
0.03	0.10	0.06	0.29	0.83	0.04	1.37	1.06	0.23	4.36	5.63	0.12	2.11	0.23	1.92
0.02	0.06	0.05	0.27	0.78	0.04	1.42	0.62	0.39	3.98	7.45	0.10	2.68	0.23	1.74
0.02	0.09	0.06	0.29	0.93	0.04	2.13	0.96	0.40	3.83	5.65	0.08	2.08	0.24	1.45
0.02	0.07	0.06	0.28	0.96	0.05	1.81	1.18	0.37	4.71	8.06	0.11	2.54	0.21	1.53
0.02	0.07	0.06	0.29	0.85	0.04	2.42	0.96	0.38	3.50	8.58	0.09	2.16	0.26	1.59
0.02	0.08	0.04	0.29	0.83	0.06	1.59	1.12	0.33	3.90	7.38	0.13	1.97	0.23	1.98
0.02	0.08	0.06	0.29	0.74	0.03	1.77	1.06	0.45	3.70	9.32	0.12	2.47	0.16	1.90
0.06	0.16	0.10	1.18	4.12	0.15	0.57	0.19	0.49	1.16	3.69	0.16	0.97	0.04	1.22
0.05	0.17	0.18	1.27	5.64	0.15	0.58	0.17	0.88	1.25	4.16	0.13	0.66	0.04	1.28
0.04	0.15	0.13	1.11	7.11	0.11	0.54	0.18	0.67	0.78	3.69	0.17	0.79	0.02	1.33
0.07	0.15	0.16	1.18	5.13	0.15	0.57	0.20	0.65	0.91	3.60	0.19	0.88	0.03	1.48
0.05	0.12	0.15	1.16	5.71	0.10	0.69	0.23	0.54	1.06	4.37	0.17	0.86	0.03	1.05
0.06	0.12	0.08	1.16	5.70	0.11	0.62	0.24	0.53	1.30	3.15	0.18	1.17	0.03	0.84
0.06	0.15	0.15	1.23	6.77	0.14	0.51	0.20	0.63	0.89	2.96	0.18	0.83	0.03	1.54
0.05	0.16	0.11	1.21	4.39	0.10	0.62	0.16	0.74	1.20	3.38	0.15	0.99	0.03	1.31

Sample	PCB 84	PCB 85	PCB 87	PCB 92	PCB 95	PCB 97	PCB 99	PCB 101	PCB 105	PCB 110	PCB 118	PCB 128	PCB 130
Aroclor 1248	0.97	1.05	0.67	0.28	1.55	1.01	1.71	2.07	1.59	3.00	2.41	0.09	0.01
Aroclor 1248	0.98	1.17	0.95	0.25	1.36	1.12	2.00	2.26	1.44	1.87	1.94	0.08	0.01
Aroclor 1248	0.89	0.89	1.14	0.27	1.72	0.98	1.67	2.34	1.48	2.60	2.50	0.08	0.01
Aroclor 1248	1.00	1.15	1.13	0.21	1.57	1.20	1.96	2.43	1.57	2.65	2.00	0.06	0.01
Aroclor 1248	1.04	1.25	1.34	0.28	1.56	1.00	1.79	1.87	1.35	2.82	2.02	0.09	0.01
Aroclor 1248	1.05	1.10	1.43	0.23	1.38	1.10	1.93	1.44	1.77	2.51	3.01	0.09	0.01
Aroclor 1248	0.97	1.29	0.75	0.29	1.78	1.23	2.28	1.73	1.07	2.32	2.98	0.08	0.01
Aroclor 1248	0.80	1.27	1.48	0.23	1.75	0.94	1.88	1.75	1.36	3.14	2.68	0.08	0.01
Aroclor 1254 (late production)	1.83	2.27	3.60	0.63	1.99	2.78	4.81	6.55	8.02	11.26	14.48	1.39	0.50
Aroclor 1254 (late production)	2.02	2.50	3.03	0.62	2.02	2.39	5.32	6.78	9.52	8.60	13.85	2.09	0.64
Aroclor 1254 (late production)	1.67	2.66	3.82	0.62	1.59	2.89	4.53	6.59	7.18	9.10	17.61	1.94	0.34
Aroclor 1254 (late production)	1.72	2.69	3.92	0.64	2.18	2.78	4.30	7.52	5.89	7.90	14.42	2.56	0.65
Aroclor 1254 (late production)	1.47	2.22	2.85	0.64	1.92	3.80	3.68	4.33	7.12	10.01	14.88	2.14	0.50
Aroclor 1254 (late production)	1.98	2.74	3.90	0.45	1.80	2.38	5.50	6.91	6.24	9.93	10.95	2.09	0.52
Aroclor 1254 (late production)	1.95	2.31	4.00	0.48	2.02	2.44	5.43	4.85	7.26	7.78	16.30	1.99	0.55
Aroclor 1254 (late production)	1.39	3.04	3.03	0.57	1.55	3.43	5.26	5.67	7.24	10.00	11.55	2.09	0.60
Aroclor 1254 (typical)	1.98	1.71	4.30	1.27	7.23	2.34	3.34	4.98	3.28	10.74	9.63	1.70	0.66
Aroclor 1254 (typical)	2.15	1.33	4.78	1.75	5.74	3.00	4.40	7.80	3.21	9.91	8.14	1.59	0.51
Aroclor 1254 (typical)	2.74	1.36	4.17	0.88	7.27	3.16	3.51	7.94	2.59	10.14	10.10	1.22	0.48
Aroclor 1254 (typical)	2.99	1.23	4.74	1.19	6.53	3.10	3.42	7.67	2.70	10.28	7.99	1.67	0.69
Aroclor 1254 (typical)	2.22	1.49	3.93	1.17	6.10	2.45	2.64	9.72	3.65	9.61	10.14	1.46	0.75
Aroclor 1254 (typical)	1.82	1.89	3.96	1.41	7.39	2.75	3.81	7.08	3.54	12.26	7.92	1.48	0.62
Aroclor 1254 (typical)	2.91	1.77	4.63	1.02	6.99	2.03	3.25	8.63	2.65	9.86	8.89	1.98	0.68
Aroclor 1254 (typical)	2.02	1.66	4.23	1.02	6.43	2.48	3.54	7.81	3.62	12.39	6.52	1.95	0.58

Table 7A.1
Continued

PCB 132	PCB 135	PCB 136	PCB 137	PCB 138	PCB 141	PCB 146	PCB 149	PCB 151	PCB 153	PCB 156	PCB 158	PCB 163	PCB 170	PCB 180
0.13	0.05	0.06	0.02	0.42	0.09	0.05	0.33	0.08	0.43	0.05	0.04	0.09	0.06	0.24
0.16	0.04	0.05	0.01	0.46	0.09	0.05	0.39	0.08	0.39	0.04	0.04	0.11	0.07	0.17
0.13	0.05	0.06	0.02	0.34	0.10	0.05	0.40	0.08	0.36	0.05	0.04	0.10	0.10	0.21
0.13	0.05	0.05	0.02	0.55	0.10	0.05	0.36	0.10	0.44	0.05	0.04	0.08	0.09	0.25
0.14	0.04	0.07	0.02	0.44	0.10	0.05	0.27	0.07	0.36	0.04	0.04	0.11	0.10	0.19
0.17	0.04	0.05	0.03	0.38	0.08	0.04	0.26	0.06	0.45	0.02	0.04	0.07	0.06	0.29
0.16	0.04	0.06	0.02	0.40	0.11	0.06	0.34	0.08	0.49	0.04	0.05	0.08	0.10	0.29
0.13	0.05	0.06	0.02	0.46	0.11	0.05	0.30	0.10	0.49	0.03	0.04	0.09	0.09	0.23
1.10	0.31	0.19	0.60	6.32	0.59	0.52	1.93	0.26	3.24	1.30	1.25	0.81	0.33	0.30
1.72	0.35	0.28	0.53	7.08	0.78	0.49	2.31	0.29	2.65	0.93	1.36	0.78	0.46	0.53
1.36	0.28	0.29	0.61	5.92	0.64	0.46	1.90	0.31	3.11	0.84	0.91	0.84	0.31	0.43
2.09	0.39	0.32	0.63	7.98	0.74	0.43	2.40	0.26	4.00	1.20	0.93	0.70	0.53	0.38
1.57	0.20	0.26	0.51	6.74	0.74	0.56	2.04	0.22	3.95	1.04	1.35	0.95	0.38	0.48
1.72	0.25	0.23	0.76	6.08	0.58	0.46	2.21	0.26	4.99	1.35	1.31	0.80	0.39	0.47
2.02	0.34	0.30	0.68	6.48	0.63	0.52	1.83	0.21	4.25	1.10	0.94	0.88	0.37	0.55
2.07	0.24	0.30	0.61	6.35	0.63	0.64	2.07	0.28	4.02	1.19	0.94	0.90	0.46	0.42
2.87	0.59	0.69	0.34	6.85	1.22	0.56	4.79	0.81	5.44	1.00	0.87	1.11	0.55	0.67
3.12	0.64	0.96	0.47	5.83	1.35	0.69	3.55	0.87	3.37	0.90	0.88	1.38	0.65	0.84
1.75	0.66	0.79	0.37	6.29	1.16	0.77	3.46	0.76	3.98	1.06	0.91	0.86	0.42	0.65
2.35	0.69	0.74	0.46	7.57	1.30	0.67	3.65	0.82	4.08	0.90	0.64	1.27	0.63	0.77
2.20	0.63	0.78	0.44	6.10	0.92	0.47	3.24	0.75	4.42	0.62	0.96	1.21	0.40	0.79
2.84	0.65	0.85	0.40	4.56	1.17	0.81	3.11	0.64	4.91	1.08	0.96	1.20	0.45	0.68
2.39	0.61	0.61	0.50	5.50	0.80	0.68	3.26	0.62	4.82	0.82	0.76	1.36	0.62	0.89
2.54	0.61	0.72	0.48	6.33	1.22	0.59	4.43	0.75	4.20	0.91	1.23	1.14	0.66	0.72

Table 7A.2

*Input matrices for artificial three-source PCB mixture. Multiplication by Equation 7.1 (X = A*F) yields error free matrix X.*

Source Contributions Matrix (Mixing Proportions) [A]				End-Member Source Compositions Matrix [F^t] (Shown graphically in Figure 12.1)			
Sample Number	**Source 1 Aroclor 1248 (%)**	**Source 2 Late Production Aroclor 1254 (%)**	**Source 3 Typical Aroclor 1254 (%)**	**IUPAC Congener**	**Source 1 Aroclor 1248**	**Source 2 Late Production Aroclor 1254**	**Source 3 Typical Aroclor 1254**
Sample 1	0.03	0.43	0.54	1 PCB 16	0.75	0.02	0.10
Sample 2	0.22	0.26	0.51	2 PCB 17	0.98	0.02	0.09
Sample 3	0.59	0.13	0.28	3 PCB 18	3.46	0.09	0.27
Sample 4	0.71	0.11	0.18	4 PCB 22	1.45	0.02	0.04
Sample 5	0.03	0.38	0.59	5 PCB 28	5.86	0.06	0.21
Sample 6	0.59	0.24	0.17	6 PCB 31	5.76	0.12	0.30
Sample 7	0.45	0.07	0.48	7 PCB 32	0.98	0.01	0.05
Sample 8	0.41	0.02	0.57	8 PCB 33	2.33	0.05	0.17
Sample 9	0.19	0.37	0.44	9 PCB 37	1.00	0.01	0.08
Sample 10	0.60	0.35	0.05	10 PCB 40	0.97	0.16	0.13
Sample 11	0.20	0.74	0.05	11 PCB 41	0.79	0.02	0.01
Sample 12	0.55	0.12	0.33	12 PCB 42	1.88	0.10	0.16
Sample 13	0.48	0.08	0.44	13 PCB 44	5.36	0.72	2.50
Sample 14	0.17	0.72	0.11	14 PCB 45	0.96	0.02	0.05
Sample 15	0.44	0.13	0.43	15 PCB 47	2.54	0.08	0.15
Sample 16	0.44	0.36	0.20	16 PCB 48	1.62	0.05	0.13
Sample 17	0.48	0.50	0.02	17 PCB 49	4.39	0.28	1.19
Sample 18	0.15	0.57	0.29	18 PCB 52	5.87	0.89	5.81
Sample 19	0.47	0.47	0.07	19 PCB 53	0.93	0.04	0.13
Sample 20	0.14	0.01	0.85	20 PCB 56	3.36	1.83	0.59
Sample 21	0.40	0.08	0.52	21 PCB 60	2.81	1.02	0.19
Sample 22	0.46	0.45	0.10	22 PCB 64	3.50	0.39	0.64
Sample 23	0.50	0.22	0.28	23 PCB 66	7.60	3.84	1.09
Sample 24	0.49	0.45	0.06	24 PCB 70	7.78	7.36	3.77
Sample 25	0.63	0.27	0.10	25 PCB 71	1.96	0.12	0.16
Sample 26	0.03	0.67	0.30	26 PCB 74	4.92	2.36	0.91
Sample 27	0.24	0.26	0.50	27 PCB 77	0.55	0.22	0.03
Sample 28	0.57	0.07	0.36	28 PCB 82	0.65	1.65	1.20
Sample 29	0.24	0.12	0.65	29 PCB 84	0.96	1.70	2.51
Sample 30	0.67	0.06	0.27	30 PCB 85	1.20	2.68	1.38
Sample 31	0.50	0.21	0.29	31 PCB 87	1.17	3.68	4.31
Sample 32	0.07	0.79	0.13	32 PCB 92	0.26	0.61	1.39
Sample 33	0.07	0.46	0.47	33 PCB 95	1.51	1.98	6.75
Sample 34	0.03	0.79	0.18	34 PCB 97	1.02	3.00	2.83

Table 7A.2
Continued

Source Contributions Matrix (Mixing Proportions) [A]				End-Member Source Compositions Matrix [F^t] (Shown graphically in Figure 12.1)			
Sample Number	**Source 1 Aroclor 1248 (%)**	**Source 2 Late Production Aroclor 1254 (%)**	**Source 3 Typical Aroclor 1254 (%)**	**IUPAC Congener**	**Source 1 Aroclor 1248**	**Source 2 Late Production Aroclor 1254**	**Source 3 Typical Aroclor 1254**
Sample 35	0.18	0.38	0.44	35 PCB 99	1.91	4.88	3.26
Sample 36	0.38	0.28	0.34	36 PCB101	1.99	5.92	8.67
Sample 37	0.34	0.30	0.36	37 PCB105	1.53	7.95	3.23
Sample 38	0.51	0.20	0.29	38 PCB110	2.68	9.08	10.04
Sample 39	0.43	0.52	0.06	39 PCB118	2.47	14.65	7.94
Sample 40	0.38	0.00	0.62	40 PCB128	0.08	1.84	1.53
Sample 41	0.07	0.28	0.65	41 PCB130	0.01	0.54	0.65
Sample 42	0.16	0.04	0.80	42 PCB132	0.15	1.62	2.47
Sample 43	0.29	0.50	0.22	43 PCB135	0.04	0.30	0.66
Sample 44	0.21	0.51	0.28	44 PCB136	0.06	0.26	0.76
Sample 45	0.56	0.43	0.01	45 PCB137	0.02	0.56	0.45
Sample 46	0.24	0.42	0.34	46 PCB138	0.43	6.42	6.27
Sample 47	0.47	0.21	0.32	47 PCB141	0.09	0.74	1.06
Sample 48	0.71	0.06	0.23	48 PCB146	0.05	0.49	0.72
Sample 49	0.37	0.39	0.24	49 PCB149	0.35	1.96	3.94
Sample 50	0.68	0.23	0.08	50 PCB151	0.08	0.24	0.75
				51 PCB153	0.45	3.55	4.07
				52 PCB156	0.04	1.22	0.89
				53 PCB158	0.04	0.97	0.88
				54 PCB163	0.08	0.75	1.11
				55 PCB170	0.08	0.38	0.56
				56 PCB180	0.22	0.45	0.72

Table 7A.3

Data Set 2 (50 samples by 56 congeners created by (1) multiplication of matrices A and F (Table 7.2) as per Equation 7.1; (2) transformation to concentration metric (ng/g); (3) addition of 10% random Gaussian noise; (4) censoring of the data based on a sample specific detection limit (censored data qualified as 'U.' Reported measurement is the detection limit); (5) simulation of data transcription errors in Sample 22; and (6) simulation of DDT coelution with PCB 141 in a subset of samples.

Sample	PCB 16	Qualifier	PCB 17	Qualifier	PCB 18	Qualifier	PCB 22	Qualifier	PCB 28	Qualifier	PCB 31	Qualifier	PCB 32	Qualifier	PCB 33	Qualifier
Sample 1	0.29		0.30		0.94		0.24	U	0.99		1.2		0.24		0.52	
Sample 2	0.66	U	0.66	U	1.6		0.66	U	2.8		2.8		0.66	U	1.3	
Sample 3	0.59		0.82		2.2		1.16		4.3		3.9		0.79		1.9	
Sample 4	1.5		1.7		6.9		2.70		12.2		9.6		1.6		4.5	
Sample 5	0.64		0.80		2.9		0.64	U	3.1		3.2		0.72		1.7	
Sample 6	1.2		1.7		5.4		2.2		9.2		9.8		1.6		3.2	
Sample 7	3.8		4.3		14.7		6.4		28.0		26.0		4.3		8.8	
Sample 8	2.7		2.6		10.2		3.2		15.4		17.9		2.9		8.2	
Sample 9	1.3	U	1.3	U	1.3	U	1.3	U	1.3	U	1.3	U	1.3	U	1.3	U
Sample 10	1.8		2.0		7.8		3.1		10.9		13.4		2.6		5.6	
Sample 11	1.4		1.3		5.8		2.4		8.6		7.9		1.6		3.4	
Sample 12	1.5		2.0		7.6		3.3		15.6		13.0		2.5		5.0	
Sample 13	2.4		3.2		11.0		4.4		15.7		18.9		3.2		7.5	
Sample 14	1.1		1.3		4.0		1.7		7.1		8.1		1.2		2.9	
Sample 15	1.5		2.2		5.2		2.8		11.3		12.1		2.1		4.9	
Sample 16	1.9		2.9		8.7		4.2		14.8		16.3		2.3		6.7	
Sample 17	1.1	U	1.1	U	1.4		1.1	U	2.3		2.6		1.1	U	1.1	U
Sample 18	0.60		0.76		3.1		1.1		4.2		4.4		0.65		2.02	
Sample 19	1.7		2.4		8.8		3.8		13.9		14.8		2.3		4.9	
Sample 20	3.5		3.6		12.8		3.9		15.9		15.6		2.9		6.9	
Sample 21	1.8		2.1		8.0		2.6		10.8		11.6		2.2		4.7	
Sample 22	3.0		3.7		17.4		5.8		20.3		20.8		4.2		9.0	
Sample 23	1.5		1.8		7.3		3.4		12.2		9.5		2.0		4.8	
Sample 24	2.0		2.3		8.1		3.4		15.2		14.3		2.7		5.0	
Sample 25	4.3		3.7		20.8		9.2		37.2		37.5		5.5		15.9	
Sample 26	0.50	U	0.50	U	1.6		0.50	U	2.0		2.3		0.50	U	1.2	
Sample 27	1.7		2.2		6.8		2.7		12.1		13.3		2.6		5.2	
Sample 28	1.6		2.0		7.2		2.2		11.2		11.0		2.0		4.2	
Sample 29	1.5		1.9		6.6		2.7		9.7		12.1		2.4		4.4	
Sample 30	1.8		2.3		7.5		3.2		14.8		12.7		2.3		5.5	
Sample 31	0.71		0.99		2.14		0.90		4.3		4.6		0.70		2.0	
Sample 32	1.2		1.4		5.2		2.2		8.0		9.3		1.5		3.9	
Sample 33	0.78	U	0.78	U	0.86		0.78	U	1.2		1.2		0.78	U	0.78	U
Sample 34	0.66	U	0.66	U	2.1		0.66	U	2.4		3.0		0.66	U	1.3	
Sample 35	1.8		1.8		8.0		2.8		11.3		10.7		2.1		5.2	
Sample 36	1.3		1.8		6.2		2.4		11.0		7.6		1.7		4.5	
Sample 37	1.0		1.6		4.0		2.4		6.6		8.6		1.5		2.9	
Sample 38	0.78		1.0		2.8		1.3		5.5		6.0		0.76		2.3	
Sample 39	0.72	U	1.1		3.0		1.4		5.5		4.6		0.92		1.8	
Sample 40	1.9		2.5		9.3		3.9		13.1		13.4		1.9		6.7	
Sample 41	0.46		0.44		1.7		0.52		2.2		3.0		0.41		1.1	
Sample 42	0.36	U	0.36	U	0.83		0.36	U	1.2		0.85		0.36	U	0.6	
Sample 43	0.74		0.83		3.3		1.2		6.0		5.1		0.85		2.6	
Sample 44	1.09		1.2		5.3		1.8		6.8		6.3		1.4		2.3	
Sample 45	2.6		2.2		10.3		3.8		20.2		19.2		3.0		8.1	
Sample 46	2.2		2.5		8.4		3.0		12.0		16.5		2.3		4.6	
Sample 47	2.3		2.4		10.0		4.2		13.3		12.5		2.7		6.9	
Sample 48	7.4		9.8		40.3		14.2		60.7		52.6		10.0		23.9	
Sample 49	2.5		3.5		14.1		5.4		21.0		19.7		4.0		8.3	
Sample 50	3.5		4.2		12.9		4.2		17.9		24.0		4.0		9.3	

Table 7A.3
Continued

PCB 37	Qualifier	PCB 40	Qualifier	PCB 41	Qualifier	PCB 42	Qualifier	PCB 44	Qualifier	PCB 45	Qualifier	PCB 47	Qualifier	PCB 48	Qualifier	PCB 49	Qualifier	PCB 52	Qualifier	PCB 53	Qualifier
0.30		0.63		0.24	U	0.66		6.0		0.24	U	0.69		0.50		3.6		12.4		0.41	
0.66	U	0.66	U	0.66	U	0.92		4.0		0.66	U	1.1		0.73		3.0		7.7		0.66	U
0.83		0.74		0.45		1.7		5.2		0.84		2.0		1.3		3.6		6.5		0.82	
1.9		1.8		1.4		3.3		10.8		1.9		4.6		3.8		9.4		14.6		2.0	
0.71		1.5		0.64	U	1.7		16.8		0.64	U	1.9		1.6		9.0		34.6		1.1	
1.6		1.6		1.2		3.1		8.6		1.6		4.6		3.1		7.8		14.2		1.3	
4.4		5.2		3.3		8.7		35.5		4.2		11.7		5.1		23.1		47.7		3.8	
3.0		3.1		2.2		6.1		27.1		2.6		7.5		4.4		14.2		39.2		2.86	
1.3	U	1.3	U	1.3	U	1.3	U	2.2		1.3	U	1.3	U	1.3	U	1.4		3.9		1.3	U
2.2		1.8		1.6		3.8		14.2		1.8		5.9		3.5		9.3		14.7		2.3	
1.4		2.5		1.2		3.2		12.8		1.4		3.9		2.9		8.3		14.9		1.8	
2.4		2.5		1.8		5.2		14.7		2.4		6.0		4.3		10.7		18.2		2.6	
2.9		3.0		2.6		5.0		24.5		3.0		8.1		3.7		16.6		32.2		2.4	
1.4		2.1		1.1		2.8		10.7		1.1		3.4		2.6		6.9		13.7		1.4	
1.8		1.8		1.5		4.5		13.8		1.7		4.9		3.3		10.7		25.3		1.7	
2.5		2.8		1.9		6.0		19.5		2.9		6.4		3.9		13.6		20.5		2.6	
1.1	U	1.1	U	1.1	U	1.1	U	2.4		1.1	U	1.2		1.1	U	1.9		2.4		1.1	U
0.85		1.1		0.61		1.7		8.8		0.92		2.4		1.5		4.4		15.4		0.84	
2.4		2.1		1.9		3.9		15.0		2.3		6.2		3.3		9.3		15.7		2.3	
3.4		3.8		1.9		6.5		45.1		3.3		8.0		5.8		25.5		96.0		3.7	
2.0		2.3		1.4		3.8		15.8		2.1		4.7		3.5		10.7		28.9		2.3	
3.9		4.4		2.8		7.7		21.6		3.2		10.4		5.3		19.8		30.9		3.2	
1.8		2.4		1.6		4.1		14.9		1.8		5.3		3.2		7.6		16.4		1.9	
2.3		2.4		2.0		4.0		14.0		2.3		5.5		4.2		10.7		18.0		2.2	
6.5		6.8		4.8		12.3		38.2		6.5		13.5		10.3		25.6		34.3		6.1	
0.50	U	1.5		0.50	U	0.98		9.0		0.50	U	1.0		0.92		4.8		16.0		0.65	
2.2		2.7		1.7		4.9		19.8		2.2		6.5		3.7		11.5		35.9		2.5	
2.0		1.6		1.5		4.1		13.2		1.9		5.6		3.1		7.9		18.9		1.9	
1.9		2.2		1.4		4.1		23.3		2.2		4.9		3.6		14.2		36.1		2.1	
2.2		2.6		1.8		5.1		11.4		2.6		4.9		3.7		10.5		18.8		2.4	
0.77		0.86		0.58	U	1.6		5.3		0.7		2.3		1.4		4.5		7.3		0.62	
1.4		4.0		1.3		3.8		19.9		1.2		3.6		2.7		11.1		29.6		1.8	
0.78	U	0.78	U	0.78	U	0.78	U	4.2		0.78	U	0.78	U	0.78	U	2.3		8.9		0.78	U
0.66	U	1.6		0.66	U	1.42		13.0		0.66	U	1.4		1.0		4.7		17.7		0.84	
1.9		2.7		1.6		4.7		20.0		1.7		4.3		3.7		14.4		38.1		2.2	
1.6		1.9		1.4		3.2		14.2		1.9		4.4		3.0		10.7		20.0		1.5	
1.2		1.7		1.0		2.4		14.1		1.3		3.9		3.0		8.4		17.1		1.4	
0.88		0.97		0.65		1.7		5.9		0.95		2.2		1.5		4.6		7.6		0.77	
1.08		1.2		0.72	U	2.0		6.0		0.74		2.4		1.4		3.8		5.6		0.81	
2.20		2.6		1.9		4.8		22.4		2.0		6.3		4.3		14.0		34.8		2.2	
0.46		0.64		0.26	U	1.1		9.3		0.43		1.0		0.79		4.2		17.1		0.61	
0.36	U	0.36	U	0.36	U	0.54		3.1		0.36	U	0.49		0.36	U	1.6		5.8		0.36	U
1.1		1.5		0.73		2.0		6.7		0.99		2.6		1.6		5.1		10.9		0.89	
1.0		1.6		0.87		2.6		11.2		1.1		3.2		2.2		6.1		20.1		1.2	
2.8		3.2		2.59		6.59		20.83		3.47		7.37		4.4		15.7		23.5		3.4	
2.0		2.7		1.77		4.55		20.20		2.04		5.57		4.1		10.9		31.3		2.5	
2.6		3.1		2.01		5.54		20.30		2.20		6.03		4.9		14.8		22.9		2.3	
11.8		11.3		8.45		18.67		65.55		9.52		27.48		17.8		47.1		88.7		10.4	
3.4		4.2		2.92		7.14		26.77		3.10		11.34		6.0		18.5		36.5		4.1	
4.1		4.1		3.25		5.79		25.26		4.25		9.21		6.4		19.2		27.6		3.4	

Sample	PCB 56	Qualifier	PCB 60	Qualifier	PCB 64	Qualifier	PCB 66	Qualifier	PCB 70	Qualifier	PCB 71	Qualifier	PCB 74	Qualifier	PCB 77	Qualifier
Sample 1	4.1		1.9		1.9		8.7		20.0		0.66		5.2		0.52	
Sample 2	2.8		2.0		2.0		6.3		10.5		0.82		3.9		0.66	U
Sample 3	3.1		2.5		2.8		6.5		7.7		1.3		4.9		0.46	
Sample 4	6.4		5.9		7.3		17.7		21.6		3.2		8.9		1.1	
Sample 5	11.1		5.5		4.5		16.0		53.0		2.0		13.7		1.1	
Sample 6	6.3		5.1		7.0		13.4		20.0		3.8		9.6		1.2	
Sample 7	15.4		14.8		15.9		39.5		42.5		9.5		22.1		2.7	
Sample 8	11.9		8.3		11.6		21.5		35.6		7.0		15.1		1.7	
Sample 9	1.5		1.3	U	1.3	U	2.8		4.8		1.3	U	1.7		1.3	U
Sample 10	9.7		6.2		8.7		20.4		27.7		4.9		15.0		1.5	
Sample 11	12.5		9.8		7.3		32.3		49.7		2.8		18.7		2.1	
Sample 12	7.1		5.9		7.3		21.2		24.5		5.1		12.9		1.5	
Sample 13	13.1		10.3		10.0		25.7		36.7		6.1		17.5		1.6	
Sample 14	13.4		7.8		6.4		27.4		46.8		2.6		18.6		1.2	
Sample 15	8.7		7.6		8.7		16.8		25.1		3.9		11.1		1.2	
Sample 16	12.9		9.9		10.6		27.6		42.8		5.1		15.5		1.7	
Sample 17	2.2		1.9		1.6		4.6		5.4		1.1	U	2.8		1.1	U
Sample 18	8.6		5.1		4.0		14.3		35.1		1.5		12.5		0.83	
Sample 19	12.1		9.4		8.8		24.8		36.7		4.9		19.1		1.4	
Sample 20	16.3		8.3		18.5		28.1		66.5		6.9		28.1		1.7	
Sample 21	8.0		6.4		7.8		20.3		24.6		4.5		15.7		1.1	
Sample 22	18.6		13.3		13.3		36.2		60.9		7.4		27.6		1.6	
Sample 23	9.2		6.3		7.2		17.1		25.1		3.9		11.4		1.1	
Sample 24	10.1		7.3		10.5		28.0		38.0		4.7		18.2		1.8	
Sample 25	25.1		20.6		23.2		49.2		67.8		13.6		33.4		4.2	
Sample 26	10.2		5.9		3.6		22.0		49.7		1.2		15.3		1.1	
Sample 27	13.7		8.0		8.6		31.8		48.8		5.0		20.0		2.0	
Sample 28	7.0		5.7		7.5		18.4		19.7		4.6		11.4		1.1	
Sample 29	9.2		5.9		8.8		22.1		37.8		3.7		16.0		1.1	
Sample 30	9.8		7.2		10.0		17.2		22.0		4.3		13.4		1.5	
Sample 31	4.3		2.8		3.1		8.4		11.0		1.8		5.1		0.58	U
Sample 32	26.1		20.0		10.3		62.3		98.3		3.8		33.6		3.4	
Sample 33	3.2		1.9		1.9		6.0		13.5		0.78	U	3.6		0.78	U
Sample 34	13.5		9.0		5.5		35.2		70.0		1.8		19.6		1.7	
Sample 35	15.4		9.4		11.0		33.0		54.2		4.1		20.0		1.8	
Sample 36	9.1		6.6		8.1		18.9		22.9		3.2		11.4		1.2	
Sample 37	6.7		5.7		5.1		17.0		22.2		3.0		9.4		1.0	
Sample 38	4.1		2.9		3.5		9.6		11.4		2.1		5.5		0.62	U
Sample 39	5.2		3.2		3.3		9.3		15.2		2.1		5.9		0.75	
Sample 40	8.7		7.1		10.8		19.2		29.2		4.7		13.3		1.2	
Sample 41	4.4		2.2		2.7		9.2		19.7		0.87		6.0		0.55	
Sample 42	1.1		0.64		1.1		1.8		4.6		0.36		1.7		0.36	U
Sample 43	6.2		4.9		4.8		11.5		22.4		2.0		8.5		0.8	
Sample 44	11.0		7.1		5.8		19.9		30.0		2.5		10.6		1.2	
Sample 45	14.8		10.0		13.0		28.1		46.0		6.8		20.5		2.9	
Sample 46	16.3		8.9		10.7		34.0		55.3		5.1		21.2		2.3	
Sample 47	11.6		8.9		10.1		22.5		31.1		5.5		17.9		1.6	
Sample 48	32.1		28.3		33.6		93.1		96.4		22.7		67.4		5.5	
Sample 49	20.5		15.1		14.4		48.2		59.9		8.0		31.6		2.9	
Sample 50	13.4		13.1		16.5		32.4		40.0		8.3		23.7		2.3	

Table 7A.3
Continued

PCB 82	Qualifier	PCB 84	Qualifier	PCB 85	Qualifier	PCB 87	Qualifier	PCB 92	Qualifier	PCB 95	Qualifier	PCB 97	Qualifier	PCB 99	Qualifier	PCB 101	Qualifier	PCB 105	Qualifier	PCB 110	Qualifier
5.0		6.7		6.1		14.0		3.0		12.2		9.9		14.3		27.3		17.0		33.2	
1.7		3.3		3.1		6.0		1.5		8.9		5.6		6.1		13.3		8.5		16.4	
1.2		2.0		1.8		3.0		0.80		3.6		2.7		3.7		5.8		3.6		6.9	
2.4		3.1		3.7		4.7		1.2		5.5		4.2		6.3		9.0		6.7		11.0	
10.8		19.8		15.7		40.2		9.1		47.2		23.7		37.6		65.7		51.0		88.3	
3.4		3.7		4.7		5.4		1.6		7.1		4.6		8.9		12.0		8.5		15.2	
9.0		16.3		12.4		24.4		7.2		38.8		17.1		24.4		45.9		27.5		48.6	
7.7		11.1		9.4		20.9		5.4		27.6		13.3		18.5		39.2		19.6		44.4	
1.3		1.9		1.6		3.0		1.3	U	3.4		2.2		3.4		6.3		4.0		6.7	
3.6		4.6		6.7		8.5		1.7		6.1		5.8		11.7		16.1		15.3		19.2	
11.6		10.9		12.3		20.7		3.7		16.6		14.1		25.8		39.7		42.6		47.6	
3.5		6.5		5.5		9.2		3.1		13.5		7.3		10.4		20.0		12.5		27.8	
5.3		12.0		8.2		17.2		4.3		23.0		10.3		17.3		31.8		16.7		37.5	
9.8		12.5		14.3		21.9		4.3		14.7		16.7		27.6		36.9		39.0		56.4	
4.6		8.3		6.1		11.5		3.0		14.5		9.8		10.3		17.3		13.8		27.1	
6.6		8.7		12.2		16.3		3.3		12.6		12.1		17.8		29.1		23.7		35.4	
1.1	U	1.1	U	1.6		2.3		1.1	U	1.6		1.7		3.0		3.5		3.1		5.2	
6.04		8.22		10.74		14.49		3.84		14.25		11.2		17.9		29.7		30.5		43.3	
6.6		6.9		9.9		11.4		2.4		11.7		10.5		21.2		19.5		22.5		31.7	
19.3		42.8		23.4		65.3		20.0		106.9		45.5		57.5		115.2		51.1		128.4	
5.2		8.1		5.9		13.1		3.9		19.1		10.6		11.7		25.9		15.2		32.5	
5.5		12.1		14.3		15.8		4.3		23.2		17.7		24.7		33.1		41.5		27.9	
3.8		5.9		5.5		10.9		2.5		9.9		6.3		10.7		18.7		10.8		18.3	
6.2		8.0		9.6		11.7		2.1		11.2		10.4		13.8		19.3		25.4		28.3	
8.1		11.5		16.7		22.0		4.5		19.5		19.1		27.3		35.0		35.9		51.7	
11.3		14.3		12.5		23.5		6.1		20.2		22.1		31.7		45.0		40.1		67.4	
10.6		16.3		13.7		24.7		6.3		31.8		19.2		30.8		50.1		36.3		74.2	
3.2		5.2		4.1		9.2		2.4		10.3		5.8		8.5		13.0		8.1		19.8	
9.3		16.2		11.6		25.1		7.8		30.6		19.6		23.6		51.0		30.0		51.3	
2.9		5.0		3.3		6.6		2.2		10.0		5.0		8.8		13.1		8.9		18.8	
1.3		2.8		2.7		4.7		1.1		5.7		3.4		5.4		7.7		6.0		8.8	
21.7		26.1		41.5		62.2		12.1		39.5		38.7		68.2		91.4		112.6		137.7	
3.5		3.9		4.8		8.3		1.8		10.0		5.5		9.8		15.8		11.0		19.4	
13.9		19.3		23.3		37.3		7.3		23.3		28.3		44.3		53.2		68.4		83.2	
13.5		16.1		14.4		34.6		8.8		34.3		20.2		38.8		68.4		42.8		72.8	
4.0		7.9		7.0		13.2		3.6		12.5		8.1		12.7		21.6		16.2		33.2	
4.9		6.1		5.8		13.0		3.1		15.1		8.2		12.7		21.2		14.4		26.7	
1.9		2.8		2.9		4.8		1.0		5.6		3.4		5.1		7.6		6.4		11.8	
2.3		2.5		4.1		5.3		1.0		3.9		4.8		8.0		9.5		8.7		13.0	
6.3		11.9		8.9		16.5		4.8		24.2		12.6		12.9		31.6		14.2		36.4	
4.2		7.3		6.2		14.1		3.8		19.9		11.8		12.0		28.6		16.4		30.4	
1.4		2.3		1.5		3.4		1.2		6.1		2.3		3.7		7.6		3.2		8.5	
3.9		6.2		6.1		9.9		2.0		9.6		8.7		9.8		19.4		12.3		27.0	
6.9		7.4		11.9		15.7		4.5		18.7		10.4		19.9		26.2		30.6		27.2	
6.0		8.8		10.1		12.2		3.0		10.6		11.8		19.3		18.8		20.6		33.3	
11.3		12.9		16.8		26.5		5.9		32.1		20.3		33.8		50.4		40.8		65.7	
5.6		8.9		9.8		15.5		4.2		16.0		10.2		15.2		29.8		16.4		40.1	
12.4		17.1		15.9		27.2		7.2		37.2		26.0		33.4		50.4		30.3		57.3	
9.8		17.0		18.4		28.5		5.7		28.0		18.5		38.5		59.5		48.6		70.7	
5.6		6.2		8.0		10.8		2.4		12.4		10.6		14.8		20.3		19.4		26.3	

Sample	PCB 118	Qualifier	PCB 128	Qualifier	PCB 130	Qualifier	PCB 132	Qualifier	PCB 135	Qualifier	PCB 136	Qualifier	PCB 137	Qualifier	PCB 138	Qualifier
Sample 1	34.3		4.9		2.1		6.9		1.6		1.5		1.7		24.9	
Sample 2	16.7		2.2		0.93		3.3		0.72		0.81		0.66		6.9	
Sample 3	6.1		0.87		0.44	U	1.3		0.44	U	0.44	U	0.44	U	3.4	
Sample 4	9.6		1.5		0.92	U	1.6		0.92	U	0.92	U	0.92	U	6.0	
Sample 5	90.2		14.5		5.3		17.6		4.6		4.8		3.7		42.5	
Sample 6	18.8		2.1		0.98	U	2.1		0.98	U	0.98	U	0.98	U	8.2	
Sample 7	47.0		7.1		3.4		11.1		3.0		3.9		2.3		34.4	
Sample 8	37.6		5.5		2.3		8.5		2.7		3.3		1.8		26.7	
Sample 9	8.1		1.3	U	1.3	U	1.6		1.3	U	1.3	U	1.3	U	5.0	
Sample 10	27.4		3.0		0.81		2.7		0.63		0.61		0.87		10.0	
Sample 11	78.1		11.2		3.1		11.0		2.0		1.6		3.0		39.1	
Sample 12	26.4		3.1		1.0		4.7		1.2		1.4		0.88		11.8	
Sample 13	29.3		5.2		2.0		6.2		1.9		2.2		1.6		16.8	
Sample 14	78.1		9.7		3.3		8.7		2.1		1.5		2.7		34.4	
Sample 15	20.9		4.4		1.2		6.4		1.3		1.8		1.2		15.4	
Sample 16	44.9		5.6		2.0		5.2		1.5		1.6		1.6		21.7	
Sample 17	5.6		1.1	U	1.1	U	1.1	U	1.1	U	1.1	U	1.1	U	2.6	
Sample 18	55.3		7.2		2.5		6.5		2.0		1.7		2.2		26.5	
Sample 19	40.6		5.6		1.3		5.5		0.96		1.1		1.4		16.2	
Sample 20	126.1		25.1		8.8		39.0		8.1		7.0		7.1		103.7	
Sample 21	27.5		4.4		1.8		8.1		2.0		1.7		1.5		18.7	
Sample 22	9.4		7.7		25.3		7.4		0.29		7		1		2	
Sample 23	26.6		2.8		1.2		4.4		1.1		1.0		1.0		13.1	
Sample 24	41.3		4.7		1.4		4.5		1.0		0.94		1.5		15.6	
Sample 25	60.0		6.0		2.1		8.6		1.7		1.4		2.1		26.3	
Sample 26	94.0		12.0		3.4		11.9		3.1		2.8		3.7		43.4	
Sample 27	77.8		9.8		4.2		16.2		3.6		3.6		2.7		41.4	
Sample 28	16.3		2.1		0.78		3.3		0.8		1.1		0.8		9.4	
Sample 29	53.7		7.3		3.7		15.0		4.7		3.5		2.4		35.1	
Sample 30	17.8		2.0		0.68		2.8		0.75		0.94		0.47		7.8	
Sample 31	11.3		1.5		0.58	U	1.6		0.58	U	0.58	U	0.58	U	5.2	
Sample 32	213.0		23.8		8.6		26.4		4.9		4.9		6.2		87.6	
Sample 33	19.3		3.3		1.4		3.7		1.1		1.1		0.96		11.7	
Sample 34	110.8		13.2		5.4		14.8		3.5		3.5		4.5		59.1	
Sample 35	109.7		14.5		4.5		15.6		4.6		4.1		3.1		57.1	
Sample 36	30.6		5.5		1.5		6.9		1.3		1.6		1.1		19.3	
Sample 37	33.4		5.0		1.5		5.4		1.2		1.5		1.4		14.9	
Sample 38	11.4		1.5		0.62	U	1.86		0.62	U	0.62	U	0.62	U	5.1	
Sample 39	17.8		2.5		0.72	U	2.34		0.72	U	0.72	U	0.72	U	7.4	
Sample 40	34.6		6.5		2.2		9.0		2.5		2.5		1.9		22.2	
Sample 41	36.9		5.7		2.3		8.5		1.9		2.2		1.8		20.1	
Sample 42	8.2		1.3		0.55		1.9		0.58		0.63		0.39		5.55	
Sample 43	32.8		3.8		1.2		4.6		0.92		1.0		1.3		13.2	
Sample 44	52.7		7.5		2.4		9.0		1.7		2.0		2.5		24.2	
Sample 45	38.2		5.3		1.6		5.4		0.84		0.9		1.5		18.3	
Sample 46	74.9		11.0		3.9		13.5		2.9		2.8		3.6		39.4	
Sample 47	40.2		4.9		1.6		6.1		1.6		1.6		1.3		20.5	
Sample 48	67.6		7.1		2.7		10.4		3.1		3.0		2.1		24.7	
Sample 49	94.5		10.8		3.1		12.4		2.9		3.2		3.3		37.2	
Sample 50	31.4		3.6		0.98		4.4		0.88		1.0		1.0		12.8	

Table 7A.3
Continued

PCB 141	Qualifier	PCB 146	Qualifier	PCB 149	Qualifier	PCB 151	Qualifier	PCB 153	Qualifier	PCB 156	Qualifier	PCB 158	Qualifier	PCB 163	Qualifier	PCB 170	Qualifier	PCB 180	Qualifier
10.5		2.0		10.3		1.4		12.1		3.4		3.2		2.7		1.8		2.1	
3.6		0.9		3.8		0.80		5.2		1.2		1.1		1.5		0.76		1.19	
2.2		0.44	U	1.7		0.44	U	1.8		0.57		0.53		0.59		0.44	U	0.46	
4.2		0.92	U	2.5		0.92	U	3.1		0.92	U	0.92	U	0.92	U	0.92	U	0.92	U
26.5		5.4		30.3		4.2		25.4		9.6		6.9		8.5		4.3		5.9	
6.0		0.98	U	4.0		0.98	U	3.94		1.16		1.18		1.30		0.98	U	1.0	
22.4		3.8		19.0		3.4		19.1		5.0		4.6		4.8		2.8		3.7	
15.7		3.2		14.9		3.0		18.2		3.8		3.8		4.6		2.2		3.5	
5.7		1.3	U	2.3		1.3	U	2.4		1.3	U	1.3	U	1.3	U	1.3	U	1.3	U
5.0		0.97		4.8		0.69		7.4		2.1		1.4		1.6		0.76		1.6	
16.9		2.7		13.8		1.6		20.4		6.3		4.6		5.0		2.3		3.3	
10.5		1.0		6.6		1.3		7.7		1.9		1.7		2.1		1.2		1.4	
10.0		2.5		9.7		2.3		15.2		2.9		2.5		3.4		1.7		2.9	
26.1		2.3		12.5		2.1		20.3		7.2		5.2		3.5		2.2		3.0	
9.2		1.6		7.9		1.6		10.6		1.9		2.0		2.5		1.4		1.9	
13.2		1.9		9.3		1.5		14.1		3.4		2.8		3.6		1.4		2.9	
6.0		1.1	U	1.2		1.1	U	1.6		1.1	U	1.1	U	1.1	U	1.1	U	1.1	U
13.8		2.5		11.1		1.6		15.0		4.7		3.8		3.3		1.7		2.6	
8.6		1.6		5.7		0.9		9.8		3.2		2.4		2.7		1.2		1.5	
57.6		10.0		62.8		9.7		59.6		13.9		10.7		11.6		8.5		11.1	
12.1		1.9		11.5		2.2		11.0		2.4		2.8		3.4		1.8		2.4	
7		17		6		0.97		17.76		4.59		5.22		3.59		2.45		2.87	
9.9		1.3		6.6		1.2		8.2		2.0		2.2		2.0		1.0		1.2	
12.2		1.3		7.4		1.1		9.4		3.5		3.0		2.0		1.2		1.5	
12.2		2.4		9.3		1.9		15.4		3.7		3.5		3.2		1.8		3.1	
31.1		3.5		20.0		2.7		25.0		7.6		6.2		6.5		2.6		4.0	
32.6		3.6		17.5		4.2		25.9		6.6		5.3		6.9		3.5		3.8	
4.4		1.1		5.6		0.98		7.0		1.8		1.3		1.9		0.79		1.3	
17.0		3.6		19.9		4.0		23.3		6.4		5.6		6.4		3.5		3.6	
6.8		0.91		4.5		0.93		4.1		1.1		1.1		1.4		0.86		1.3	
2.5		0.58	U	2.7		0.58	U	3.5		0.85		0.82		0.94		0.58	U	0.60	
58.9		8.6		34.4		5.1		51.4		17.0		13.6		12.3		5.5		7.5	
9.1		1.3		5.5		0.93		8.5		2.3		2.0		1.9		1.0		1.4	
35.4		5.1		22.3		2.3		37.7		11.3		8.9		7.3		3.8		3.7	
43.3		4.7		26.3		4.6		30.5		7.9		6.9		8.0		3.5		4.8	
2.8		1.6		8.7		1.5		9.7		3.0		2.7		2.4		1.6		1.9	
2.3		1.8		8.1		1.5		10.7		2.6		2.5		2.4		1.2		1.8	
0.93		0.62	U	2.7		0.62	U	4.9		0.86		0.66		0.94		0.62	U	0.74	
0.98		0.72	U	3.3		0.72	U	4.6		1.2		1.1		0.91		0.72	U	0.80	
4.3		2.4		16.1		2.5		15.8		3.4		3.0		4.2		2.5		3.2	
3.3		2.2		10.1		2.1		14.0		3.0		3.3		3.1		1.9		2.5	
0.95		0.54		3.0		0.76		3.5		0.75		0.75		1.0		0.40		0.66	
1.7		1.2		6.1		0.99		9.0		2.7		2.1		2.0		1.1		1.4	
3.3		2.2		9.1		2.2		14.5		4.8		3.8		3.4		1.9		2.6	
2.1		1.3		5.7		0.95		9.6		3.4		2.5		2.7		1.4		1.8	
5.0		3.6		17.8		3.1		25.0		7.4		6.0		6.0		3.6		3.5	
3.0		2.5		11.5		1.9		11.1		3.2		2.8		2.8		1.7		2.2	
5.2		3.5		20.0		3.0		22.9		4.5		3.5		5.5		2.9		5.6	
5.8		3.4		21.4		2.6		21.5		6.4		5.0		5.6		2.9		3.8	
1.8		1.3		6.2		1.0		9.8		2.4		2.1		1.7		1.2		1.9	

CHAPTER 8

RECEPTOR MODELS FOR SOURCE APPORTIONMENT OF SUSPENDED PARTICLES

John G. Watson and Judith C. Chow

8.1 Introduction 273
8.2 Chemical Mass Balance (CMB) Receptor Models 275
8.2.1 Tracer Solution and Enrichment Factors 276
8.2.2 Single Sample Least Squares Minimization Solution 277
8.2.3 Multiple Sample Solutions 278
8.3 Empirical Receptor Models 280
8.3.1 Temporal and Spatial Correlation Eigenvectors 280
8.3.2 Neural Networks 281
8.3.3 Time Series 281
8.4 Physical, Chemical, and Temporal Properties 282
8.4.1 Particle Size 282
8.4.2 Chemical Composition 284
8.5 Specific Organic Compounds 287
8.5.1 Temporal and Spatial Variability 290
8.6 Source Apportionment Examples 293
8.6.1 Sources of Wintertime $PM_{2.5}$ in Denver, CO 293
8.6.2 Causes of Haze in the Mt. Zirkel Wilderness Area 294
8.6.3 Particle Deposition Near a Cement Production Facility 296
8.7 Summary and Conclusions 297
References 298

8.1 INTRODUCTION

Particulate air pollutants adversely affect human health (Vedal, 1997), visibility (Watson, 2002), materials (Grossi and Brimblecombe, 2002), and the earth's climate (Jacobson, 2002). Most suspended particles are too small to be seen with the unaided eye. Even when they can be viewed by magnification their origins are not often evident. A variety of sources contribute to a specific event, and the proportions of these contributions change from event to event. Some particles retain the forms in which they were emitted, but others are created from emitted gases through chemical reactions, or they undergo

chemical transformations that change their chemical and physical characteristics. Receptor models use the variability of chemical composition, particle size, and concentration variations in space and time to identify source types and to quantify source contributions that affect particle mass concentrations, light extinction, or deposition. This chapter describes several source apportionment receptor models, identifies measurements that can be taken to apply them, and shows examples of how these models were used to make pollution control decisions. Of necessity, these descriptions are brief, and you are guided to more complete works on each topic through selected references.

Receptor models (Cooper and Watson, 1980; Watson, 1984; Hopke, 1985, 1991, 1999, 2003; Henry, 1997a, 2002; Chow and Watson, 2002; Watson *et al.*, 2002; Brook *et al.*, 2004; Watson and Chow, 2004, 2005) provide a theoretical and mathematical framework for quantifying source contributions. These models interpret measurements of physical and chemical properties taken at different times and places to infer the possible or probable sources of excessive concentrations and to quantify the contributions from those sources. Receptor models contrast with source models (e.g., Zanneti, 2003, 2005) that combine source emission rates with meteorological transport and chemical changes to estimate concentrations at a receptor. Receptor and source models are imperfect representations of reality, and input data are seldom complete. Using both types of model helps to identify and quantify their inaccuracies and to focus further investigation on the areas of greatest discrepancy.

Some receptor models establish empirical relationships between variables that are not necessarily explainable by known physical processes. The most believable and useful receptor models have a sound physical basis, although they may apply some of the same multivariate mathematical solutions that are used in the empirical models. Examples from each type are described in the following subsections.

As with all forensic investigations, source apportionment by receptor or source models begins with plausible theories about the causes of an excessive concentration or effect. A conceptual model (e.g., Vickery, 2004) provides reasonable, though not necessarily accurate, explanations of (1) potential sources; (2) size, chemical, and temporal characteristics of particle and precursor gas emissions from these sources; (3) particle and precursor gas emissions rates; (4) meteorological conditions that affect emissions, transport, and transformation; and (5) frequency and intensity of the effect. The conceptual model is derived from previous experience (e.g., tests on similar sources, particle movement under similar meteorological conditions), the nature of the problem (neighborhood complaints, consistently poor visibility over a local area or large region, exceedance of an air quality standard), and available measurements (from air quality and meteorological sensors). Sometimes the

conceptual model yields an obvious culprit and appropriate remedial actions. If not, the conceptual model guides the location of monitoring sites, sampling periods, sampling frequencies, sample durations, selection of samples for laboratory analysis, and the particle properties that are quantified in those samples to test and discard theories until only one remains.

8.2 CHEMICAL MASS BALANCE (CMB) RECEPTOR MODELS

The CMB model (Hidy and Friedlander, 1971; Watson *et al.*, 1997, 1998a) provides the mathematical basis for most receptor models. The CMB expresses ambient chemical concentrations as the sum of products of species abundances and source contributions.

$$C_{ikl} = \sum_{j=1}^{J} F_{ij} S_{jkl} \quad \text{for } i = 1 \text{ to } I \tag{8.1}$$

The indices are defined as:

i = Pollutant, representing any quantifiable property such as an element (e.g., silicon, iron, titanium, etc.), water soluble ions (e.g., sulfate, nitrate, ammonium, sodium, etc.), an operationally defined carbon fraction (e.g., organic carbon [OC], elemental carbon [EC]), or a particle property (e.g., 10 to 30 nm diameter, long barrel-shaped with protrusions).

j = The source type, a grouping of individual source emissions with similar compositions that differ from the compositions of other source types. Common source types include geological material, sea salt, vegetative burning, cooking, diesel exhaust, and gasoline engine exhaust.

k = The sampling period, a part of the day, day of the week, season, or during a special event such as a fire or dust storm.

I = The receptor location, often selected to be exposure dominated (where people live, work, and play), transport dominated (between major source areas), source dominated (e.g., near roadways or other emitters), or boundary dominated (to determine what is entering a monitoring domain).

Using these indices, the variables in Equation (8.1) and following solutions are defined as follows:

C_{ikl} = Concentration of pollutant i for time period k at location l (unit of $\mu g/m^3$, ng/m^3, ppm, or ppb). This is the measured receptor concentration.

F_{ij} = Fractional quantity of pollutant i in source type j (unitless). These are elements of the source profile for different source types as determined by sampling effluents from representative emitters. For PM measurements, profile abundances usually are normalized to mass emissions from a source in the desired size range and averaged over several source tests.

S_{jkl} = Contribution from source type j for time period k at receptor l ($\mu g/m^3$, ng/m^3, ppm, or ppb). Source contributions are calculated by the receptor model.

σ_{Cikl} = Uncertainty of the measured receptor concentrations, usually determined by replicate analysis and propagation of analytical and flow rate errors uncertainties.

σ_{Fijl} = Uncertainty of source profile abundances, usually estimated as the standard deviation of the average from several source tests.

σ_{Sjkl} = Uncertainty of the calculated source contribution estimate, estimated by error propagation or by Monte Carlo simulation.

The CMB is physically based and can be derived from the fundamental theory of atmospheric dispersion, as detailed by Watson and Chow (2005), with the following assumptions: (1) compositions of source emissions are constant over the period of ambient and source sampling; (2) chemical species do not react with each other (i.e., they add linearly); (3) all sources with a potential for contributing to the receptor have been identified and have had their emissions characterized; (4) the number of sources or source categories is less than or equal to the number of species ($J < I$); (5) the source profiles are linearly independent of each other; and (6) measurement uncertainties are random, uncorrelated, and normally distributed. Although there are substantial real-world deviations from these assumptions, sensitivity tests demonstrate that these are often tolerable and are adequately reflected in the calculated σ_{Sjkl}. The CMB equations can be solved in several different ways, as described next.

8.2.1 TRACER SOLUTION AND ENRICHMENT FACTORS

Hidy and Friedlander (1971) assumed that there was a unique chemical component at a fixed ratio in each source profile. At the time, these were lead for vehicle exhaust, vanadium for residual oil combustion, and so on. In this case, the summation sign in Equation 8.1 disappears because all F_{ij} except the one for $i = j$ is equal to zero, and

$$S_{jkl} = C_{jkl} / F_{jj} \tag{8.2}$$

The tracer solution often is applied when using carbon-14 isotope abundances (^{14}C) to distinguish particles arising from biogenic and fossil fuel

sources (Hidy *et al.*, 2004). Owing to its radioactive decay, fossil fuel emissions contain zero ^{14}C, whereas biogenic emissions contain detectable amounts owing to their ingestion of relatively recent carbon dioxide (CO_2) that has been transformed by interaction with cosmic rays.

The Enrichment Factor (EF) model (Sturges, 1989; Rashid and Griffiths, 1993) is a variant of the tracer solution that compares ratios of chemical concentrations in the atmosphere to the same ratios in a reference material, such as average crustal composition or sea salt for a region (Mason, 1966; Lawson and Winchester, 1979). These ratios often are normalized to the silicon (Si) concentrations in air and in soil, as Si is the most abundant element in suspended dust. Reference ratios vary substantially between bulk soil and suspendable particles, and among different regions (Rahn, 1976). Differences are explained in terms of other sources, but not usually with quantitative attribution as in the CMB. Heavy metal enrichments usually are attributed to industrial emitters. Sulfur enrichment is attributed to secondary sulfate particles formed from gaseous sulfur dioxide emissions. Potassium, and sometimes chloride, enrichment is attributed to burning and cooking. Vanadium and nickel enrichments indicate contributions from residual oil combustion or refinery catalyst crackers, and enriched selenium often can be attributed to coal-fired power stations. Iron, manganese, copper, zinc, and lead enrichments usually indicate steel mill, smelting, or plating contributions. Calcium often is enriched near cement manufacture or the use of cement products in construction. Sodium and chloride are enriched near the coast, dry lake beds, and after de-icing materials are applied to streets. This method has also been used to examine different geological strata to determine how specific elements, such as lead, may have deposited over long time periods (Weiss *et al.*, 1999).

A common, and quantitative, application of the EF model (Chu, 2005) estimates the fraction of OC that derives from direct emissions (primary OC) and that which forms from conversion of organic gases during atmospheric transport (secondary OC). An OC to EC ratio of 2 to 4 often is found in directly emitted particles, but OC/EC ratios of 7 to 10 are found in some ambient samples. The excess OC is assumed to be caused by the secondary organic aerosol. The accuracy of this apportionment depends on the consistency of the OC/EC ratios in the source profiles and on consistent OC and EC measurements for source and receptor samples (Watson *et al.*, 2005).

8.2.2 SINGLE SAMPLE LEAST SQUARES MINIMIZATION SOLUTION

As toxic trace elements (e.g., lead) have been removed from many fuels, the tracer solution has become less applicable for source apportionment. What is

sought is a complex pattern, similar to a fingerprint, of components in the source profile that can also be measured at the receptor, thereby requiring a multivariate solution to the CMB equations. This is accomplished by minimizing the weighted squared differences between the calculated and measured concentrations.

$$\sum_{i=1}^{I}\left\{\left(C_{ikl}-\sum_{j=1}^{J}F_{ij}S_{jkl}\right)\Big/W_{ikl}\right\}^2 = Minimum \tag{8.3}$$

where:

$$W_{ikl}=\sqrt{\sigma_{Cikl}^2+\sum_{j=1}^{J}\sigma_{Fij}^2S_{jkl}^2} \tag{8.4}$$

The "effective variance" weighting of Equation 8.4 (Watson *et al.*, 1984) includes both the uncertainties of the receptor concentrations and the source profiles. This weighting gives less influence to source and receptor components (i) that have higher uncertainties, and more influence to those that are more precise. Since the S_{jkl} are unknown, initially they are set to zero. This is the ordinary weighted least squares solution to the CMB that was first used by Friedlander (1973), but has the disadvantage that the source profile uncertainties are often much larger than the receptor uncertainties. An iterative solution, implemented by commonly available software (Watson *et al.*, 1997), substitutes values for S_{jkl} into Equation 8.4 and repeats the process until convergence (usually within ±1%) is attained. In the single sample CMB solution, C_{ikl}, F_{ij}, σ_{Cikl}, and σ_{Fijl} are included in sample input files, and S_{jkl} and σ_{Sjkl} are returned as model outputs.

8.2.3 MULTIPLE SAMPLE SOLUTIONS

Many researchers have acquired receptor concentrations, but not the source profiles that represent their area of study. When many samples are available for different locations or time periods (k and l), the number of equations exceeds the unknowns and can be solved for the F_{ij} as well as the S_{jkl}. This requires additional assumptions that the C_{ikl} are uncorrelated in space and/or time and that the F_{ij} are constant for all the different samples. Solutions are quite sensitive to deviations from these assumptions. When pollutant concentrations covary, as they often do with changes in wind direction and dispersion, the F_{ij} derived from multiple sample solutions reflect source mixtures rather than individual source types.

The Multiple Linear Regression (MLR) model (Kleinman *et al.*, 1980; White, 1986; Lowenthal and Rahn, 1989) is the multivariate analog to the single sample tracer solution.

$$C_{(j+1)kl} = \sum_{j=1}^{J} F_{jj}^{-1} C_{jkl} \quad \text{for} \quad k = 1 \text{ to } K \text{ or } L = l \text{ to } L \tag{8.5}$$

MLR expresses particle mass, a component of mass such as sulfate, light extinction, or other observables ($C_{(j+1)kl}$) as a linear sum of unknown regression coefficients (F_{jj}^{-1}) times source marker concentrations measured at a receptor (C_{jkl}). The solution is subject to the least squares estimation constraint

$$\sum_{k=1}^{K} \left(C_{(j+1)kl} - \sum_{j=1}^{J} F_{jj}^{-1} C_{jkl} \right)^2 = Minimum \tag{8.6}$$

A substitution of l for k in the summation makes this a spatial, rather than temporal, MLR. After the tracer abundance F_{jj}^{-1} are calculated, the source contribution estimates are

$$S_{jkl} = F_{jj}^{-1} C_{jkl}, \; J = 1 \text{ to } J \tag{8.7}$$

As with the tracer method, the markers, which may be chemical elements or compounds, must originate only in the source type being apportioned. The regression coefficients represent the inverse of the source profile chemical abundance of the marker species in the source emissions. The total number of samples (K or L) must be large compared to the number of sources (J), typically more than 100.

MLR has been used to estimate source contributions to excessive particle concentrations (Morandi *et al.*, 1991), to relate light extinction to power station and smelter emissions (Malm *et al.*, 1990), and as the major epidemiological tool in relating adverse health effects to particle concentrations (Lipfert and Wyzga, 1997; Vedal, 1997; Spix *et al.*, 1999).

The Positive Matrix Factorization (PMF) solution uses samples over multiple time periods or multiple locations with the intent to derive source profiles as well as source contributions from the ambient concentrations by solving the CMB equations minimizing (Paatero and Tapper 1994; Paatero, 1998, 2000; Eberly, 2005).

$$\sum_{k=1}^{K}\sum_{i=1}^{I}\left\{\left(C_{ikl}-\sum_{j=1}^{J}F_{ij}S_{jkl}\right)\Big/W_{ikl}\right\}^2 = Minimum \tag{8.8}$$

The weighting factors can be set to the ambient concentration uncertainty or to other user-defined uncertainties to give different amounts of influence to different samples and compositions. The F_{ij} and S_{jkl} matrices are obtained by an iterative minimization algorithm. A nonnegativity constraint is built into the PMF so that source contributions and source profile abundances are always greater than zero. Selected factors are interpreted as source profiles by comparison of factor loadings with source measurements.

The UNMIX model (Henry, 2000, 2003) solves the CMB equations by placing edge constraints on potential solutions. Edges are constant ratios among chemical components that are detected in multidimensional space. Some samples must have no contribution from the sources to define an edge. The edges detected by models such as UNMIX are extensions of self-modeling curve resolution to multiple dimensions that can be associated with source profiles from independent measurements (Kim and Henry, 1999, 2000).

Although it is often said that multivariate solutions do not require source profiles, this is untrue. The derived F_{ij} must be compared with measured F_{ij} to assign them a source type. Owing to the intercorrelation among ambient data cited earlier, these associations are often uncertain and disputable.

8.3 EMPIRICAL RECEPTOR MODELS

Empirical receptor models determine relationships among ambient concentrations and other data that classify samples into different categories. Sometimes these categories correspond to chemical species groupings that can be associated with a source profile, but this is not always the case. These models can include gas concentrations and meteorological measurements that also add information concerning potential sources.

8.3.1 TEMPORAL AND SPATIAL CORRELATION EIGENVECTORS

Principal Component Analysis (PCA), Factor Analysis (FA), and Empirical Orthogonal Functions (EOF) are in this category, and there is a large variability in how these methods are applied and interpreted (e.g., Hopke, 1981, 1988; Hopke *et al.*, 1983; Henry, 1986, 1987; Henry *et al.*, 1991). The patterns from these methods are empirically derived, as a basic-principles derivation has not yet been formulated. PCA, FA, and EOF are applied to correlations between variables over time or space. Temporal correlations are calculated

from a long time series (>100 samples) of chemical concentrations at one or more locations. Eigenvectors of this correlation matrix (a smaller matrix of coordinates that can reproduce the ambient concentrations when added with appropriate coefficients) are determined and a subset is rotated to maximize and minimize correlations of each vector with each measured species. These rotated vectors (also called principal components, factors, or orthogonal functions) are interpreted as source profiles by comparison of factor loadings with source measurements. Several different normalization and rotation schemes have been used.

For a large spatial network (25 to 50 locations), spatial correlations are calculated from chemical measurements taken on simultaneous samples at the monitoring sites (Henry, 1997b,c; White, 1999). Eigenvectors of this correlation matrix represent a spatial distribution of source influences over the area, providing that the samplers have been located to represent the gradients in source contributions. As with temporal correlation models, several normalization and rotation schemes have been applied.

Eigenvector receptor models have been used to evaluate contributions to excessive particle concentrations that might affect health (Thurston and Spengler, 1985) and to determine sulfate contributions to light extinction over large regions (Henry, 1997c).

8.3.2 NEURAL NETWORKS

Known inputs and outputs are presented to a neural network that is intended to simulate the human pattern-recognition process (Wienke and Hopke, 1994; Reich *et al.*, 1999; Perez *et al.*, 2000). Training sets that have known source-receptor relationships are used to establish the linkages in the neural net that are then used to estimate source contributions for data sets with unknown relationships. The network assigns weights to the inputs that reproduce the outputs. Neural networks can provide functional relationships that are solutions to the MLR and CMB equations (Song and Hopke, 1996). Neural networks are empirical models and their results are only as good as the known relationships with the training sets and the extent to which the unknown data sets share the same patterns with the training sets.

8.3.3 TIME SERIES

Spectral analysis (Perrier *et al.*, 1995; Hies *et al.*, 2000), intervention analysis (Jorquera *et al.*, 2000), lagged regression analysis (Hsu, 1997), and trend analysis (Somerville and Evans, 1995) models separate temporal patterns for a single variable and establish temporal relationships between different

variables. These models have been used to identify sources, to forecast future pollutant concentrations, and to infer relationships between causes and effects. It is especially important to include meteorological indicators in time series models (e.g., Weatherhead *et al.*, 1998; Shively and Sager, 1999) and to use data sets with comparable measurement methods and sampling frequencies.

8.4 PHYSICAL, CHEMICAL, AND TEMPORAL PROPERTIES

The receptor models described earlier require multivariate measurements that are as specific as possible as to particle sizes, chemical and physical characteristics, temporal variation, and locations of sources. Chow (1995), Cohen and Hering (1995), McMurry (2000), and Watson *et al.* (1998b) provide details on methods that measure these properties with varying levels of time resolution, accuracy, precision, validity, and expense. Aerosol measurement technology is tending toward *in situ* monitors capable of sequential samples of one hour or less duration. However, specific chemical characterization still requires collection of particles on a substrate over periods of six to 24 hours to obtain enough material for subsequent laboratory analysis.

8.4.1 PARTICLE SIZE

Figure 8.1, an elaboration on concepts first discovered by Whitby *et al.* (1974), shows how particle mass concentrations typically are distributed according to size when suspended in the atmosphere. Particles present in different portions of this spectrum indicate their origins as well as their effects. The integral mass quantities $PM_{2.5}$ and PM_{10} currently are regulated by National Ambient Air Quality Standards (NAAQS, U.S. EPA, 1997) to protect public health and are monitored in urban areas throughout the United States. The $PM_{2.5}$ integral contains the particles that most efficiently scatter and absorb light (Watson, 2002), and baseline concentrations and compositions are measured within and around U.S. National Parks (Eldred *et al.*, 1997) to support U.S. regional haze regulations (U.S. EPA, 1999). The $PM_{0.1}$ integral, and the large number of ultrafine particles that it contains, may be another indicator of adverse health effects (Oberdorster *et al.*, 2005), as might specific chemical components of the mass (Mauderly *et al.*, 1998). The Total Suspended Particulate (TSP) integral once constituted the U.S. particulate NAAQS (U.S. EPA, 1971) and is still monitored in many parts of the world today.

Ultrafine ($PM_{0.1}$) particles are abundant in fresh combustion emissions, but they rapidly coagulate with each other and with larger particles shortly after

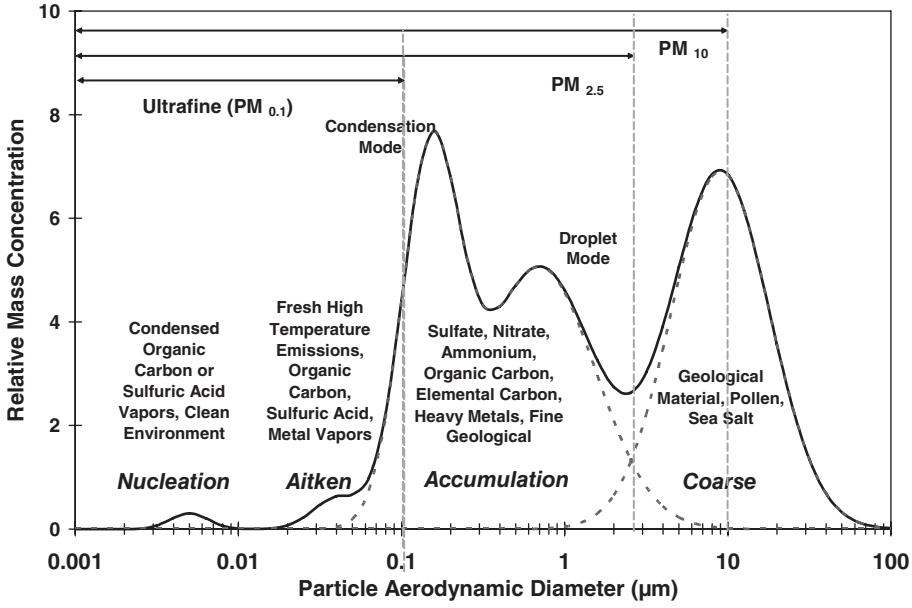


Figure 8.1
Major features of the mass distribution of atmospheric particles by particle size. This idealization was created by weighted sums of five log-normal distributions. Ultrafine particles have short lifetimes, indicating close proximity to a source.

emission (Preining, 1998). Buzorius *et al.* (1999) measured 10,000 to 50,000 particles/cm^3 near a heavily traveled street, whereas levels at a more pristine site were typically 2000 to 3000 particles/cm^3. $PM_{0.1}$ mass is <10% of $PM_{2.5}$ or PM_{10} mass, but the number of particles in this size range constitutes >80% of all particles typically found in an urban area (Keywood *et al.*, 1999).

Coarse particles, often reported as TSP minus $PM_{2.5}$ or PM_{10} minus $PM_{2.5}$ mass, result from the disaggregation of material, mostly fugitive dust. Concentrations are bounded at larger particle sizes by gravitational settling. This settling results in dustfall nuisances near coarse particle emitters such as heavily traveled unpaved roads, construction sites, and facilities that handle large amounts of minerals. As a result of settling, the peak of the coarse distribution shifts to smaller particles with transport distance from the point of emissions. Coarse particles can travel over long distances when they are injected high into the atmosphere by severe wind storms (Perry *et al.*, 1997).

The accumulation mode contains most of the $PM_{2.5}$ mass and most of its chemical variability. These particles are of both primary and secondary origins. Owing to the similarity between accumulation mode particle diameters and the wavelengths of visible light (0.3 to 0.7 μm), these particles cause most of the haze in urban areas and national parks. The two submodes in Figure 8.1 indicate secondary particle formation processes (John *et al.*, 1990). The condensation mode consists of smaller particles that form from condensation and coagulation, and the larger droplet mode forms from gases absorbed by and reacted in cloud and fog droplets. When the water in the droplet

evaporates, a larger amount of sulfate or nitrate remains than is in the condensation mode. The mode occupied by sulfate in the size spectrum allows inferences about the history of the gases that formed it.

Particle size measurements are made with size-selective inlets (Watson and Chow, 2001a) that pass material to a collection filter to be weighed or chemically analyzed. Impactors (Marple *et al.*, 1993; Hering, 1995) use several of these inlets in series to obtain a full size spectrum on substrates that can be analyzed chemically. Continuous detectors quantify the number of particles in discrete ranges for the ultrafine, accumulation, and coarse modes by a combination of their diffusion characteristics, electrical mobility, and optical properties (Cheng *et al.*, 1993; Rader and O'Hern, 1993).

8.4.2 CHEMICAL COMPOSITION

Chemical composition of particles provides the most information on their origins. The most abundant and commonly measured components of $PM_{2.5}$ or PM_{10} mass are OC, EC, geological material, sulfate, nitrate, and ammonium (Chow, 1995; Wilson *et al.*, 2002). Sodium chloride and other soluble salts often are found near the ocean, open playas, and after road de-icing. Analysis for these basic components indicates the extent to which an excessive mass concentration results from fugitive dust, secondary aerosols, or carbon-generating combustion activities. More specific chemical characterization can then be focused on the components with the largest mass fractions.

As noted previously, source profiles must contain chemical abundances (mass fractions) that differ among source types, that do not change appreciably during transport between source and receptor (or that allow changes to be simulated by measurement or modeling), and that are reasonably constant among different emitters and source operating conditions. Minor chemical components, constituting less than 1% of particle mass, are needed for quantitative apportionment as these are more numerous and more likely to occur with patterns that allow differentiation among source contributions.

Figure 8.2 compares profiles from several common sources. These are a few of the hundreds of profiles that have been compiled in various libraries (e.g., Chow *et al.*, 2004a). Aluminum, silicon, potassium, calcium, and iron are most abundant in geological material, although the abundances vary by type and use of soil. Figure 8.2a is from a paved road, and lead deposited from exhaust and exhaust systems from prior use of leaded gasoline is evident at low levels. OC from deposited exhaust, oil drippings, asphalt, and ground-up plant detritus is more abundant in this sample than in pristine desert soils. Several of the trace metals such as copper and zinc probably originate from brake and clutch linings as well as metals used in vehicle construction. Water-soluble

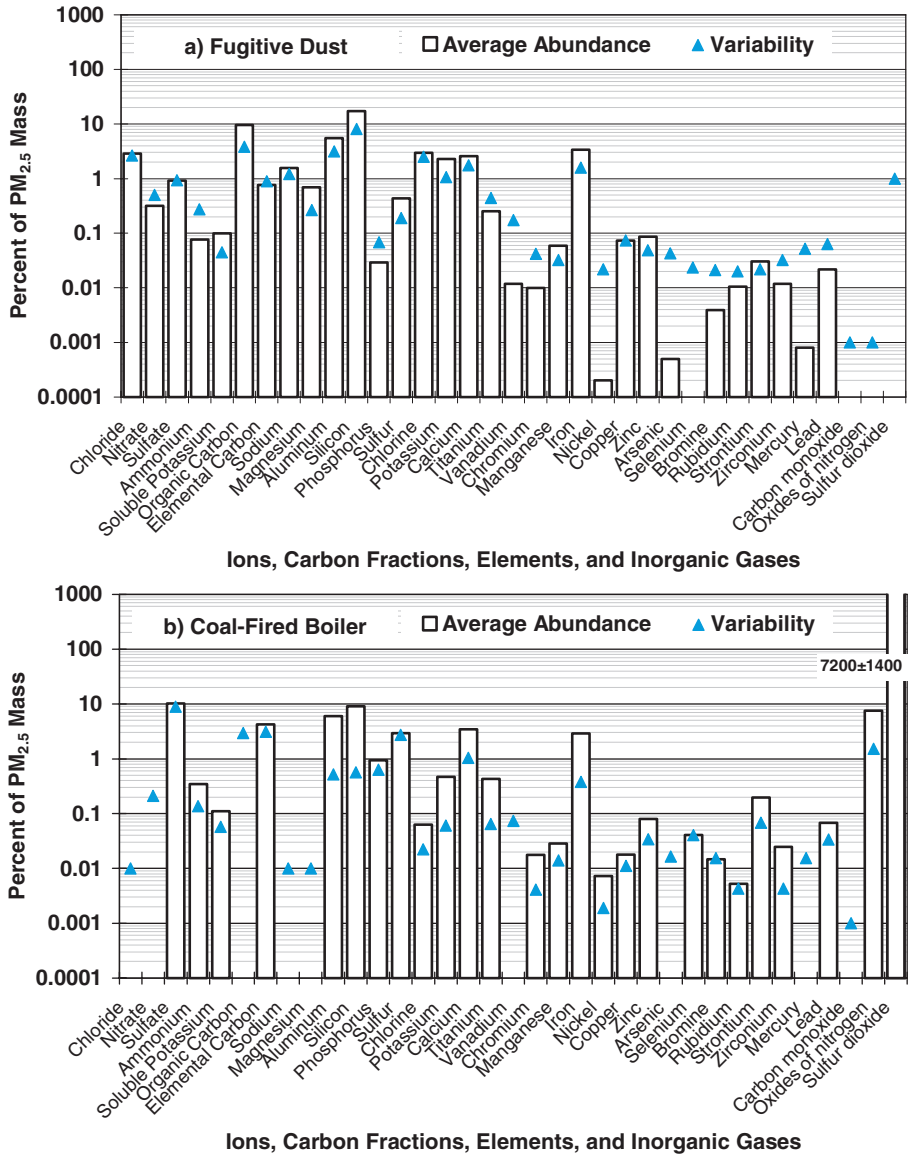
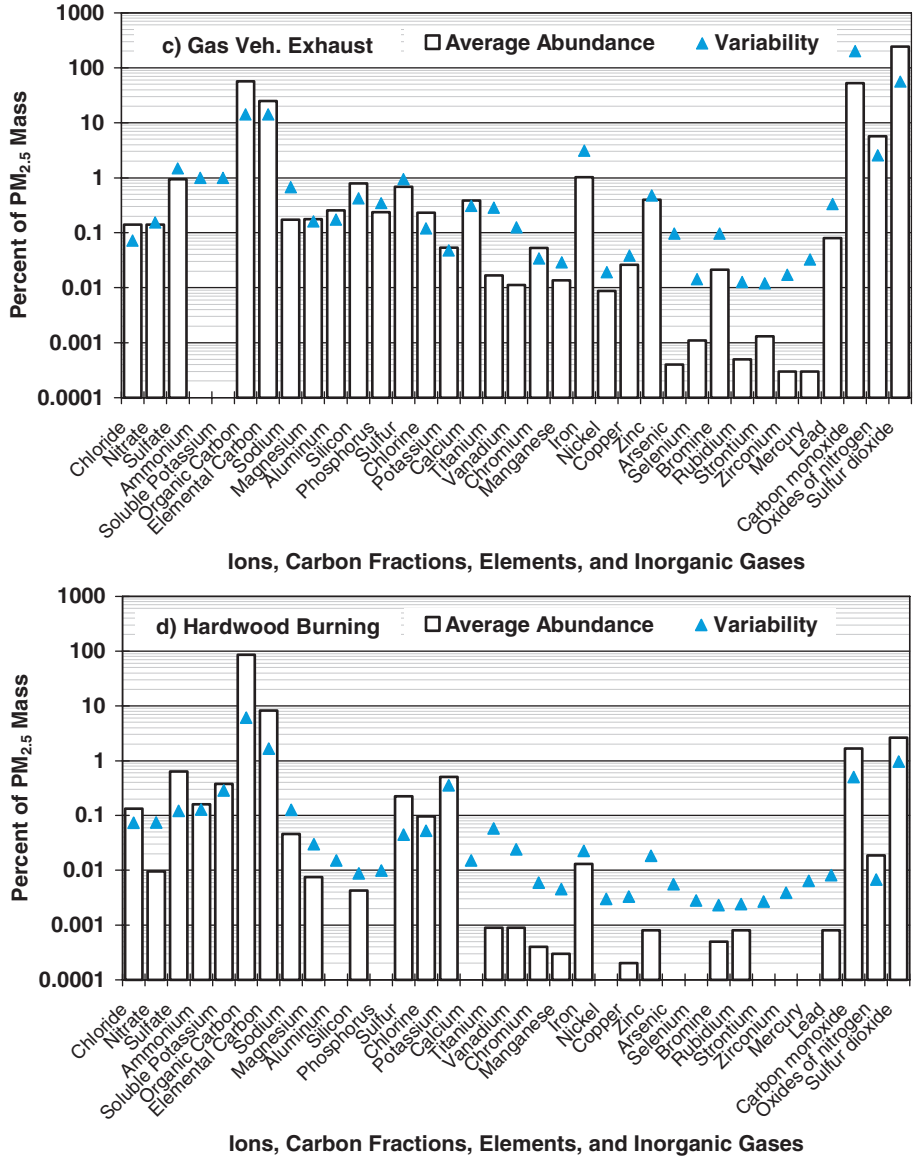


Figure 8.2

Examples of $PM_{2.5}$ source profiles determined from carbon, ion, and elemental analyses of samples from representative sources for: a) fugitive dust, b) a coal-fired power station, c) hot-stabilized gasoline vehicle exhaust, and d) hardwood burning typical of Denver, CO during winter, 1995 (Fujita et al., *1998). The height of the bar represents the average abundance of each species as determined by the average from several source tests and the triangle represents the standard deviation of the average derived from the same tests. Ratios of carbon monoxide, oxides of nitrogen, and sulfur dioxide to $PM_{2.5}$ mass emissions can also be included in source profiles.*

potassium in dust profiles is about one-tenth of the total potassium abundance. Figure 8.2b from a coal-fired power station contains many of the same elements found in fugitive dust, but in different proportions. Selenium, strontium, and lead are enriched over the dust profile, and OC is depleted. The ratio of sulfur dioxide gas to $PM_{2.5}$ emissions in this sample is large, and this would be a good indicator of fresh emissions from this source type. This ratio would change with time, however, as sulfur deposits more rapidly than $PM_{2.5}$

Figure 8.2
Continued



and augments $PM_{2.5}$ mass as it transforms into secondary sulfate within a day after emission.

The gasoline vehicle exhaust profile in Figure 8.2c shows OC and EC as its major components, with OC two to three times EC. A wide variety of different elements is found in these emissions, but they are highly variable. This is consistent with many of these metals originating from exhaust system deterioration and motor oil that will change substantially from vehicle to vehicle. The carbon monoxide to $PM_{2.5}$ ratio is highest for gasoline vehicle exhaust, but it

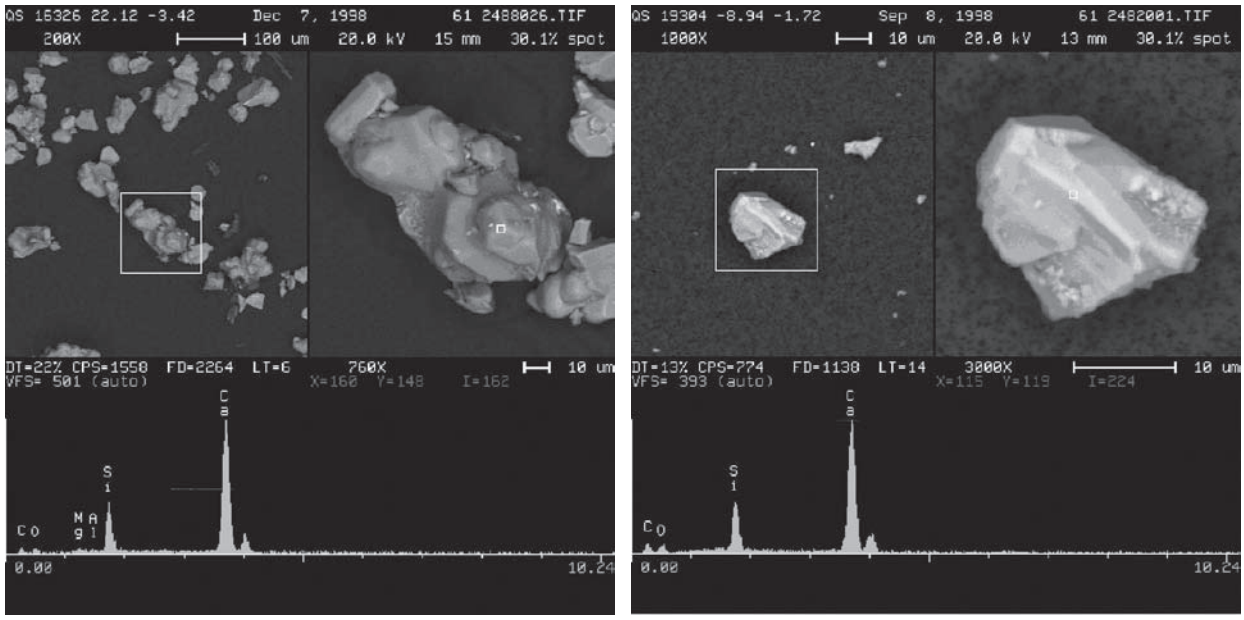
is also highly variable. Figure 8.2d shows that woodburning emissions are also dominated by carbon, but the OC abundance is usually much higher than the EC fraction. Most of the potassium is water-soluble, in contrast to the insoluble potassium in fugitive dust. Other soluble compounds containing sulfate, nitrate, ammonium, and chloride are evident. Trace metals from dust deposited on the burned wood and from stove or fireplace deterioration are often evident in these emissions.

As noted in the discussion of enrichment factors earlier, trace elements have been used in the past as markers for different source contributions. However, having identified and quantified residual oil combustion, leaded gasoline combustion, and metal and refining processes as contributors to excessive particle mass and toxic element concentrations, these trace elements have been removed from many fuels and processes. Although good for the environment, these modifications have degraded the utility of elemental abundances to distinguish among source contributions in receptor models. Additional chemical and physical properties beyond the commonly measured elements, ions, and carbon need to be examined to determine the ones that are useful in practical source apportionment applications.

Optical microscopy magnifies particles between ~2 and 30 μm so that their shapes, sizes, mineralogy, and optical properties can be related to similar particles from representative sources (McCrone and Delly, 1973; Draftz, 1982). Electron microscopy can be applied to smaller particles (Casuccio *et al.*, 1983, 1989; also see Chapter 13). As shown in Figure 8.3, both shape and elemental content are determined by this method. In this case, the elemental spectra from a clinker and a cement particle are very similar, but the more polished surface of the cement clearly distinguishes it from the clinker that precedes it in the cement production process (Greer *et al.*, 2000). Results from microscopic examination are semiquantitative, as thousands of particles must be examined to obtain an adequate statistical representation of the millions deposited on a typical air sampling filter. Computer-automated scanning systems with pattern recognition methods are being developed to perform these analyses automatically (Hopke and Casuccio, 1991).

8.5 SPECIFIC ORGANIC COMPOUNDS

The OC components shown in Figure 8.2 are composed of thousands of individual compounds, but the detected compounds amount to only 10% to 20% of the organic compound mass. Specific organic compounds may be analyzed by solvent extraction followed by gas or liquid chromatography with different detectors (e.g., Mazurek *et al.*, 1987). Figure 8.4 compares a variety of different organic compound abundances in hardwood burning with those in meat



a b

Figure 8.3

Backscattered electron images and elemental spectrum illustrating: a) clinker particles and b) finished cement particles (Watson et al.*, 1999). The elemental composition of the particles shows them both to be rich in calcium and silicon, but their appearances are quite different.*

cooking. Meat cooking profiles for elements, ions, and carbon are similar to those for wood burning (e.g., Figure 8.2d). The additional organic compounds shown in Figures 8.4a and 8.4b are quite different, however. Hardwood burning is rich in guaiacols and syringols, but low in sterols such as steroid-m and cholesterol. Just the opposite is true for the meat cooking, where cholesterol is among the most abundant species. Syringols are more abundant in hardwoods, such as oak or walnut, and they are depleted in softwoods such as pine, thereby allowing even greater differentiation to be achieved in source apportionment. Simoneit (1999) cites several examples of odd and even numbered carbon molecules in the n-alkane series as indicating the presence or absence of OC from manmade sources in the presence of ubiquitous contributions from natural sources. Biogenic material, such as plant waxes or secondary organic aerosols from vegetative hydrocarbon emissions, tend to have more molecules with odd numbers of carbon atoms, whereas carbon from combustion processes has nearly equal quantities of even and odd carbon-numbered molecules. A Carbon Preference Index (CPI), the ratio of odd to even n-alkane masses, is used to estimate the relative abundances of odd and even compounds to separate natural from manmade contributions. Some organic compound source profiles have been measured for gasoline and diesel

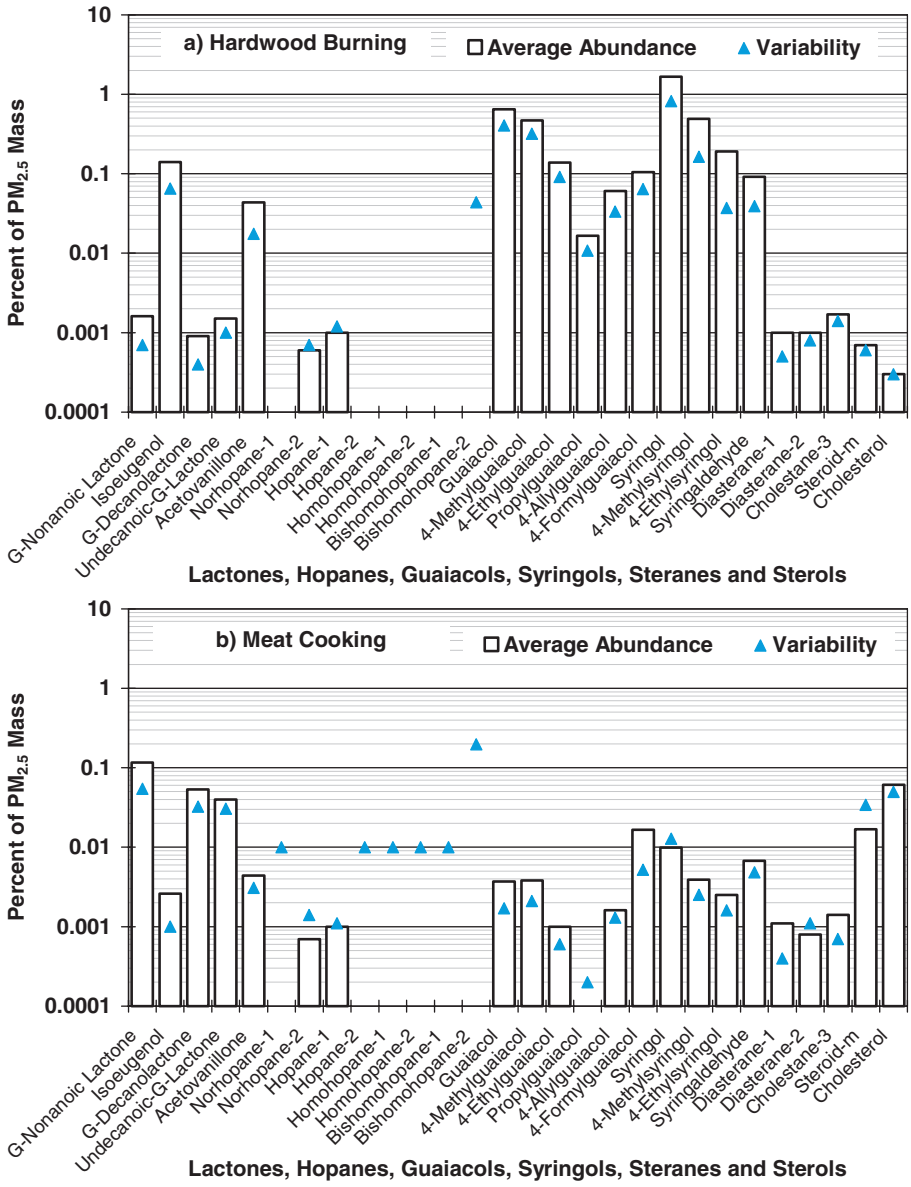


Figure 8.4

Examples of $PM_{2.5}$ source profiles for lactones, hopanes, guaiacols, syringols, steranes, and sterols for: a) hardwood combustion and b) meat cooking (Fujita et al., *1998). These are among many different categories of molecular organic markers measured by filter extraction and gas chromatography with mass spectrometric detection.*

vehicle exhaust (Rogge *et al.*, 1993a; Schauer *et al.*, 1999a), meat cooking (Nolte *et al.*, 1999; Schauer *et al.*, 1999b), natural gas combustion (Rogge *et al.*, 1993b), coal combustion (Oros and Simoneit, 2000), oil combustion (Rogge *et al.*, 1997), wood burning and forest fires (Standley and Simoineit, 1994; Rogge *et al.*, 1998; Simoneit *et al.*, 1999; Elias *et al.*, 1999; McDonald *et al.*, 2000), and fugitive dust (Rogge *et al.*, 1993c). Most of these results are from emitters in southern California. Different investigators measure and use different markers for these sources in receptor models.

Solvent extraction methods require large samples (often composites of many samples) and are costly. Thermal desorption methods have been perfected (Hays *et al.*, 2004; Ho and Yu, 2004) that evaporate OC directly from a small sample, only a fraction of the normal 47 mm diameter circular filters that are acquired in speciation networks. Compounds in the desorbed materials are separated by a gas chromatograph and detected with a mass spectrometer. There are clear differences between the patterns from different sources, as seen in Figure 8.5. As more source samples are analyzed by these methods, it will become clearer which patterns (spikes and humps in Figure 8.5) are similar within a source type, but different between types. This method holds great potential for source apportionment, even in the absence of associating each pattern with a specific chemical component.

The OC and EC groupings in Figure 8.2 are operationally defined by a thermal evolution method that also monitors the changing reflectance of the sample (Chow *et al.*, 1993, 2001, 2004b). This is only one of many ways in which carbon can be separated into functional and repeatable categories for which the exact composition is not precisely known (Watson *et al.*, 2005). Watson *et al.* (1994) found differences in the quantities of carbon that evolved at different temperatures, with diesel exhaust yielding more carbon at higher temperatures than gasoline vehicle exhaust. Cao *et al.* (2005) and Kim and Hopke (2004a,b) demonstrate how even these thermal fractions can be used to distinguish among source contributions.

Isotopic abundances differ among source material depending on its formation process or geologic origin, with ^{14}C having been cited earlier as an indicator of biogenic origin. Estimates of fuel-age must be made, as wood that grew before outdoor nuclear testing in the 1950s and early 1960s has much lower ^{14}C abundances than plant life that grew during and after that period. Other receptor model applications using isotopic abundances include ^{04}He as a marker of continental dust (Patterson *et al.*, 1999), ^{210}Pb to determine sources of lead in house dust (Adgate *et al.*, 1998), and ^{34}S to determine sources of acid deposition (Turk *et al.*, 1993), and to distinguish secondary sulfate contributions from residual oil combustion (Newmann *et al.*, 1975). Isotopes such as nonradioactive ^{34}S or radioactive ^{35}S follow the transformations of sulfur dioxide to sulfate from a source type and offer a way to directly apportion the secondary aerosol in cases where there are large differences in profiles among contributing sources.

8.5.1 TEMPORAL AND SPATIAL VARIABILITY

Sharp spikes of 1- to 60-minute duration indicate contributions from nearby sources (Watson *et al.*, 2001). Diurnal patterns show variations over durations

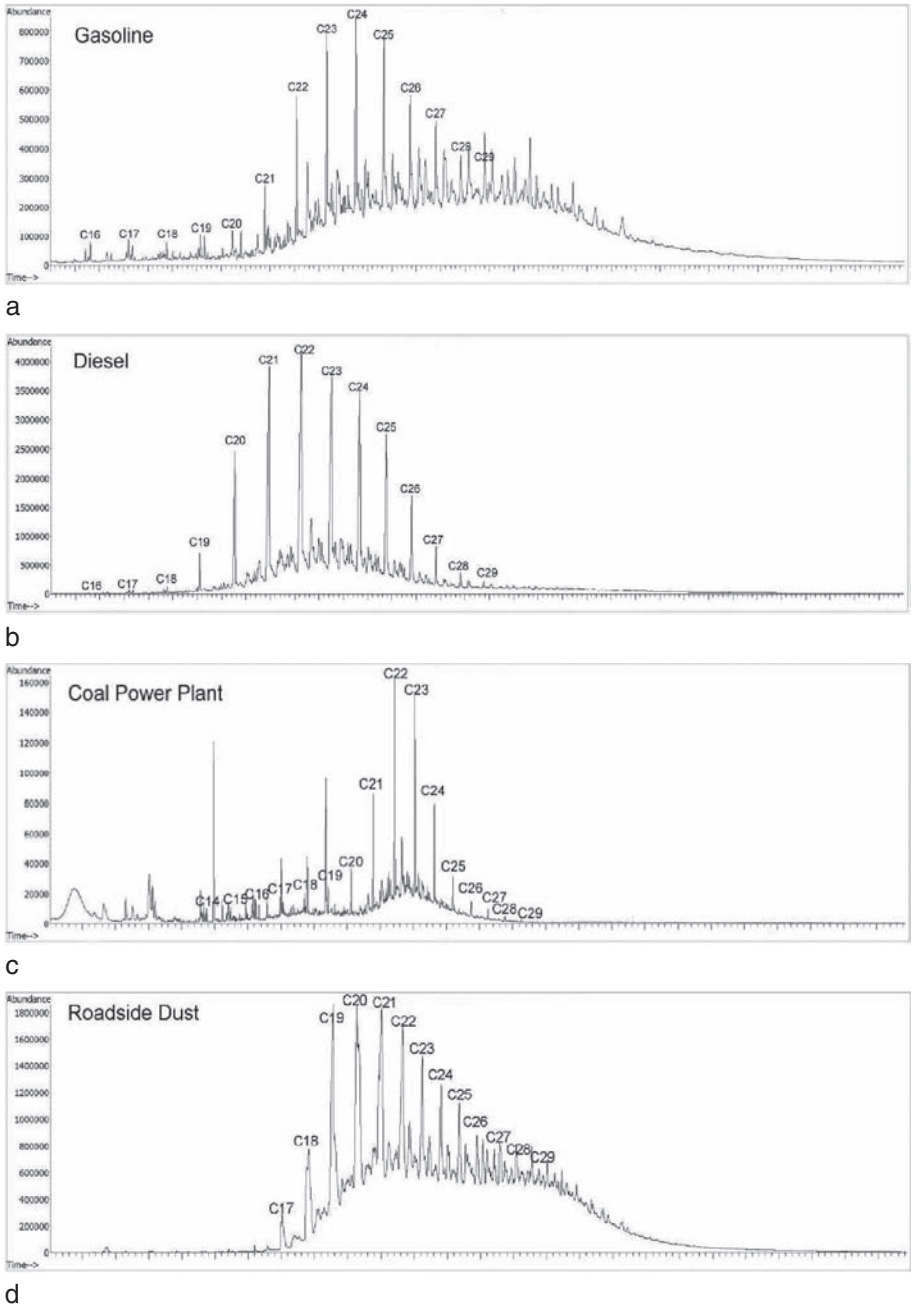


Figure 8.5

Thermal desorption gas chromatographic (Ho and Yu, 2004) mass spectra for ion 57, representative of alkanes. The C labels indicate the number of carbon atoms in each peak. Similar spectra are produced for a large number of ions. Note differences in the peak structure as well as in the size and shape of the hump under the peaks for different source types.

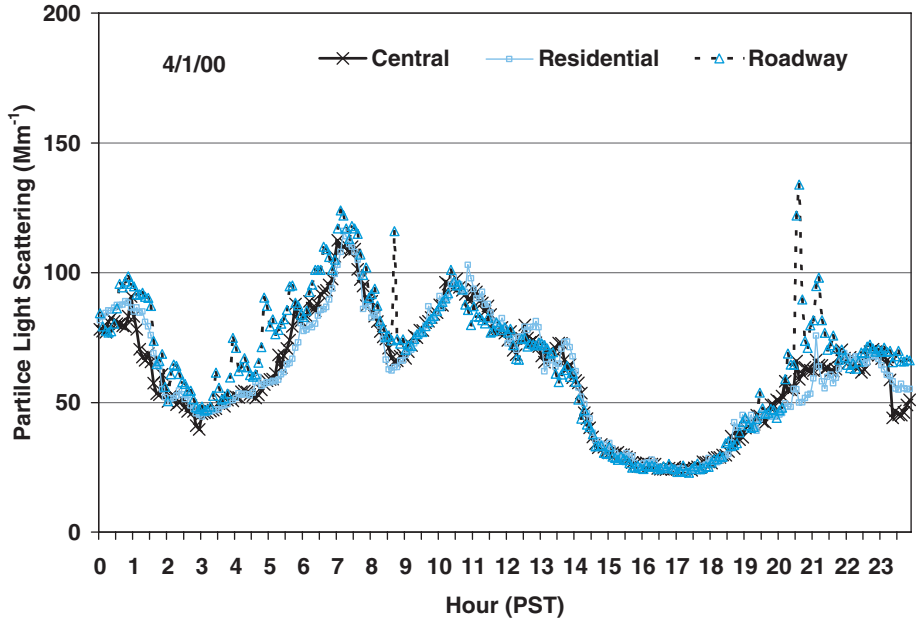


Figure 8.6
Particle light scattering measurements at a central Fresno, CA, compliance monitoring sites compared to satellite sites in a residential neighborhood ~0.5 km east and near a heavily traveled road ~1 km west (Watson et al.*, 2000a). Dividing light scattering by 2 to 5 g/m^3 provides an estimate of $PM_{2.5}$ mass, a conversion that can be made more precise by comparison with periodic collocated filter sampling with battery-powered monitors.*

of several hours. A morning hump in $PM_{2.5}$ or PM_{10} concentrations often is observed on weekdays in urban areas. This hump disappears on weekends when traffic volumes are more evenly distributed throughout the day (e.g., Chow *et al.*, 1999). High PM_{10} levels that correspond to high wind speeds are often indicators of windblown dust. Short-term concentration spikes that correspond to specific wind directions indicate the direction of the source with respect to the receptor.

Figure 8.6 compares light scattering measurements from nearby locations taken with an inexpensive integrating nephelometer. This nephelometer was equipped with a smart heater that raises the temperature of the airstream when relative humidities rise above 70% to maintain it at less than 70% relative humidity (Richards *et al.*, 2001). This heating minimizes interferences from liquid water absorbed by soluble particles while retaining volatile particles such as ammonium nitrate that evaporate at temperatures exceeding ~15˚C. This system can be battery- and solar-powered for mounting on power poles with minimal expense. It is a cost-effective way to resolve environmental justice concerns in which some communities believe that their exposure to particles is not well-represented by a compliance monitor in another neighborhood.

Even though the monitors in Figure 8.6 are separated by more than 1.5 km, the general diurnal pattern appears as if they were right next to each other. Short-duration spikes show the effects of local emissions that affect one monitor but not the others. At the roadway site these represent contributions

from nearby vehicle exhaust or road dust. At the residential site during spring, nighttime spikes from home heating are evident.

8.6 SOURCE APPORTIONMENT EXAMPLES

The CMB has been applied to evaluate particle control effectiveness (e.g., Chow *et al.*, 1990; Engelbrecht *et al.*, 2000), determine causes of excessive human exposure (Vega *et al.*, 1997; Schauer and Cass, 2000), quantify contributions to light extinction (e.g., Watson *et al.*, 1988, 1996), and to estimate sources of particle deposition (Feeley and Liljestrand, 1983). Although most often applied to ambient air samples for PM and organic gases, CMB modeling software (Watson *et al.*, 1990, 1997) has also been used for forensic applications in quantifying contributions to oil spills (Pena-Mendez *et al.*, 2001), toxic components in lake-bed sediments (Kelly and Nater, 2000), and water quality. The following examples are from actual problems that the authors were commissioned to solve using the methods described here. Each summary provides a statement of the problem, the technical approach, study results, and the decisions that were taken.

8.6.1 SOURCES OF WINTERTIME $PM_{2.5}$ IN DENVER, CO

The Denver Brown Cloud has been the subject of five major studies, and numerous smaller ones, since 1973. The 1987–1988 study (Watson *et al.*, 1988) identified secondary ammonium nitrate and carbon as the largest chemical components, with wood smoke and vehicle exhaust being the largest contributors to the carbon as determined by the CMB model. With the available measurements, diesel contributions could not be distinguished from gasoline vehicle exhaust and wood-burning could not be distinguished from cooking contributions. The 1996–1997 study (Watson *et al.*, 1998c) measured organic markers that were sufficient to separate these source contributions, as well as gasoline-exhaust contributions from cold-starts and poorly maintained vehicles. Relative contributions from the emissions inventory and CMB source apportionment are compared in Figure 8.7.

The 1996–1997 study showed that measures to reduce wood burning emissions implemented after 1988 had been effective. Approximately half of the $PM_{2.5}$ that would have been attributed to wood smoke without the organic markers was found to derive from meat cooking, as evidenced by the ambient sterol levels. Fugitive dust emissions were overestimated by the inventory with respect to their proportions in ambient air, typically of most inventories that overestimate fugitive dust (Watson *et al.*, 2000b).

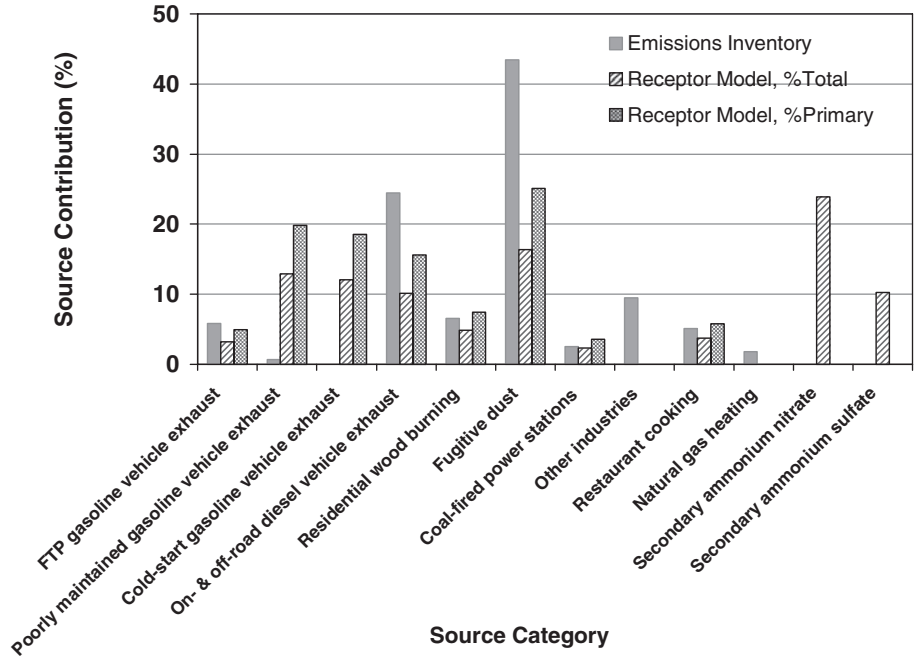


Figure 8.7
Comparison of fractional source contributions to $PM_{2.5}$ during 1996 in Denver, CO, (Watson et al., *1998c) from an emissions inventory, receptor model source contributions as fraction of total $PM_{2.5}$, and receptor model source contributions as fraction of primary $PM_{2.5}$ ($PM_{2.5}$ minus secondary ammonium nitrate and ammonium sulfate contributions).*

The most important result in Figure 8.7 is the large contribution from gasoline vehicle exhaust, and that most of this is from off-cycle operations and a relatively low number of very high emitters. The inventory grossly underestimates these contributions and they were not considered in control strategy development; most of the emphasis was placed on diesel exhaust, which the inventory showed to contribute ~80% of vehicular emissions.

As a result of this study, inventory improvements are intended to better quantify gasoline vehicle exhaust, cooking, and fugitive dust emissions. Although diesel emissions reductions are still being considered, additional control measures are being investigated for the other sources.

8.6.2 CAUSES OF HAZE IN THE MT. ZIRKEL WILDERNESS AREA

The Mt. Zirkel Wilderness Area in northwestern Colorado is defined as a Class I area by the United States Clean Air Act, as are most National Parks and Wilderness Areas in the United States, for which visibility is a protected resources. U.S. EPA regulations require remediation measures when significant visibility impairment in these areas has been certified and is "reasonably attributable" to an existing industrial source or a group of sources. The Yampa Valley, to the west of the Wilderness, contains two coal-fired power stations with five generating units that were identified as potential contributors to the

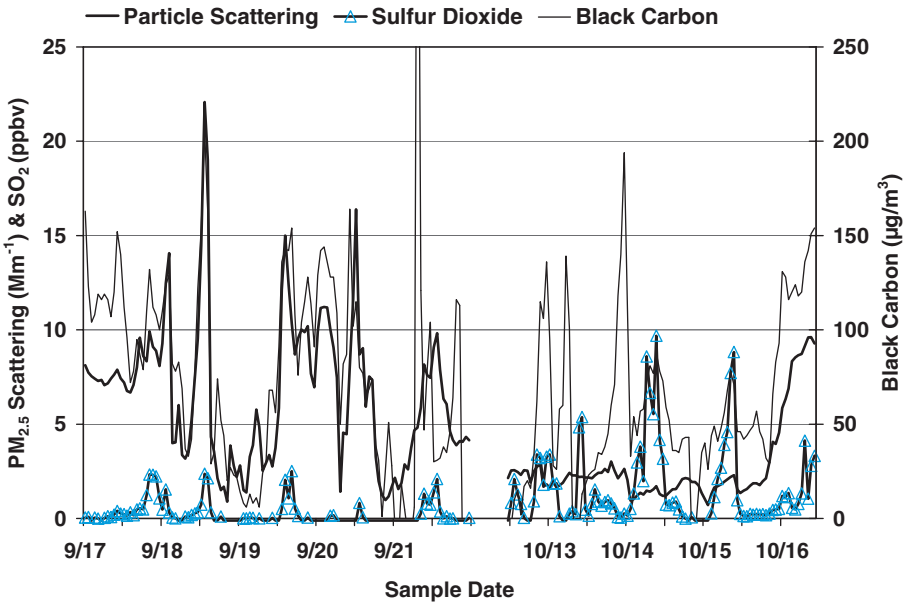


Figure 8.8
Time series of 1-hour average particle light scattering, sulfur dioxide, and black carbon, at the Buffalo Pass site near the Mt. Zirkel Wilderness area for episodes of 9/17/95 through 9/21/95 and of 10/13/95 through 10/16/95 (Watson et al., *1996).*

haze in the Wilderness. Watson *et al.* (1996) conducted a study to determine the frequency and intensity of hazes at the Wilderness caused by contributions from different sources. Many of these approaches were used together, including source dispersion modeling and CMB modeling with simulated aging for source profiles.

High time resolution measurements were also taken for sulfur dioxide, black carbon, and light scattering, as shown in Figure 8.8. The 9/17/95 through 9/21/95 episode shows a correspondence between increases in sulfur dioxide concentrations, indicative of local power station emissions, and light scattering. Black carbon, typical of contributions from nonpower station combustion, also covaries with scattering. The early afternoon increases corresponded to the lifting and transport of emissions that had accumulated in the Yampa Valley during overnight fogs that accelerated the transformation of gaseous sulfur dioxide to sulfate. This was confirmed by chemical measurements within the valley as well as measurement of upper air winds and modeling. The phenomenon could also be seen on time-lapsed video.

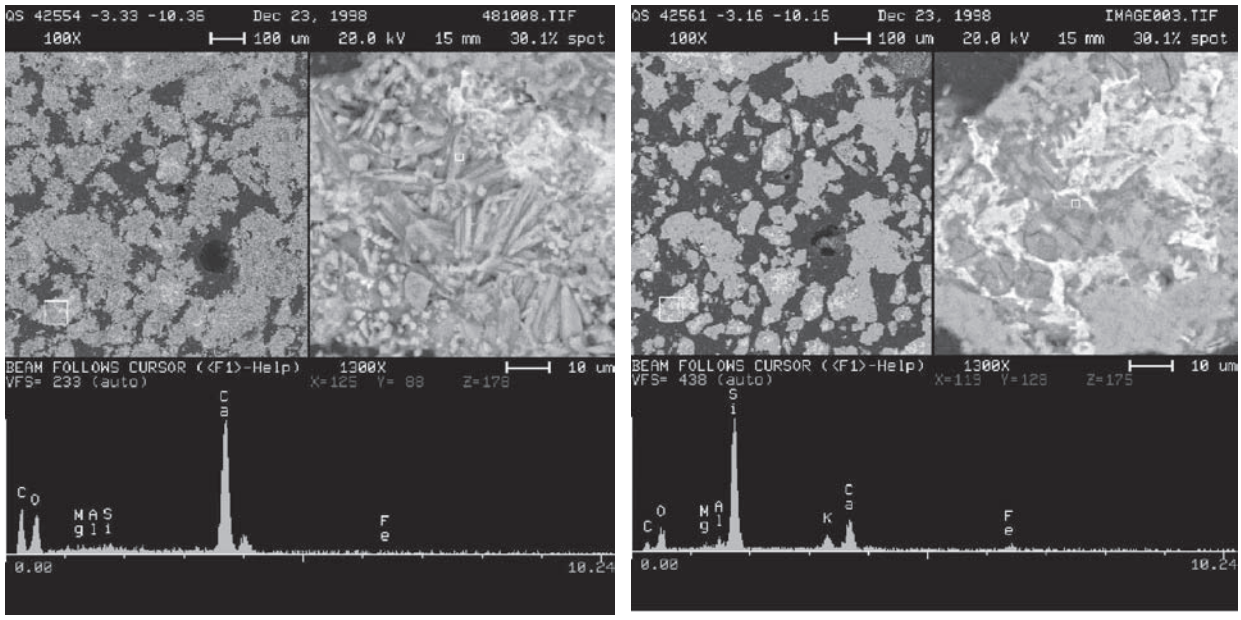
The 10/13/95 through 10/16/95 episode shows a clear presence of power station emissions at the Wilderness site in the high sulfur dioxide pulses, but it shows no indication of corresponding increases in light scattering. Visibility was good for this case and there was insufficient time for gas-to-particle conversion in the absence of aging in clouds and fogs. As a result of this study and other considerations a settlement was negotiated to add state-of-the art sulfur dioxide scrubbers to uncontrolled Yampa Valley power stations.

8.6.3 PARTICLE DEPOSITION NEAR A CEMENT PRODUCTION FACILITY

Residents of a community near a cement production facility along the California coast experienced episodes of deposited dust that adhered to their houses, windows, and vehicles, making it difficult to remove. Several sources of fugitive dust existed in the area, including agricultural tilling and harvesting, traffic suspension from paved and unpaved roads, construction, cement production, and wind erosion from open ground. Previous studies of PM_{10} in the area (Watson, 1998) had found no concentrations exceeding health standards with minor contributions from the cement plant. Deposition plates and scanning electron microscopy were used for long-term monitoring and during events to determine the sources of larger particles that created deposition.

The study showed that the deposited dust consisted of a mixture of substances, including suspended soil, sea salt, limestone, clinker, and cement. Calcium-rich crystalline particles were present in those samples that citizens found to adhere to surfaces. The deposited particles were not similar in appearance or chemical composition to the primary emissions from any of the potential sources, including the clinker and finished cement shown in Figures 8.3a and b. Samples of the different source materials were deliberately placed on deposition plates and left in a moist outside environment for several weeks. The sea salt dissolved, and most of the mineral materials retained their original shapes and compositions. The cement and clinker samples, however, reacted with other deposited material to form a coating similar to that of the deposition samples. The silicate portion was missing from the electron-excited x-ray spectrum, as shown in Figure 8.9a. When the surface layer was removed from the test sample, however, it was found that the crystalline calcium surrounded a more complex core of silicon-rich particles that were typical of the clinker and cement, as seen in Figure 8.9b. Although the deposited particles did not look like the primary particles, the changes and interactions with weather and other particles could be simulated to determine the origin.

A survey was conducted of the clinker and cement operations that resulted in (1) paving and sweeping the road to the clinker storage shed, (2) installing a retaining wall around the base of the storage shed to reduce damage to the shed walls by heavy equipment, (3) creating a curtained airlock with a dust collector at the shed exit, (4) overhauling dust-collectors on the cement storage silos, and (5) installing automatic detection and shutoff systems to prevent overfilling of silos. The deposition surveillance program is being continued by community members to determine improvement trends and to identify additional episodes that can be analyzed to further reduce emissions.



a b

Figure 8.9

Backscattered electron image and elemental spectrum of a) a polished cross-section of cement particles after exposure to a moisture-enriched environment and b) the same sample after several microns were removed from the surface (Watson et al.*, 1999). Note that the elemental spectrum of Figure 8.9b is similar to the spectra for clinker and cement particles in Figure 8.3a.*

8.7 SUMMARY AND CONCLUSIONS

Particle size, chemical, temporal, and spatial patterns have been used to solve several air quality problems related to public health, visibility, and deposition nuisances. These have employed a variety of measurement technologies and mathematical receptor models that interpret the measurements in terms of source contributions and/or their locations. The advent of *in situ* continuous monitoring for different chemical components as a function of particle size offers new opportunities to combine these different properties into a unified model. Portable and inexpensive monitors, especially for light scattering, offer opportunities to develop spatial and temporal patterns close to sources and in different neighborhoods that might be affected. Challenges remain to quantify most of the molecular compounds that constitute OC, one of the largest particle components; only 10% to 20% of the OC mass has been identified to date. Although many source types have been characterized, the measurements reported by different investigators often differ from each other and from measurements available at receptors. Particle transformation between source and receptor must be accounted for by mathematical or empirical

simulation. These limitations argue for using particle pattern recognition methods in conjunction with other methods, such as source dispersion modeling, that can provide independent source attribution. Reconciling the results of these independent approaches then allows uncertainties to be estimated and for experiments designed to reduce those uncertainties.

REFERENCES

Adgate, J.L., Willis, R.D., Buckley, T.J., Chow, J.C., Watson, J.G., Rhoads, G.G., and Lioy, P.J. (1998) Chemical mass balance source apportionment of lead in house dust. *Enivron. Sci. Technol.* 32, 108–114.

Brook, J.R., Vega, E., and Watson, J.G. (2004) Chapter 7: Receptor Methods. In *Particulate Matter Science for Policy Makers—A NARSTO Assessment, Part 1*. J.M. Hales and G.M. Hidy, eds. Cambridge University Press, London, UK, 235–281.

Buzorius, G., Hämeri, K., Pekkanen, J., and Kulmala, M. (1999) Spatial variation of aerosol number concentration in Helsinki City. *Atmos. Environ.* 33, 553–565.

Cao, J.J., Wu, F., Chow, J.C., Lee, S.C., Li, Y., Chen, S.W. *et al.* (2005) Characterization and source apportionment of atmospheric organic and elemental carbon during fall and winter of 2003 in Xi'an, China. *Atmos. Chem. Physics* 5, 3127–3137, 1680-7324/acp/2005-5-3127.

Casuccio, G.S., Janocko, P.B., Lee, R.J., Kelly, J.F., Dattner, S.L., and Mgebroff, J.S. (1983) The use of computer controlled scanning electron microscopy in environmental studies. *J. Air Poll. Control Assoc.* 33, 937–943.

Casuccio, G.S., Schwoeble, A.J., Henderson, B.C., Lee, R.J., Hopke, P.K., and Sverdrup, G.M. (1989) The use of CCSEM and microimaging to study source/receptor relationships. In *Transactions, Receptor Models in Air Resources Management*, J.G. Watson, ed. Air & Waste Management Association, Pittsburgh, PA, 39–58.

Cheng, M.D., Hopke, P.K., and Zeng, Y. (1993) A receptor-oriented methodology for determining source regions of particulate sulfate observed at Dorset, Ontario. *J. Geophys. Res.* 98, 16839–16850.

Chow, J.C., Watson, J.G., Egami, R.T., Frazier, C.A., Lu, Z., Goodrich, A., and Bird, A. (1990) Evaluation of regenerative-air vacuum street sweeping on geological contributions to PM_{10}. *J. Air Waste Manage. Assoc.* 40, 1134–1142.

Chow, J.C., Watson, J.G., Pritchett, L.C., Pierson, W.R., Frazier, C.A., and Purcell, R.G. (1993) The DRI Thermal/Optical Reflectance carbon analysis system: Description, evaluation and applications in U.S. air quality studies. *Atmos. Environ.* 27A, 1185–1201.

Chow, J.C. (1995) Critical review: Measurement methods to determine compliance with ambient air quality standards for suspended particles. *J. Air Waste Manage. Assoc.* 45, 320–382.

Chow, J.C., Watson, J.G., Green, M.C., Lowenthal, D.H., DuBois, D.W., Kohl, S.D. *et al.* (1999) Middle- and neighborhood-scale variations of PM_{10} source contributions in Las Vegas, Nevada. *J. Air Waste Manage. Assoc.* 49, 641–654.

Chow, J.C., Watson, J.G., Crow, D., Lowenthal, D.H., and Merrifield, T.M. (2001) Comparison of IMPROVE and NIOSH carbon measurements. *Aerosol Sci. Technol.* 34, 23–34.

Chow, J.C. and Watson, J.G. (2002) Review of $PM_{2.5}$ and PM_{10} apportionment for fossil fuel combustion and other sources by the chemical mass balance receptor model. *Energy & Fuels* 16, 222–260.

Chow, J.C., Watson, J.G., and Chen, L.-W.A. (2004a) Contemporary inorganic and organic speciated particulate matter source profiles for geological material, motor vehicles, vegetative burning, industrial boilers, and residential cooking. Prepared for Pechan and Associates, Inc., Springfield, VA, Desert Research Institute, Reno, NV.

Chow, J.C., Watson, J.G., Chen, L.-W.A., Arnott, W.P., Moosmüller, H., and Fung, K.K. (2004b) Equivalence of elemental carbon by Thermal/Optical Reflectance and Transmittance with different temperature protocols. *Environ. Sci. Technol.* 38, 4414–4422.

Chu, S.H. (2005) Stable estimate of primary OC/EC ratios in the EC tracer method. *Atmos. Environ.* 39, 1383–1392.

Cohen, B.S. and Hering, S.V. (1995) *Air Sampling Instruments for Evaluation of Atmospheric Contaminants*. Am. Conf. of Govnmtl. Industrial Hygienists, Inc., Cincinnati, OH.

Cooper, J.A. and Watson, J.G. (1980) Receptor oriented methods of air particulate source apportionment. *J. Air Poll. Control Assoc.* 30, 1116–1125.

Draftz, R.G. (1982) Distinguishing carbon aerosols by microscopy. In *Particulate Carbon: Atmospheric Life Cycle*, G.T. Wolff and R.L. Klimisch, eds. Plenum Press, New York, NY, 261–271.

Eberly, S. (2005) EPA PMF 1.1 User's Guide. U.S. Environmental Protection Agency, Research Triangle Park, NC.

Eldred, R.A., Cahill, T.A., and Flocchini, R.G. (1997) Composition of $PM_{2.5}$ and PM_{10} aerosols in the IMPROVE Network. *J. Air Waste Manage. Assoc.* 47, 194–203.

Elias, V.O., Simoneit, B.R.T., Pereira, A.S., Cabral, J.A., and Cardoso, J.N. (1999) Detection of high molecular weight organic tracers in vegetation smoke samples by high-temperature gas chromatography-mass spectrometry. *Environ. Sci. Technol.* 33, 2369–2376.

Engelbrecht, J.P., Swanepoel, L., Zunckel, M., Chow, J.C., Watson, J.G., and Egami, R.T. (2000) Modelling PM_{10} aerosol data from the Qalabotjha low-smoke fuels macro-scale experiment in South Africa. *Ecological Modelling* 127, 235–244.

Feeley, J.A. and Liljestrand, H.M. (1983) Source contributions to acid precipitation in Texas. *Atmos. Environ.* 17, 807–814.

Friedlander, S.K. (1973) Chemical element balances and identification of air pollution sources. *Enivron. Sci. Technol.* 7, 235–240.

Fujita, E.M., Watson, J.G., Chow, J.C., Robinson, N.F., Richards, L.W., and Kumar, N. (1998) Northern Front Range Air Quality Study. Volume C: Source apportionment and simulation methods and evaluation. Prepared for Colorado State University, Cooperative Institute for Research in the Atmosphere, Ft. Collins, CO, Desert Research Institute, Reno, NV.

Greer, W.L., Dougherty, A., and Sweeney, D.M. (2000) Mineral products industry—Portland Cement. In *Air Pollution Engineering Manual*, 2e, W.T. Davis, ed. John Wiley & Sons, Inc., New York, 664–681.

Grossi, C.M. and Brimblecombe, P. (2002) The effect of atmospheric pollution on building materials. *Journal de Physique Iv* 12, 197–210.

Hays, M.D., Smith, N.D., and Dong, Y.J. (2004) Nature of unresolved complex mixture in size-distributed emissions from residential wood combustion as measured by thermal desorption-gas chromatography-mass spectrometry. *J. Geophys. Res.* 109.

Henry, R.C. (1986) Fundamental limitations of receptor models using factor analysis. In *Transactions, Receptor Methods for Source Apportionment: Real World Issues and Applications*, T.G. Pace, ed. Air Pollution Control Association, Pittsburgh, PA, 68–77.

Henry, R.C. (1987) Current factor analysis receptor models are ill-posed. *Atmos. Environ.* 21, 1815–1820.

Henry, R.C., Wang, Y.J., and Gebhart, K.A. (1991) The relationship between empirical orthogonal functions and sources of air pollution. *Atmos. Environ.* 25A, 503–509.

Henry, R.C. (1997a) History and fundamentals of multivariate air quality receptor models. *Chemom. Intell. Lab. Sys.* 37, 37–42.

Henry, R.C. (1997b) Receptor model applied to patterns in space (RMAPS) Part I—Model description. *J. Air Waste Manage. Assoc.* 47, 216–219.

Henry, R.C. (1997c) Receptor model applied to patterns in space (RMAPS) Part II—Apportionment of airborne particulate sulfur from Project MOHAVE. *J. Air Waste Manage. Assoc.* 47, 220–225.

Henry, R.C. (2000) UNMIX Version 2 Manual. Ronald C. Henry, Ph.D., West Hills, CA.

Henry, R.C. (2002) Multivariate receptor models—Current practice and future trends. *Chemom. Intell. Lab. Sys.* 60, 43–48, doi:10.1016/S0169-7439(01)00184-8.

Henry, R.C. (2003) Multivariate receptor modeling by *N*-dimensional edge detection. *Chemom. Intell. Lab. Sys.* 65, 179–189, doi:10.1016/S0169-7439(02)00108-9.

Hering, S.V. (1995) Impactors, cyclones, and other inertial and gravitational collectors. In *Air Sampling Instruments—For Evaluation of Atmospheric Contaminants*, B.S. Cohen and S.V. Hering, eds. Am. Conf. of Governmental Industrial Hygienists, Cincinnati, OH, 279–321.

Hidy, G.M. and Friedlander, S.K. (1971) The nature of the Los Angeles aerosol. In *Proceedings of the Second International Clean Air Congress*, H.M. Englund and W.T. Beery, ed. Academic Press, New York, 391–404.

Hidy, G.M., Eatough, D.J., and Klouda, G.A. (2004) Design scenario for the radioisotopic estimation of the biogenic component of airborne particles. *J. Air Waste Manage. Assoc.* 54, 600–613.

Hies, T., Treffeisen, R., Sebald, L., and Reimer, E. (2000) Spectral analysis of air pollutants Part 1. Elemental carbon time series. *Atmos. Environ.* 34, 3495–3502.

Ho, S.S.H. and Yu, J.Z. (2004) In-injection port thermal desorption and subsequent gas chromatography-mass spectrometric analysis of polycyclic aromatic hydrocarbons and *n*-alkanes in atmospheric aerosol samples. *J. Chromatogr. A.* 1059, 121–129.

Hopke, P.K. (1981) The application of factor analysis to urban aerosol source resolution. In *Atmospheric Aerosol: Source/Air Quality Relationships*, E.S. Macias and P.K. Hopke, eds. American Chemical Society, Washington, DC, 21–49.

Hopke, P.K., Alpert, D.J., and Roscoe, B.A. (1983) FANTASIA—A program for target transformation factor analysis to apportion sources in environmental samples. *Computers & Chem.* 7, 149–155.

Hopke, P.K. (1985) *Receptor Modeling in Environmental Chemistry.* John Wiley & Sons, Inc., New York.

Hopke, P.K. (1988) Target transformation factor analysis as an aerosol mass apportionment method: A review and sensitivity study. *Atmos. Environ.* 22, 1777–1792.

Hopke, P.K. (1991) An introduction to receptor modeling. In *Receptor Modeling for Air Quality Management*, P.K. Hopke, ed. Elsevier, Amsterdam, The Netherlands, 1–10.

Hopke, P.K. and Casuccio, G.S. (1991) Scanning electron microscopy. In *Receptor Modeling for Air Quality Management*, P.K. Hopke, ed. Elsevier, Amsterdam, The Netherlands, 149–212.

Hopke, P.K. (1999) An introduction to source receptor modeling. In *Elemental Analysis of Airborne Particles*, S. Landsberger and M. Creatchman, eds. Gordon and Breach Science, Amsterdam, 273–315.

Hopke, P.K. (2003) Recent developments in receptor modeling. *J. Chemometrics* 17, 255–265, doi:10.1002/cem.796.

Hsu, K.J. (1997) Application of vector autoregressive time series analysis to aerosol studies. *Tellus* 49B, 327–342.

Jacobson, M.Z. (2002) Control of fossil-fuel particulate black carbon plus organic matter, possibly the most effective method of slowing global warming. *J. Geophys. Res.* 107, ACH 16-1-ACH 16–22, doi:10.1029/2001JD001376.

John, W., Wall, S.M., Ondo, J.L., and Winklmayr, W. (1990) Modes in the size distributions of atmospheric inorganic aerosol. *Atmos. Environ.* 24A, 2349–2359.

Jorquera, H., Palma, W., and Tapia, J. (2000) An intervention analysis of air quality data at Santiago, Chile. *Atmos. Environ.* 34, 4073–4084.

Kelley, D.W. and Nater, E.A. (2000) Historical sediment flux from three watersheds into Lake Pepin, Minnesota, USA. *J. Environ. Quality* 29, 561–568.

Keywood, M.D., Ayers, G.P., Gras, J.L., Gillett, R.W., and Cohen, D.D. (1999) Relationships between size segregated mass concentration data and ultrafine particle number concentrations in urban areas. *Atmos. Environ.* 33, 2907–2913.

Kim, B.M. and Henry, R.C. (1999) Extension of self-modeling curve resolution to mixtures of more than three components—Part II: Finding the complete solution. *Chemom. Intell. Lab. Sys.* 49, 67–77, doi:10.1016/S0169-7439(99)00029-5.

Kim, B.M. and Henry, R.C. (2000) Extension of self-modeling curve resolution to mixtures of more than three components—Part III: Atmospheric aerosol data simulation studies. *Chemom. Intell. Lab. Sys.* 52, 145–154, doi:10.1016/S0169-7439(00)00077-0.

Kim, E. and Hopke, P.K. (2004a) Improving source identification of fine particles in a rural northeastern US area utilizing temperature-resolved carbon fractions. *J. Geophys. Res.* 109, 1–13, doi:2003JD004199.

Kim, E. and Hopke, P.K. (2004b) Source apportionment of fine particles at Washington, DC, utilizing temperature-resolved carbon fractions. *J. Air Waste Manage. Assoc.* 54, 773–785.

Kleinman, M.T., Pasternack, B.S., Eisenbud, M., and Kneip, T.J. (1980) Identifying and estimating the relative importance of sources of airborne particulates. *Enivron. Sci. Technol.* 14, 62–65.

Lawson, D.R. and Winchester, J.W. (1979) A standard crustal aerosol as a reference for elemental enrichment factors. *Atmos. Environ.* 13, 925–930.

Lipfert, F.W. and Wyzga, R.E. (1997) Air pollution and mortality: The implications of uncertainties in regression modeling and exposure measurement. *J. Air Waste Manage. Assoc.* 47, 517–523.

Lowenthal, D.H. and Rahn, K.A. (1989) The relationship between secondary sulfate and primary regional signatures in northeastern aerosol and precipitation. *Atmos. Environ.* 23, 1511–1515.

Malm, W.C., Iyer, H.K., and Gebhart, K.A. (1990) Application of tracer mass balance regression to WHITEX data. In *Transactions, Visibility and Fine Particles*, C.V. Mathai, ed. Air & Waste Management Association, Pittsburgh, PA, 806–818.

Marple, V.A., Rubow, K.L., and Olson, B.A. (1993) Inertial, gravitational, centrifugal and thermal collection techniques. In *Aerosol Measurement: Principles, Techniques and Applications*, K. Willeke and P.A. Baron, eds. Van Nostrand Reinhold, New York, 206–232.

Mason, B. (1966) *Principles of Geochemistry*, 3e. John Wiley & Sons, Inc., New York, NY.

Mauderly, J.L., Neas, L.M., and Schlesinger, R.B. (1998) PM monitoring needs related to health effects. In *Atmospheric Observations: Helping Build the Scientific Basis for Decisions Related to Airborne Particulate Matter*, D.L. Albritton and D.S. Greenbaum, eds. Health Effects Research Institute, Cambridge, MA, 9–14.

Mazurek, M.A., Simoneit, B.R.T., Cass, G.R., and Gray, H.A. (1987) Quantitative high-resolution gas chromatography and high-resolution gas chromatography/mass spectrometry analyses of carbonaceous fine aerosol particles. *Int. J. Environ. Anal. Chem.* 29, 119–139.

McCrone, W.C. and Delly, J.G. (1973) *The Particle Atlas, Volume I: Principles and Techniques*. Ann Arbor Science Publishers, Inc., Ann Arbor, MI.

McDonald, J.D., Zielinska, B., Fujita, E.M., Sagebiel, J.C., Chow, J.C., and Watson, J.G. (2000) Fine particle and gaseous emission rates from residential wood combustion. *Environ. Sci. Technol.* 34, 2080–2091.

McMurry, P.H. (2000) A review of atmospheric aerosol measurements. *Atmos. Environ.* 34, 1959–1999.

Morandi, M.T., Lioy, P.J., and Daisey, J.M. (1991) Comparison of two multivariate modeling approaches for the source apportionment of inhalable particulate matter in Newark, NJ. *Atmos. Environ.* 25A, 927–937.

Newmann, L., Forrest, J., and Manowitz, B. (1975) The application of an isotopic ratio technique to a study of the atmospheric oxidation of sulphur dioxide in the plume from an oil-fired power plant. *Atmos. Environ.* 9, 969–977.

Nolte, C.G., Schauer, J.J., Cass, G.R., and Simoneit, B.R.T. (1999) Highly polar organic compounds present in meat smoke. *Environ. Sci. Technol.* 33, 3313–3316.

Oberdorster, G., Oberdorster, E., and Oberdorster, J. (2005) Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. *Environmental Health Perspectives* 113, 823–839.

Oros, D.R. and Simoneit, B.R.T. (2000) Identification and emission rates of molecular tracers in coal smoke particulate matter. *Fuel* 79, 515–536.

Paatero, P. and Tapper, U. (1994) Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5, 111–126.

Paatero, P. (1998) *User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3 Part 1: Tutorial*. University of Helsinki, Helsinki, Finland.

Paatero, P. (2000) *User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3 Part 2: Reference*. University of Helsinki, Helsinki, Finland.

Patterson, D.B., Farley, K.A., and Norman, M.D. (1999) (super 4) He as a tracer of continental dust: A 1.9 million year record of aeolian flux to the West Equatorial Pacific Ocean. *Geochimica et Cosmochimica Acta* 63, 615–625.

Pena-Mendez, E.M., Astorga-Espana, M.S., and Garca-Montelongo, F.J. (2001) Chemical fingerprinting applied to the evaluation of marine oil pollution in the coasts of Canary Islands (Spain). *Environ. Poll.* 111, 177–187.

Perez, P., Trier, A., and Reyes, J. (2000) Prediction of $PM_{2.5}$ concentrations several hours in advance using neural networks in Santiago, Chile. *Atmos. Environ.* 34, 1189–1196.

Perrier, V., Philipovitch, T., and Basdevant, C. (1995) Wavelet spectra compared to Fourier spectra. *J. Math. Phys.* 36, 1506–1519.

Perry, K.D., Cahill, T.A., Eldred, R.A., and Dutcher, D.D. (1997) Long-range transport of North African dust to the eastern United States. *J. Geophys. Res.* 102, 11225–11238.

Preining, O. (1998) The physical nature of very, very small particles and its impact on their behaviour. *J. Aerosol Sci.* 29, 481–495.

Rader, D.J. and O'Hern, T.J. (1993) Optical direct-reading techniques: *In situ s*ensing. In *Aerosol Measurement: Principles, Techniques and Applications*, K. Willeke and P.A. Baron, eds. Van Nostrand Reinhold, New York, NY, 345–380.

Rahn, K.A. (1976) Silicon and aluminum in atmospheric aerosols: Crust-air fractionation? *Atmos. Environ.* 10, 597–601.

Rashid, M. and Griffiths, R.F. (1993) Ambient K, S, and Si in fine and coarse aerosols of Kuala Lumpur, Malaysia. *J. Aerosol Sci.* 24, S5–S6.

Reich, S.L., Gomez, D.R., and Dawidowski, L.E. (1999) Artificial neural network for the identification of unknown air pollution sources. *Atmos. Environ.* 33, 3045–3052.

Richards, L.W., Lehrman, D.E., Weiss, R.E., Bush, D., Watson, J.G., McDade, C.E., and Magliano, K. (2001) Light scattering measurements during the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study. In *Proceedings, International Specialty Conference on Regional Haze and Global Radiation Balance—Aerosol Measurements and Models: Closure, Reconciliation and Evaluation*. Air & Waste Management Association, Pittsburgh, PA.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T. (1993a) Sources of fine organic aerosol—2. Noncatalyst and catalyst-equipped automobiles and heavy-duty diesel trucks. *Enivron. Sci. Technol.* 27, 636–651.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T. (1993b) Sources of fine organic aerosol—5. Natural gas home appliances. *Enivron. Sci. Technol.* 27, 2736–2744.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T. (1993c) Sources of fine organic aerosol—3. Road dust, tire debris, and organometallic brake lining dust: Roads as sources and sinks. *Enivron. Sci. Technol.* 27, 1892–1904.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T. (1997) Sources of fine organic aerosol—8. Boilers burning No. 2 distillate fuel oil. *Enivron. Sci. Technol.* 31, 2731–2737.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., and Simoneit, B.R.T. (1998) Sources of fine organic aerosol—9. Pine, oak, and synthetic log combustion in residential fireplaces. *Enivron. Sci. Technol.* 32, 13–22.

Schauer, J.J., Kleeman, M.J., Cass, G.R., and Simoneit, B.R.T. (1999a) Measurement of emissions from air pollution sources—2. C_1 through C_{30} organic compounds from medium duty diesel trucks. *Environ. Sci. Technol.* 33, 1578–1587.

Schauer, J.J., Kleeman, M.J., Cass, G.R., and Simoneit, B.R.T. (1999b) Measurement of emissions from air pollution sources—1. C_1 through C_{29} organic compounds from meat charbroiling. *Environ. Sci. Technol.* 33, 1566–1577.

Schauer, J.J. and Cass, G.R. (2000) Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. *Environ. Sci. Technol.* 34, 1821–1832.

Shively, T.S. and Sager, T.W. (1999) Semiparametric regression approach to adjusting for meteorological variables in air pollution trends. *Environ. Sci. Technol.* 33, 3873–3880.

Simoneit, B.R.T. (1999) A review of biomarker compounds as source indicators and tracers for air pollution. *Environ. Sci. Pollut. Res.* 6, 159–169.

Simoneit, B.R.T., Schauer, J.J., Nolte, C.G., Oros, D.R., Elias, V.O., Fraser, M.P. *et al.* (1999) Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. *Atmos. Environ.* 33, 173–182.

Somerville, M.C. and Evans, E.G. (1995) Effect of sampling frequency on trend detection for atmospheric fine mass. *Atmos. Environ.* 29, 2429–2438.

Song, X.H. and Hopke, P.K. (1996) Solving the chemical mass balance problem using an artificial neural network. *Enivron. Sci. Technol.* 30, 531–535.

Spix, C., Heinrich, J., Dockery, D.W., Schwartz, J., Völksch, G., Schwinkowski, K. *et al.* (1999) Air pollution and daily mortality in Erfurt, East Germany, from 1980–1989. *Environ. Health Perspect.* 101, 518–526.

Standley, L.J. and Simoneit, B.R.T. (1994) Resin diterpenoids as tracers for biomass combustion aerosols. *Atmos. Environ.* 28, 1–16.

Sturges, W.T. (1989) Identification of pollution sources of anomalously enriched elements. *Atmos. Environ.* 23, 2067–2072.

Thurston, G.D. and Spengler, J.D. (1985) A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmos. Environ.* 19, 9–25.

Turk, J.T., Campbell, D.H., and Spahr, N.E. (1993) Use of chemistry and stable sulfur isotopes to determine sources of trends in sulfate of Colorado lakes. *Water Air and Soil Pollution* 67, 415–431.

U.S. EPA (1971) National primary and secondary ambient air quality standards. Federal Register 36, 8186.

U.S. EPA (1997) National ambient air quality standards for particulate matter: Final rule. Federal Register 62, 38651–38760.

U.S. EPA (1999) 40 CFR Part 51—Regional haze regulations: Final rule. Federal Register 64, 35714–35774.

Vedal, S. (1997) Critical review—Ambient particles and health: Lines that divide. *J. Air Waste Manage. Assoc.* 47, 551–581.

Vega, E., García, I., Apam, D., Ruíz, M.E., and Barbiaux, M. (1997) Application of a chemical mass balance receptor model to respirable particulate matter in Mexico City. *J. Air Waste Manage. Assoc.* 47, 524–529.

Vickery, J. (2004) Conceptual models of PM for North American regions. In *Particulate Matter Science for Policy Makers—A NARSTO Assessment*, P. McMurry, M. Shepherd, and J. Vickery, eds. Cambridge University Press, Cambridge, UK, 355–413.

Watson, J.G. (1984) Overview of receptor model principles. *J. Air Poll. Control Assoc.* 34, 619–623.

Watson, J.G., Cooper, J.A., and Huntzicker, J.J. (1984) The effective variance weighting for least squares calculations applied to the mass balance receptor model. *Atmos. Environ.* 18, 1347–1355.

Watson, J.G., Chow, J.C., Richards, L.W., Andersen, S.R., Houck, J.E., and Dietrich, D.L. (1988) The 1987–88 Metro Denver Brown Cloud Air Pollution Study, Volume III: Data interpretation. Prepared for Greater Denver Chamber of Commerce, Denver, CO, Desert Research Institute, Reno, NV.

Watson, J.G., Robinson, N.F., Chow, J.C., Henry, R.C., Kim, B.M., Pace, T.G. *et al.* (1990) The USEPA/DRI chemical mass balance receptor model, CMB 7.0. *Environ. Software* 5, 38–49.

Watson, J.G., Chow, J.C., Lowenthal, D.H., Pritchett, L.C., Frazier, C.A., Neuroth, G.R., and Robbins, R. (1994) Differences in the carbon composition of source profiles for diesel- and gasoline-powered vehicles. *Atmos. Environ.* 28, 2493–2505.

Watson, J.G., Blumenthal, D.L., Chow, J.C., Cahill, C.F., Richards, L.W., Dietrich, D. *et al.* (1996) Mt. Zirkel Wilderness Area reasonable attribution study of visibility

impairment, Vol. II: Results of data analysis and modeling. Prepared for Colorado Department of Public Health and Environment, Denver, CO, Desert Research Institute, Reno, NV.

Watson, J.G., Robinson, N.F., Lewis, C.W., Coulter, C.T., Chow, J.C., Fujita, E.M. *et al.* (1997) Chemical mass balance receptor model version 8 (CMB) user's manual. Prepared for U.S. Environmental Protection Agency, Research Triangle Park, NC, Desert Research Institute, Reno, NV.

Watson, J.G. (1998) Implications of new suspended particle standards for the cement industry. In *Conference Record, 40th Cement Industry Technical Conference*, M.W. Jasberg, ed. Portland Cement Association, Skokie, IL, 331–342.

Watson, J.G., Robinson, N.F., Lewis, C.W., Coulter, C.T., Chow, J.C., Fujita, E.M. *et al.* (1998a) CMB8 applications and validation protocol for $PM_{2.5}$ and VOCs. Report No. 1808.2D1, Prepared for U.S. Environmental Protection Agency, Research Triangle Park, NC, Desert Research Institute, Reno, NV.

Watson, J.G., Chow, J.C., Moosmüller, H., Green, M.C., Frank, N.H., and Pitchford, M.L. (1998b) Guidance for using continuous monitors in $PM_{2.5}$ monitoring networks. Report No. EPA-454/R-98-012, Prepared for U.S. EPA, Research Triangle Park, NC, Desert Research Institute, Reno, NV.

Watson, J.G., Fujita, E.M., Chow, J.C., Zielinska, B., Richards, L.W., Neff, W.D., and Dietrich, D. (1998c) Northern Front Range Air Quality Study. Final report. Prepared for Colorado State University, Fort Collins, CO, Desert Research Institute, Reno, NV.

Watson, J.G., Coulombe, W.G., Egami, R.T., Casuccio, G.S., Hanneman, H.G., and Badger, S. (1999) Causes of particle deposition in Davenport, California. Report No. 99-0954.1F2, Prepared for RMC LONESTAR, Davenport, CA, Desert Research Institute, Reno, NV.

Watson, J.G., Chow, J.C., Bowen, J.L., Lowenthal, D.H., Hering, S., Ouchida, P., and Oslund, W. (2000a) Air quality measurements from the Fresno Supersite. *J. Air Waste Manage. Assoc.* 50, 1321–1334.

Watson, J.G., Chow, J.C., and Pace, T.G. (2000b) Fugitive dust emissions. In *Air Pollution Engineering Manual*, 2e, W.T. Davis, ed. John Wiley & Sons, Inc., New York, 117–135.

Watson, J.G. and Chow, J.C. (2001a) Ambient air sampling. In *Aerosol Measurement: Principles, Techniques, and Applications*, 2e, P. Baron and K. Willeke, eds. John Wiley & Sons, New York, NY, 821–844.

Watson, J.G. and Chow, J.C. (2001b) Estimating middle-, neighborhood-, and urban-scale contributions to elemental carbon in Mexico City with a rapid response aethalometer. *J. Air Waste Manage. Assoc.* 51, 1522–1528.

Watson, J.G. and Chow, J.C. (2001) Source characterization of major emission sources in the Imperial and Mexicali valleys along the U.S./Mexico boarder. *Sci. Total Environ*. 276, 33–47.

Watson, J.G. (2002) Visibility: Science and regulation. *J. Air Waste Manage. Assoc*. 52, 628–713.

Watson, J.G., Zhu, T., Chow, J.C., Engelbrecht, J.P., Fujita, E.M., and Wilson, W.E. (2002) Receptor modeling application framework for particle source apportionment. *Chemosphere* 49, 1093–1136.

Watson, J.G. and Chow, J.C. (2004) Receptor models for air quality management. EM 10, 27–36.

Watson, J.G. and Chow, J.C. (2005) Receptor models. In *Air Quality Modeling—Theories, Methodologies, Computational Techniques, and Available Databases and Software. Vol. II—Advanced Topics*, P. Zannetti, ed. Air and Waste Management Association and the EnviroComp Institute, Pittsburgh, PA, 455–501.

Watson, J.G., Chow, J.C., and Chen, L.-W.A. (2005) Summary of organic and elemental carbon/black carbon analysis methods and intercomparisons. *Aerosol and Air Quality Research* 5, 69–102.

Weatherhead, E.C., Reinsel, G.C., Tiao, G.C., Meng, X.L., Choi, D., Cheang, W.K. *et al.* (1998) Factors affecting the detection of trends: Statistical considerations and applications to environmental data. *J. Geophys. Res.* 103, 17149–17161.

Weiss, D., Shotyk, W., Appleby, P.G., Kramers, J.D., and Cheburkin, A.K. (1999) Atmospheric Pb deposition since the Industrial Revolution recorded by five Swiss peat profiles: Enrichment factors, fluxes, isotopic composition, and sources. *Environ. Sci. Technol.* 33, 1340–1352.

Whitby, K.T., Charlson, R.J., Wilson, W.E., and Stevens, R.K. (1974) The size of suspended particle matter in air. *Science* 183, 1098–1099.

White, W.H. (1986) On the theoretical and empirical basis for apportioning extinction by aerosols: A critical review. *Atmos. Environ.* 20, 1659–1672.

White, W.H. (1999) Phantom spatial factors: An example. *J. Air Waste Manage. Assoc.* 49, 345–349.

Wienke, D. and Hopke, P.K. (1994) Projection of Prim's minimal spanning tree into Kohonen's neural network for identification of airborne particle sources by their multielement trace patterns. *Anal. Chim. Acta.* 291, 1–18.

Wilson, W.E., Chow, J.C., Claiborn, C.S., Fusheng, W., Engelbrecht, J.P., and Watson, J.G. (2002) Monitoring of particulate matter outdoors. *Chemosphere* 49, 1009–1043.

Zannetti, P. (2003) *Air Quality Modeling: Theories, Methodologies, Computational Techniques, and Available Databases and Software—Vol. I—Fundamentals*. Air and Waste Management Association, Pittsburgh, PA.

Zannetti, P. (2005) *Air Quality Modeling: Theories, Methodologies, Computational Techniques, and Available Databases and Software—Vol. II—Advanced Topics*. Air and Waste Management Association, Pittsburgh, PA.

CHAPTER 9

CHEMICAL FINGERPRINTING METHODS

Gregory S. Douglas, Stephen D. Emsbo-Mattingly,
Scott A. Stout, Allen D. Uhler, and Kevin J. McCarthy

9.1 Introduction 312
9.2 Overview of Chemical Fingerprinting Methodology 315
9.2.1 EPA Reference Methods 317
9.2.2 Technical Challenges 319
9.2.3 Method Specialization 320
9.2.4 Tiered Approach 322
9.3 Quality Assurance and Quality Control 322
9.3.1 Quality Assurance 325
9.3.2 Quality Control 326
9.3.2.1 QC Parameters for Sample Preparation 326
9.3.2.2 QC Parameters for Analytical Procedures 331
9.4 Volatile Hydrocarbon Fingerprinting Methods 337
9.4.1 Volatile Hydrocarbons in Air and Vapor by GC/MS 340
9.4.2 Collection of Solid and Liquid Samples for Volatile Hydrocarbon Analysis 345
9.4.2.1 Soil Samples 347
9.4.2.2 Water Samples 348
9.4.2.3 NAPL Samples 348
9.4.3 PIANO and Nonhydrocarbon Analysis by Purge-and-Trap GC/MS 348
9.4.3.1 Standard Preparation 349
9.4.3.2 Sample Preparation 349
9.4.3.3 Sample Analysis 351
9.4.4 PIANO and Nonhydrocabon Analysis by Direct Injection GC/MS 353
9.4.4.1 Standard Preparation 355
9.4.4.2 Sample Preparation 355
9.4.4.3 Sample Analysis 355
9.4.5 Alkyl Lead Gasoline Additives by Direct Injection GC/MS 357
9.4.6 Diamondoid Analysis by Direct Injection GC/MS 360

9.4.7 Oxygenate Analysis by Direct Injection GC/MS in NAPL 362
9.4.8 Analysis of Other Polar Fuel Additives 369
9.4.9 Full Range Whole Oil Analysis 369
9.5 Semivolatile Hydrocarbon Fingerprinting Methods 371
9.5.1 Sample Preparation 372
9.5.1.1 Sample Extractions 372
9.5.1.2 Extract Cleanup and Purification 376
9.5.2 Sample Analysis 379
9.5.2.1 High Resolution Hydrocarbon Fingerprinting by GC/FID 379
9.5.2.2 Polycyclic Aromatic Hydrocarbons by GC/MS/SIM 389
9.5.2.3 Saturates and Biomarkers by GC/MS/SIM 402
9.5.3 Ancillary Hydrocarbon Fingerprinting Techniques 413
9.5.3.1 Conventional Simulated Distillation 413
9.5.3.2 High Temperature Simulated Distillation 414
9.5.3.3 Stable Isotopes of Bulk Hydrocarbons 415
9.5.3.4 Compound-Specific Stable Isotope Measurements 418
9.5.3.5 Two-Dimensional Gas Chromatography 419
9.5.3.6 Electrospray Ionization Mass Spectrometry (ESI-MS) 420
9.6 Polychlorinated Biphenyl Fingerprinting Methods 421
9.6.1 Current Practices 427
9.6.2 PCB Forensic Methods 430
9.6.2.1 Sample Preparation 431
9.6.2.2 Extract Cleanup 431
9.6.2.3 Internal Standards (EPA Method 8000B) 433
9.6.2.4 PCB Analysis by GC/ECD (EPA Method 8082) 433
9.6.2.5 PCB Analysis by GC/MS (EPA Method 8270C/680) 435
9.6.2.6 PCB Analysis by GC/HRMS (EPA Method 1668A) 436
References 439

9.1 INTRODUCTION

This chapter deals with chemical fingerprinting methods and applications for hydrocarbons and polychlorinated biphenyls (PCBs). These contaminants have been selected because the most frequently identified organic chemicals

of concern at contaminated sites are those found in petroleum, coal, thermally decomposed hydrocarbons, and PCBs. Careful chemical analysis of these contaminants, whether found as nonaqueous phase liquids (NAPLs), or as constituents of soils, sediments, ground water, surface water, atmospheric particulates, vapors, or biological tissues, can yield data useful for describing the detailed nature and source of the contamination. Armed with this chemical fingerprinting data, and drawing upon the available historic (both regulatory and operational), geologic, hydrologic, or meteorologic data, the environmental forensic investigator is in a strong position to determine the origin(s) or source(s) of the contamination and distinguish contamination from naturally occurring or anthropogenic background chemicals. In some instances, this same information may be helpful in constraining the most likely duration of time that has passed since the contaminants were released into the environment. Thus, chemical fingerprinting analysis as supported by the analytical methods discussed in this chapter often can answer the most often asked questions in environmental forensic investigations—what is the contamination, where did it come from, and when was it released to the environment?

High quality analytical chemistry data is the cornerstone of chemical fingerprinting and a key component of most environmental forensic investigations. As such the availability or selection of the appropriate data or measurement methodologies often determines the degree to which environmental forensic questions can be answered. With few exceptions, the data requirements for a forensic investigation differ from those of regulatory driven studies. In regulatory driven assessments (e.g., Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA) or Resource Conservation and Recovery Act of 1976 (RCRA) type investigations), only a limited number of parameters and chemicals of concern are measured to determine the nature and extent of contamination (Uhler *et al.*, 1998–1999). Examples of these types of measurements include total petroleum hydrocarbons (TPH), concentrations of water-soluble benzene, toluene, ethyl-benzene, and *o*-, *m*-, and *p*-xylenes (BTEX) using EPA Method 8260B, the 16 Priority Pollutant polycyclic aromatic hydrocarbons (PAHs) by EPA Method 8270C, or total Aroclors by EPA Method 8082. As described later, these compliance-driven measurements—although adequate for gross descriptions of the extent and types of contaminants found at a site (e.g., TPH or BTEX above or below a regulatory action limit)—are largely insufficient to address the questions fundamental to an environmental forensics investigation; that is, a detailed understanding of the chemical nature, source, and age of site contamination. Achieving robust answers to these questions requires that the appropriate chemical fingerprinting data, beyond that typically obtained with standard

regulatory chemical analyses, be developed using appropriate and technically defensible methodologies (Douglas and Uhler, 1993). To obtain this type of data, environmental forensic chemists have expanded or otherwise modified basic EPA methods of analyses, or developed new methods to more thoroughly characterize the nature of site contamination. Most of these modified or new methods have been published in the peer-reviewed scientific literature, authoritative books, or in promulgated treatises (i.e., in the United States Government's Federal Register). It is our intention in this chapter to describe the analytical methods most commonly used in environmental forensics investigations involving hydrocarbons and PCBs.

The inclusion of vapor and air forensic chemistry methods in this chapter is a notable addition to the first edition (Stout *et al.*, 2002) of *Introduction to Environmental Forensics* (Murphy and Morrison, 2002). Vapor intrusion and indoor air contamination is a rapidly evolving technical issue that has been receiving substantial attention during the past decade. Specifically, subsurface contamination of groundwater and soil by petroleum and nonpetroleum materials containing volatile hydrocarbons (e.g., automotive gasolines and manufactured gas plant residues, respectively) occurs commonly throughout the United States. During the past five to 10 years, there has been significant interest in the potential transport of subsurface vapors emanating from volatile subsurface contamination into residential homes and industrial structures. Central to this concern is the recognition that volatile organic compounds such as those that comprise automotive gasoline can partition from water, soil, or nonaqueous phase liquid (NAPL) into the subsurface gas phase and then migrate to the ground surface and into buildings. Thus, determining the potential for, occurrence of, and source of subsurface vapors and indoor air is an emerging area of environmental forensics.

In this chapter we also discuss the technical issues associated with the analysis of PCBs in the environment and how to generate the type of data required to support a forensic investigation. PCBs are a class of man-made chlorinated hydrocarbons that were produced commercially in the United States from the late 1920s to 1970s. PCBs are not a single chemical, but a mixture of 209 possible chlorinated isomers of biphenyl. In the United States, PCBs were produced as nine commercial formulations under the trade name *Aroclor*, each distinguished by a different weight percentage of chlorine in the formulation. Aroclors had myriad uses, including dielectric fluids, heat transfer fluids, plasticizers, adhesives, dedusting agents, inks, cutting oils, and pesticide extenders (Durfee, 1976). Many of the individual PCB isomers or congeners that comprise Aroclors have long environmental half-lives (i.e., decades or more); thus, these chemicals remain important environmental contaminants of concern, even long after their industrial production was

halted. Our contribution lies in the description of various methods currently used for the analysis of PCBs in environmental samples and the selection of the most appropriate methods to support environmental forensic studies.

In this second edition of *Introduction to Environmental Forensics*, we provide a detailed survey of the analytical methods most commonly used in the production and interpretation of chemical fingerprinting data for the detailed characterization of hydrocarbon and PCB contamination. As such, we present a review of the types of analytical methods needed to adequately fingerprint hydrocarbon and PCB contamination in environmental media. We also include recommendations for quality control (QC) and quality assurance (QA) for these chemical measurements, developed in light of the rigors to which these data often are evaluated in the course of potential or existing litigation. In the later sections of this chapter we present examples of methods by which chemical fingerprinting data are interpreted and used in environmental forensics investigations. Most of the interpretative guidelines surrounding petroleum product refining, PAH fingerprinting, MGP operations, environmental weathering, and considerations of "background" that were discussed in the first edition of *Introduction to Environmental Forensics* (Stout *et al.*, 2002) have been removed and we refer you to the first edition for a detailed review of these issues. In keeping with the theme of methods, herein we offer a review of the analytical methods most often utilized in environmental forensic studies.

9.2 OVERVIEW OF CHEMICAL FINGERPRINTING METHODOLOGY

Chromatography lies at the heart of most chemical fingerprinting methods used in the characterization of hydrocarbons and PCBs in the environment. The heavy reliance on this technique is a function of its extraordinary ability to separate complex mixtures of organic compounds and measure isolated analytes using a wide range of specialized detectors. It must be remembered that petroleum- and coal-derived materials contain thousands of compounds that range in concentration over several orders of magnitude. Similarly, PCB Aroclors are complex mixtures of over 100 individual chlorinated aromatic compounds with a maximum of 209 possible isomers. Specialized high resolution chromatography methods are required to measure hydrocarbon and PCB constituents with demonstrable accuracy and precision.

The use of chromatographic signatures to identify and classify natural materials began in the early 1900s when the Russian scientist Mikhail Tswett demonstrated compositional distinctions among plant pigments (Tswett, 1906). Instrumentation scientists earned Nobel prizes and other significant

awards over the following decades for the development of new detectors, separation columns, and instrumentation components that culminated in 1954 with the development of the gas chromatograph (GC) (Jennings, 1987). The following decade, the scientific community experienced an explosion of new GC detectors, including the flame ionization detector (FID) (Issac, 2001) and mass spectrometer (MS) (McLafferty, 1993). Each analytical advancement was closely followed by discoveries in an ever-widening range of scientific disciplines—including the characterization of oil and related products.

In the mid to late 1960s, petroleum geochemists identified hundreds of compounds in crude oil with varying degrees of stability in the environment and clear associations with precursor compounds that originated from living organisms (Eglinton *et al.*, 1964). Having survived the forces of nature over geological timescales, the composition and relative abundances of numerous diagnostic compounds in crude oil (i.e., their "fingerprints") allowed geochemists to improve oil exploration and production success (Peters *et al.*, 2005).

In the 1970s and 1980s, environmental scientists began to commonly use GC techniques in a more limited capacity to identify and quantify the distribution of potentially hazardous chemicals (not just hydrocarbons) in the environment. Over this time period improvements in the nature of the chromatographic columns—particularly the transition from packed columns to capillary columns—tremendously improved the resolution of compounds within complex mixtures (Jennings, 1980, 1987; Lautamo *et al.*, 1998). The U.S. Environmental Protection Agency (EPA) and other regulatory bodies recognized the reliable use of capillary GC methods during the promulgation of compliance monitoring methods that became codified from 1980 to the present (EPA, 1980; EPA, 2004; EPA, 2005). As noted earlier, over the past 15 to 20 years these regulatory compliant and codifed methods have been modified in order to generate the chemical specificity that permits forensic chemists to identify the source of fugitive hydrocarbons and PCBs in the environment.

In recent decades, substantial technological innovations advanced the science of analytical chemistry. Instrument manufacturers developed new, remarkably sensitive and selective detectors (e.g., GC/MS; Buddle and Eichelberger, 1979), which when combined with new higher resolution chromatography methods improved the speed and accuracy required to identify a single organic compound in a complex matrix. For example, modern quadrupole GC/MS instruments can be operated in two primary modes. The scanning mode (Scan), used for the measurement of organic compounds, involves the sequential acquisition of ion masses between m/z 35 and m/z 500 (or 465 masses scanned) every one second or less, using 70 electron volts

(nominal) ionization in the positive electron impact ionization mode (EPA, 1997). The scanning cycle repeats for the duration of the instrument run (e.g., 60 minutes), generating approximately 3600 scans per ion per run. The GC is operated such that at least 10 scans are collected per peak (e.g., the minimum peak width would be approximately 10 seconds long for a MS that generates 1 scan per second) (Hites, 1997). Recent versions of the GC/MS detector have increased the number of scans per peak and exhibit greater sensitivity.

In addition, instrumental analysts developed methods for boosting MS sensitivity still further by running the instrument in selected ion monitoring (SIM) mode. The SIM method modifies the Scan approach by reducing the number of masses from 465 to less than about 30, which translates into faster cycle times and more than an order of magnitude increase in sensitivity (minimally 465/30 = 15.5 times faster scan rate). In practice, the actual number of SIM ions can range between 1 and 30 on most instruments. The dramatic increase in sensitivity over Scan greatly favors the use of SIM when low level analyte detections are required (e.g., Sauer and Boehm, 1991; Federal Register, 1994; Douglas *et al.*, 2004). Its use is supported by the inclusion of secondary confirmation ions and established relative retention times for known chemical patterns. Standard methods like EPA Method 8270C endorse the use of SIM analysis in this way when the objectives of the measurement require detection limits below the concentrations achievable in Scan mode (EPA, 1997a). The flexibility written into these EPA standard performance-based methods has expedited the development and acceptance of the primary analytical methods (e.g., GC/MS SIM) currently utilized by forensic investigators for identifying the type, source, and degree of weathering in environmental samples.

9.2.1 EPA REFERENCE METHODS

In the last 20 years, chromatographic techniques—particularly gas chromatography using simple flame ionization detectors (GC/FID) and as sophisticated as mass spectrometers (GC/MS)—have evolved for the measurement of volatile and extractable (e.g., semivolatile) hydrocarbon compounds in water, soil, sediment, tissue, oil, air, and other media. In the regulatory environment, the U.S. EPA has developed and codified a series of gas chromatographic techniques to measure industrial chemicals of concern using GC methods. Those most frequently cited are the EPA's so-called SW-846 Methods (EPA, 1997). The standard EPA methods of greatest interest to the environmental chemist include:

- *Compendium Method TO-15, Determination of Volatile Organic Compounds (VOCs) in Air Collected in Specially-Prepared Canisters and Analyzed by Gas Chromatography/Mass*

Spectrometry (EPA, 1999). This GC/MS technique measures an extensive list of potential indoor air pollutants, including a wide range of halogenated organics and a limited number of hydrocarbons, in air and vapor samples.

- *Method 8015B, Non-Halogenated Organics Using GC/FID (EPA, 1997; SW-846)*. This method is used to measure, among other things, total petroleum hydrocarbons (TPH), particularly TPH as diesel range organics (TPH_d). This gas chromatography with flame ionization detection (GC/FID) technique often is used for gross petroleum product identification.
- *Method 8260B, Volatile Organic Compounds by Gas Chromatography/Mass Spectrometry (EPA, 1997a; SW-846)*. This purge-and-trap GC/MS method is designed to measure a limited number of volatile-range hydrocarbons including the petroleum-related BTEX; TPH as gasoline (TPH_g) can be determined using this technique.
- *Method 8270C, Semivolatile Organic Compounds by Gas Chromatography/Mass Spectrometry (EPA, 1997a; SW-846)*. This powerful GC/MS method was designed as a multiresidue analytical technique for the quantification of a broad range of industrial chemicals. The target compound list for the standard method includes over 100 well-known chemicals of potential concern. Of interest to the forensic chemist is the inclusion of the 16 so-called Priority Pollutant PAHs and confirmation of PCB Aroclor mixtures.
- *Method 8310, Polynuclear Aromatic Hydrocarbons (EPA, 1997a; SW-846*). This high-performance liquid chromatography (HPLC) method was designed for the low-level measurement of the 16 EPA Priority Pollutant PAH. The method is used less frequently than Method 8270C, because of the greater selectivity of the mass spectrometric method.
- *Method 8082, Polychlorinated Biphenyls (PCBs) by Gas Chromatography (EPA, 1997a; SW-846)*. This method is likely the most common analytical technique used for environmental investigations of PCB contamination. This method describes a procedure for the analysis of sample extracts on a gas chromatograph equipped with an electron capture detector (GC/ECD) or an electrolytic conductivity detector (GC/ELCD). Flexibility exists within the method to measure PCBs as Aroclors (using about 5 peaks to quantify each Aroclor) or as representative congeners (minimum of 19 representing the more abundant Aroclor constituents). The high potential for matrix interferences with these detectors necessitates the confirmation of detected PCBs on a second capillary column with a different stationary phase (dual column) or confirmation by analysis on a GC/MS instrument.
- *Method 680, Determination of Pesticides and PCBs in Water and Soil/Sediment by Gas Chromatography/Mass Spectrometry (EPA, 1985)*. This method was designed to measure pesticides, PCB congeners, and PCB homologs using GC/MS/SIM calibrated with at least nine PCB congeners. The analyte identification is

determined by the quality of retention time and mass spectra match between the sample and calibration standards.

- *Method 1668A, Chlorinated Biphenyl Congeners in Water, Soil, Sediment, and Tissue by HRGC/HRMS (EPA, 1999b).* This method describes the congener-specific determination of the PCB congeners designated as toxic by the World Health Organization, with capability to measure more than 150 additional congeners. The toxic PCBs and the beginning and ending level-of-chlorination PCBs are determined by isotope dilution high resolution gas chromatography/high resolution mass spectrometry (HRGC/HRMS). The remaining PCB congener concentrations are measured using an internal standard quantitation technique.

In addition to these standard chromatographic techniques, there have been two historic wet chemical methods used for estimating total hydrocarbons in water and soil/sediment that the environmental forensics investigator might encounter: *EPA Method 413.1, Total Recoverable Oil and Grease* and *EPA Method 418.1, Total Recoverable Petroleum Hydrocarbons by Infra-Red Spectroscopy* (EPA, 1983). In May 1999, these two methods were effectively replaced by *EPA Method 1664, N-Hexane Extractable Material (HEM; Oil and Grease) and Silica Gel Treated N-Hexane Extractable Material (SGT-HEM; Non-polar Material) by Extraction and Gravimetry* (EPA, 1999a). In these simple methods, EPA uses the term "total petroleum hydrocarbons'' (TPH) to designate the substances that remain after the *n*-hexane extractable material is exposed to silica gel. These screening methods rarely, if ever, are used in a detailed forensic site assessment.

Many of these analytical methods provide the framework for the forensic methods that will be discussed later in this chapter. The technical challenges associated with these methods are primarily due to the fact that the goals of the EPA (regulatory) and the forensic chemist (source identification) are usually different (Douglas and Uhler, 1993). As a result, the target compounds of interest and the sensitivity required to measure those compounds in the environment are often at issue. As discussed next, these issues are easily resolved through the incorporation of important technical changes to the EPA methods to produce highly sensitive and compound-specific forensic tools.

9.2.2 TECHNICAL CHALLENGES

The preponderance of hydrocarbons in crude oil, refined petroleum products, and coal-derived materials are principally nonpolar volatile and semivolatile compounds. From a practical standpoint, petroleum- and coal-derived distillates and wastes are ideally suited for detailed chemical characterization using modern high resolution capillary gas chromatography (Nordtest, 1991; Douglas and Uhler, 1993; Boehm *et al.*, 1998; Wang *et al.*, 1999). The standard

EPA chromatographic methods described in this chapter are excellent base techniques for the analysis of hydrocarbon residues and PCBs in environmental media.

The technical challenges faced by environmental forensics investigators using standard EPA methods include the absence of appropriate target analytes for forensic studies, relatively high detection limits, and requirements for the generic operation of GC instruments (Douglas and Uhler, 1993; Stout *et al.*, 2002). Fortunately, the EPA reference methods from SW-846 are classified as performance-based measurement systems (PBMS), which encourage them to be adapted for optimal service to the project objectives so long as established performance criteria are satisfied and demonstrated (EPA, 2001). Section 2.1 of SW-846 (EPA, 1997) states, "if an alternative analytical procedure is employed, then EPA expects the laboratory to demonstrate and document that the procedure is capable of providing appropriate performance for its intended application. This demonstration must not be performed after the fact." Although, strictly for business purposes, most commercial laboratories are not interested in modifying standard "production line" chemistry methods, the necessity for measuring different suites of target analytes at low detection limits requires forensic laboratories to alter standard methods to meet project goals with the understanding that the standard method guidelines will be observed to the maximum extent practicable. This means that the modified EPA methods described in this chapter can be used to support both forensic investigations and regulatory monitoring.

9.2.3 METHOD SPECIALIZATION

Chemists have invested many years developing forensic methods and adapting standard EPA methods to yield the necessary data to support detailed forensics investigations (e.g., Overton *et al.*, 1981; Sauer and Boehm, 1991; Douglas and Uhler, 1994; Wang *et al.*, 1999; Emsbo-Mattingy *et al.*, 2001a, 2001b; Stout *et al.*, 2002). The *first* technical challenge for forensic methods is development of the target analyte list. The appropriate target analytes for the analysis of petroleum, coal, and PCB Aroclors include numerous compounds that are not listed in EPA standard methods. That said, the analytes used in forensic investigations (e.g., hydrocarbons and PCB congeners) behave in a very similar fashion to many of the regulated compounds listed in the EPA reference methods; therefore, the standard EPA reference methods are suitable as the starting point for more sophisticated environmental forensic methods. The procedures described in this chapter detail the specific compounds and fingerprint distributions that have been used successfully in past forensic studies to identify and differentiate hydrocarbon and PCB sources.

The *second* technical challenge is method sensitivity and analyte detection limits. This ability to detect low concentrations of target analytes is critical for materials composed of complex mixtures, where analytes may range in concentration over several orders of magnitude. The requirement for low detection limits and a wide dynamic measurement range is discussed in the individual method sections of this chapter. Essentially, these methods optimize the fit between the sample preparation, cleanup, and instrumentation methods. For sample preparation, the method must accurately and aggressively extract target analytes from the environmental sample without dramatically altering the chemical signature of the sample. For the cleanup of semivolatile extracts, interfering material should be removed to the maximum extent practicable. For instrumentation, the optimal technique maximizes the resolution of target analytes using careful GC separation strategies. In addition, a range of instrument detectors (e.g., FID, MS/Scan, MS/SIM, ECD, and HRMS) are employed to most accurately determine analyte concentrations with low limits of detection.

The *third* technical challenge is the development of routine procedures for assuring the generation of high quality data over time. The methods used to measure forensic analytes must be associated with an active system for monitoring and demonstrating that the target analytes are measured with known levels of accuracy and precision. Section 9.3 provides specific details about this method monitoring and corrective action system.

Collectively, the measurement of appropriate analytes using optimized procedures under an established quality control and assurance program satisfies the PBMS guidelines for determining the acceptability of the forensic methods described in this chapter (EPA, 2001). It is for this reason that the forensic methods described herein are able to produce high quality data that withstands scientific peer review. It is also for this reason that environmental agencies (e.g., EPA, NOAA, and states) have adopted many of these methods to resolve source identification questions of their own (EPA, 1993; Federal Register, 1994; NOAA, 2002). Finally, the use of the forensic methods discussed in this chapter combined with a robust quality assurance system permits the submission of forensic data for risk assessment and regulator compliance monitoring when the forensic analytes are regulated compounds of potential concern (e.g., benzene, toluene, ethyl benzene, PAHs, PCB congeners, and others). For example, a common laboratory approach for dealing with oiled sediment samples is to heavily dilute the final sample extract resulting in mostly nondetect results for the chemicals of concern. Therefore, the environmental scientist is faced with estimating environmental risk from some fraction of the high method detection limit (e.g., half of the method detection limit or MDL). The forensic methods discussed in the following sections are

designed to provide sensitive detection limits even within highly contaminated samples, therefore providing a more useful data set to the risk assessment professional. The following sections provide detailed guidance concerning the successful development and application of these modified analytical tools to support environmental forensics programs.

9.2.4 TIERED APPROACH

The quantification of hydrocarbons in the environment, whether as free phase products or as contaminants dispersed in soil, sediment, water, tissue, vapor, or air, can be carried out in a tiered fashion (Uhler *et al.*, 1998–1999; Wang *et al.*, 1999). The objective of a tiered analytical approach is to gather sufficient detail concerning the composition of the contaminant(s) under investigation so that the forensic questions of concern pertinent to the site under study can be answered. An example of a tiered approach, summarized in Figure 9.1, gives the environmental forensic investigator the flexibility to gather as little, or as much, information as necessary to address site- or incident-specific questions about the nature and extent, and ultimately source(s), of hydrocarbon contamination. The utility of the various hydrocarbon forensic methods identified in the tiered approach is summarized in Table 9.1.

The application of the tiered approach shown in Figure 9.1 is really a division of methods based upon the analytical technique(s) employed. These techniques primarily rely upon the volatility (or boiling point) of the hydrocarbons of concern and on their relative abundance in the mixture—with lower abundances requiring more sensitive analytical techniques. Starting in Section 9.4, the analytical methods most commonly used in the tiered characterization of hydrocarbon contamination and hydrocarbon sources are described.

9.3 QUALITY ASSURANCE AND QUALITY CONTROL

Overarching any analytical chemistry investigation is an appropriate quality assurance (QA) program that employs an array of quality control (QC) parameters to document the overall reliability of the laboratory data. Like any high quality laboratory work, analytical chemistry measurements in support of environmental forensics investigations must be conducted within the framework of a defensible, robust QA/QC program (Maney and Wait, 1991; Wait and Douglas, 1995). Beyond the obvious, the practical reason a laboratory should operate under a well-defined QA/QC program is the need for the confidence in the quality of the data because the data interpretation could be used to determine who may be responsible for the contamination. An

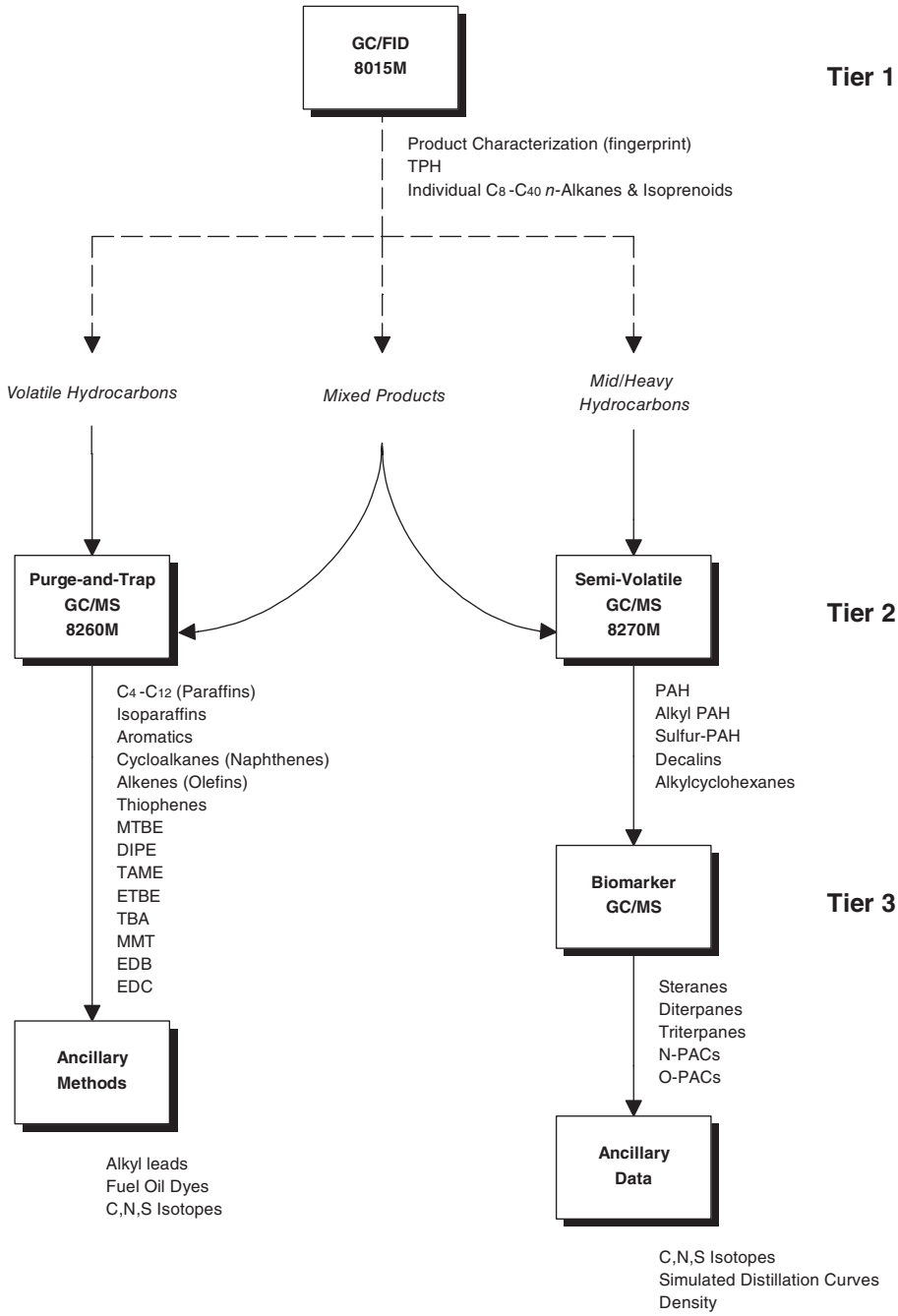


Figure 9.1
Tiered analytical approach to hydrocarbon fingerprinting.

Table 9.1

Useful analytical methods to support environmental forensics hydrocarbon investigations.

Measurement Method	Target Compounds	Utility for Environmental Forensics
Modified EPA Method 8015 Measurement of total petroleum hydrocarbon (TPH) product identification screen and measurement of individual saturated alkanes by high resolution gas chromatography with flame ionization detection (GC/FID)	■ Total extractable hydrocarbons ■ Total petroleum hydrocarbons ■ C_8 to C_{44} normal and branched-chain hydrocarbons ■ High resolution product fingerprints	■ Accurate determination of total extractable hydrocarbons and petroleum in the light distillate to residual petroleum product range ■ Development of diagnostic indices (e.g., pristine, phytane) ■ Accurate product identification in the light distillate to residual petroleum product range ■ Evaluation of weathering state of petroleum product(s) identification in the light distillate to residual petroleum product range
Modified EPA Method 8260B Measurement of a broad range of volatile petroleum hydrocarbons by purge-and-trap and direct injection GC/MS	■ Measurement of C_5 to C_{12} hydrocarbons and related chemicals ■ Paraffins ■ Isoparaffins ■ Aromatics ■ Naphthenes ■ Olefins ■ Oxygenates ■ Selected Additives	■ Product identification and differentiation in the light distillate range ■ Evaluation of weathering ■ Evaporation ■ Water-washing ■ Biodegradation ■ Presence/absence of important automotive gasoline blending agents and additives
Modified EPA 8270C Measurement cyclic hydrocarbons, polycyclic aromatic hydrocarbons (PAH), heteroatomic PAH, and biomarker compounds by high resolution GC/MS by selected ion monitoring (SIM)	■ Measurement of petroleum, coal, and combustion-derived PAH ■ Alkyl homologs of parent PAH ■ S-, N- containing PAH ■ Sterane, triterpane and other selected biomarkers	■ Detailed chemical indices used for evaluation of ■ Product identification and differentiation ■ Product mixing and allocation ■ Weathering ■ Evaporation ■ Water washing ■ Biodegradation ■ Petroleum vs. coal vs. combustion sources
Modified EPA 8270C High resolution GC/MS fingerprint of nonaqueous phase liquid (NAPL) in the C_4 to C_{44} hydrocarbon range	■ Broad boiling point range GC/MS analysis of petroleum, coal-derived, or other NAPL	■ "Snapshot" view of entire environmentally important hydrocarbon boiling point range ■ Allows investigator to visualize mixing of different hydrocarbon products ■ GC/MS allows for the identification of specific peaks in the chromatogram
Ancillary Measurements	■ C, H, S stable isotopes ■ Lead speciation by GC/MS ■ Fuel dyes ■ Density, viscosity ■ New methods (GCIRMS, GC × GC, etc.) ■ Simulated distillation ■ High temperature simulated distillation	■ Source/manufacturer differentiation ■ Age/production period classification ■ Data for fate and transport modeling ■ Useful data for mixing models

environmental forensics investigator must be aware that undocumented errors, omissions, or deviations are all potential fodder for undermining the otherwise defensible quality of an environmental data set. The best defense against this kind of criticism is to ensure that the environmental forensic laboratory operates under strict QA/QC guidelines.

9.3.1 QUALITY ASSURANCE

Quality assurance (QA) refers broadly to an integrated system of sample preparation procedures, monitoring programs, and corrective actions that guarantee compliance with established quality controls. A QA program assures laboratory management and project investigators that documented standards for the quality of facilities, equipment, personnel training, and work performance are attained or corrected quickly. EPA advocates a Quality System to manage the quality of environmental data collection, generation, and use, and is an excellent foundation upon which to build a QA/QC program (EPA, 2002a).

A laboratory's QA program should be described in the laboratory's Quality Management Plan (QMP). The QMP should describe the laboratory's policy for management system reviews, quality control parameters, data quality objectives, QA project plans, standard operating procedures, training, procurement of items and services, documentation, computer hardware and software, planning and implementation of project work, assessment and response, corrective action and continuous improvement. A QA system should consist of a minimum of six components, namely;

- A Project Plan that describes all work and quality control activities associated with a project.
- Accurate standard operating procedures (SOPs) that describe all technical activities conducted by the laboratory.
- A training program to ensure and document that all project personnel are qualified to perform project activities before independent activities begin. Personnel training records should be maintained by the QA Unit and include records of qualifications, prior experience, professional training, and internal training procedures.
- A record keeping system that facilitates sample and data tracking.
- A quality assessment program for all projects, conducted through management system reviews, technical system audits, performance evaluation samples, data validation, laboratory inspections, and independent data audits. An independent QA Unit within a laboratory should conduct the latter two activities.
- A continuous improvement program, facilitated through quality assurance audits, a formal corrective action program, and routine performance assessments.

In addition, all laboratory deliverables should receive an independent review that includes a QA and technical component. The QA review should ensure that the data are (1) complete, (2) in compliance with the procedures defined in the WP or QAPP, (3) in compliance with regulatory statutes, and (4) technically sound. The technical review should ensure that the results and interpretations are (1) objective, (2) defensible among peers, and (3) presented in a manner that conveys the conclusions clearly.

9.3.2 QUALITY CONTROL

Quality control refers to discrete activities, measurements, and acceptance criteria that ensure the production of high quality data. The quality of analytical chemistry data is an essential component of any environmental forensic investigation. Because formal protocols for the assessment of forensic data do not exist, forensic laboratories employ different degrees of quality control in the course of sample collection, analysis, and reporting. The practicalities of running a commercial laboratory often compete with QC compliance of the data. Therefore, minimum QC expectations should be established for all forensic data. This section describes the QC requirements for generating data of known quality based on EPA protocols (EPA, 1993a, 1995, 1997, 2002a).

9.3.2.1 QC Parameters for Sample Preparation

Following EPA (1997) guidance, samples are grouped for analysis in batches of 20 or fewer samples. This batch size is practical in most forensic laboratory settings and assures that QC performance is monitored with at least a 5% field sample frequency. The field sample batch is then accompanied by a suite of QC samples through the sample preparation, cleanup, and analysis steps. After data acquisition, these batch QC samples demonstrate the level of accuracy and precision associated with the applicable field samples.

The minimal sample preparation QC parameters should include:

- One procedure blank
- One laboratory control sample (e.g., blank spike)
- One field sample duplicate
- Surrogate spikes in field and QC samples
- Holding times

The procedure blank (PB) identifies any sources of contamination associated with the sample preparation, analysis, and storage in the laboratory. It demonstrates that the laboratory results represent the analyte concentrations

in the sample without bias, and can be used to temper data interpretation in those cases where low level contamination is identified in the PB. The PB sample is created by the laboratory from contaminant-free laboratory material that most closely resembles the field sample matrix. Laboratories typically use reagent water for water blanks and anhydrous sodium sulfate or precleaned sand for solid, tissue, and wipe blanks.

The laboratory control spike (LCS) demonstrates the level of accuracy to which a laboratory can recover the target analytes. The LCS sample is made from the same matrix as the procedural blank. Unlike the procedural blank, however, the LCS is spiked with known amounts of the target analytes (or a representative subset of target analytes).

The field sample duplicate (Dup) is used to quantify the degree of heterogeneity in the measurement of target analytes in addition to the laboratory precision represented by the surrogate recoveries. The duplicate sample is simply a second aliquot of the field sample. The duplicate sample is particularly effective at demonstrating method precision when using low-level forensic methods because the analytes described in this chapter are ubiquitous and numerous target analyte detections can be evaluated in almost any sample.

The matrix spike (MS) sample is an optional QC sample because the utility of MS data in the assessment of data quality is very limited. It is comprised of a field sample spiked with target analytes (or a representative subset of target analytes). The practical limitations of this QC sample stem from the highly variable range of analyte concentrations in the field samples. Unless the spike amount exceeds the native sample concentration by approximately five times, the overwhelming track record for poor MS recoveries is attributed to bias associated with high native analyte concentrations and sample heterogeneity. In short, it is impractical for laboratories to formulate analyte spiking solution on a sample specific basis. In addition, the measurement of extraction efficiency is already represented by the surrogate recoveries and the sample heterogeneity is represented by the duplicate precision. Even if the MS recoveries represented the extraction efficiency of the sample matrix, which is questionable as described previously, the typical corrective action (analyte flagging in the report) applies only to the native sample selected for spiking and fails to represent the condition of other samples in the batch. In practice, the MS sample is employed only when environmental regulators require it, because the MS sample is typically not worth the cost, and its actual quality value (extraction efficiency) is readily represented by other QC samples.

Other sample preparation QC procedures are practiced as well. For example, every field and batch QC sample is spiked with surrogate(s) at the beginning

of the sample preparation procedure to monitor the extraction efficiency and sample handling accuracy (e.g., dilutions, solvent losses). A surrogate is a compound that is not typically found in nature. It possesses structural similarities to the target analytes but it also contains a slight alteration (e.g., the substitution of deuterium, halogens, or stable isotopes) that differentiates it from the target analytes.

Lastly, the samples should be prepared in a timely fashion in order to maintain the representativeness of the field condition. EPA established holding times to assure that samples collected for regulatory compliance were analyzed quickly and field operations were not delayed. Although these holding times were not strictly based on studies of analyte degradation during sample storage, they provide reasonable guidelines for conducting environmental investigations. In the event that an additional sample is required outside the holding time, the preferred action is to collect the samples again. However, it is recognized that some projects necessitate the analysis of samples outside of holding times because of access limitations or the removal of soil, sediment, and NAPL. One option is to reanalyze a sample outside of holding time and demonstrate comparability by evaluating the concentrations of representative analytes before and after the holding time expired. Another option is to freeze the sample (tissue and solids) before the holding time expires. Freezing is considered a suitable method for eliminating biodegradation and volatilization of semivolatile hydrocarbons and PCBs for at least one year (EPA, 1993a; NOAA, 1998, 2002; Quinn and Latimer, 2003). Again, longer preservation periods may be acceptable if the condition of the sample is demonstrably unchanged over the storage interval. There are no regulatory holding times for NAPL samples and therefore they may be stored in sealed containers at 4°C or less for extended periods of time. Headspace should be minimized when possible, and the level of the NAPL in the vial should be clearly marked to monitor any potential evaporation during storage.

A summary of sample preparation QC parameters and associated data quality objectives (DQOs) is presented in Table 9.2. The purpose of the DQOs is the initiation of a data quality review to identify any QA issues and implement subsequent corrective action to fix the problem and prevent it from occurring again. Most of the minor corrective actions should be performed automatically by the laboratory analysts (e.g., review peak integrations, evaluate matrix effects, and perform instrument maintenance). However, major performance anomalies should be reviewed in a timely fashion by a qualified chemist to determine if the data quality will adversely impact data interpretation. Any quality problems should be identified in the laboratory report and in the analytical summary data tables along with the corrective actions that were performed by the laboratory.

Table 9.2

Data quality objectives for sample preparation batch QC parameters.

QC Parameter	Definition	Type	Performance Criteria	Recommended Corrective Action
Procedure Blank (PB)	Analyte Free Media Surrogate and Internal Standard	Prep Batch Required	Analyte detections must be: less than sample equivalent of low calibration standard, less than 5× detect in field samples, no hydrocarbon or target analyte pattern indicative of laboratory contamination	Review integration Evaluate retention time shift Review with senior chemist Reanalyze sample Reprepare sample batch Flag sample hits within 5× the blank concentration
Laboratory Control Sample (LCS)	Analyte Free Media Surrogate and Internal Standard Commercially available mix from independent source of cat stds	Prep Batch Required	50%–130% recovery VOCs for 90% of the spiked analytes greater than n-C6i50%–130% recorvery SVOCs	Review integration Evaluate retention time shift Review with senior chemist Reanalyze sample Reprepare sample batch Flag outliers and associated sample hits
Field Sample Duplicate (Dup)	Second aliquot of field sample	Prep Batch Required	<30% RPD for analytes at concentrations above the reporting limit	Review integration Evaluate coelution potential Compare analyte and surrogate precision to determine matrix effect Review with senior chemist Flag outliers within the calibration range
Field Sample Triplicate (Tri)	Third aliquot of field sample (NAPL only)	Prep Batch Optional	<25% RSD for analytes more than 5 times the reporting limit	Review integration Evaluate coelution potential Review with project manager Flag outliers within the calibration range

Table 9.2
Continued

QC Parameter	Definition	Type	Performance Criteria	Recommended Corrective Action
Matrix Spike (MS)	Second aliquot of field sample (No NAPL) Commercially available analyte spike from independent source of cal stds	Prep Batch Optional	50%–150% recovery for 90% of the analytes spike more than 5× ambient sample concentration	Review integration Evaluate coelution potential Review with senior chemist Flag outliers within the calibration range in reference field sample only
Surrogate (Sur)	Non-native analyte spike ~3 representative compounds Mid-calibration level	Prep Batch Required	70%–130% recovery VOCs 50%–130% recorvery PAHs 50%–130% recovery for other SVOCs	Review integration Evaluate coelution potential Review with senior chemist Reanalyze sample to confirm Reprepare the sample
Holding Time (HT)	Sample Preparation	Prep Batch Required	Waters 7-days store at 4°C VOC and SVOCs Waters acid preserved 14-days VOC and SVOC[1] Soils/sediments/tissue 14-days store 4°C VOC and SVOC Soils prepare soil/water slurry 14-days store frozen EPA Method 5035 Soils/sediments/tissue 1-year store frozen SVOC NAPL No holding times store at 4°C	Collect new sample Reanalyze sample—compare with results generated within HT Review with senior chemist Quality results as estimates

[1]SVOC compounds added based on Authors experience.

9.3.2.2 QC Parameters for Analytical Procedures

QC procedures and samples are also used to demonstrate the proper operation of the analytical instruments according to established laboratory protocols. The minimum analytical QC parameters include:

- Mass spectrometer tune (MS only)
- Multilevel initial calibration standards (5 different concentrations recommended)
- Continuing calibration standards (run before and after the samples)
- Independent check standard (that contain the calibration analytes at known concentrations from an independent source)
- Reference sample (well characterized sample that demonstrates comparability over time)
- Internal standards (recommended)

The instrument tune is performed on a mass spectrometer before the analysis of calibration standards and samples. Thereafter, the tune conditions used to analyze the calibration standards must remain unchanged for the analysis of QC and field samples. This tuning procedure assures that the ion source, mass filter, and electronic settings produce mass spectra that matches a previously determined standard. Many mass spectrometers are tuned to a standard set of conditions based on perfluorotributylamine (PFTBA, $(CF_3(CF_2)_3)_3N$, CAS # 311-89-7). When tuned against PFTBA, mass fragments can be efficiently resolved from one another and the mass spectra of unknown peaks can be identified based on the match with mass spectra reference libraries. Tunes with other reference compounds, such as decafluorotriphenylphosphine (DFTPP, CAS # 5074-71-5, $C_6H_5[C_6F_5]_2P$), can be used to mimic PFTBA. The use of instrument manufacturer software to standardize the instrument to PFTBA using DFTPP generates a low bias for high molecular weight compounds. Thus, it is preferable to tune with the standard PFTBA compound in conjunction with the tune criteria in Table 9.3. Although there are minor differences in the relative proportions of the ions across the spectrometers dynamic mass range between these two tuning methods, both tunes produce spectra of high quality that match reference spectra library such as the NIST/EPA/NIH Mass Spectral Library (NIST, 2006).

The determination of analyte concentrations is critical for generating data of known quality. The definitive measurement of analyte concentrations must be based on a series of calibration standards prepared at multiple concentrations. The analysis of these multilevel calibration standards demonstrates the linear range of the instrument including the lower limit of definitive measurement (i.e., the reporting limit). Initial calibration (ICal) curves are especially important for certain sophisticated detectors, like mass spectrometers, that

Table 9.3

Data quality objectives for analytical batch QC parameters.

QC Parameter	Definition	Type	Performance Criteria	Recommended Corrective Action
MS tune	PFTBA SVOC	Analytical Batch Required	m/z 69: Base peak >100,000 counts m/z 219: 30%–60% base peak counts m/z 502: 5%–11% base peak counts	Rerun tune Clean source Review with senior chemist
	BFB VOC	Analytical Batch Required	Per SW846 8260B	Rerun tune Clean source Review with senior chemist
	DFTPP SVOC	Analytical Batch (Optional)	Per "Maximum Sensitivity" tune criteria for PCBs Per SW846 8270C for other SVOCs	Rerun tune Clean source Review with senior chemist
Initial Calibration (Ical)	Analyte mix for instrument calibration 5 concentration levels (Nominal)	Analytical Batch Required	%RSD ≤ 25% for 90% of analytes %RSD ≤ 35% for all analytes >n-C6	Review integrations Rerun tune (MS) Rerun ICal Instrument maintenance Clean source Review with senior chemist
Continuing Calibration (Ccal)	Analyte mix for instrument calibration Mid concentration level Bracket every ≤10 samples	Analytical Batch Required	%D ≤ 25% for 90% of analytes %D ≤ 35% for all analytes >n-C6	Review integrations Rerun CCal Instrument maintenance Clean source Review with senior chemist

Table 9.3

Continued

QC Parameter	Definition	Type	Performance Criteria	Recommended Corrective Action
Independent Check (IC)	Mixture of target analytes in calibration standard prepared from a source that is independent of the Ical mixture	One per initial calibration Required	80%–120% recovery	Review Integations Rerun IC Instrument maintenance Clean source Prepare IC from a different source Review with senior chemist
Reference Sample (RS)	Routinely analyzed field sample Gasoline for VOCs Crude Oil for PAHs, Biomarkers, and Alkanes Aroclor for PCBs	Analytical Batch Required	50%–130% recovery for representative VOCs 65%–135% recovery PAHS and Hopane for Biomarkers Alkanes—For Chromatogram only	Review integrations Reanalyze RS Review with senior chemist Reanalyze analytical batch
Internal Standard (IS)	Non-native analyze spike 1–5 representative compounds Mid-calibration level	Analytical Batch Required	50%–200% of the area of the IS in the associated calibration standard	Review integrations Reanalyzed sample Review with senior chemist
Standard Reference Material (SRM)	Procured from NIST/Other Gasoline PAH PCB	Analytical Batch (Optional)	65%–135% recovery from certified value	Review integrations Reanalyze SRM Review with senior chemist Reanalyze analytical batch
Holding Time (HT)	Analytical	Analytical Batch Required	40 day storage before analysis (extracts)	Collect new sample Reanalyze sample—compare with results generated within HT Review with senior chemist Qualify results as estimates

possess a more narrow linear range with higher variability at the low and high extremities of the response curve. In the absence of an ICal, it is difficult to determine if the analyte responses are measured within the linear range of the instrument. It is also not possible to determine if low level and high level detections are made with the same accuracy and precision. Unless the analyte measurement is being measured to determine its presence or absence, the accuracy and precision of the method is a critical underlying assumption behind the statistical or numerical assessment of the data.

An independent check (IC) standard is run after the ICal and before the field and QC samples to independently verify the accuracy of the ICal. The independent check standard is formulated with the target analytes (or representative subset of target analytes) originating from a source that is independent of the target analytes in the ICal. Because the forensic analytes include many isomers that are not commercially available, the IC sample is intentionally formulated with analytes that are more readily available. The IC is used in concert with a reference sample that contains a wider range of forensic analytes, described as follows.

A reference sample (RS) is also run after the ICal and before the field and QC samples to confirm the measurement precision for target analytes, especially those that are not commercially available (e.g., alkylated PAHs). A commonly used RS sample in hydrocarbon analysis is a crude oil such as Alaska North Slope Crude oil that has been well characterized for a wide range of PAH, saturated hydrocarbon, and geochemical biomarker compounds by several independent laboratories (Page *et al.*, 1995; Douglas *et al.*, 2004). Over time, reference concentrations and standard variances can be computed and used to evaluate the accuracy of subsequent ICals, and to confirm that instruments are generating high quality analytical measurements for a full range of target analytes. A continuing calibration standard (CCal) is analyzed daily before and after samples to demonstrate the stability of the initial calibration over time. The CCal is usually an aliquot of an initial calibration standard from the middle range of the ICal curve. The concentrations of the CCal analytes are compared to the ICal to determine the degree to which analyte concentrations can be analyzed accurately at the time of sample analysis.

The use of other analyte measurement techniques (e.g., raw peak areas or peak heights) is discouraged in forensics investigations for many reasons. First, peak height is a parameter that is controlled in part by instrument operating conditions; for example, two peaks of equal concentration that elute at the beginning and end of the run may have different heights due to the gradual peak widening that occurs throughout the analytical run. Peak height is also a function of concentration; for example, high concentrations of an analyte can exceed the capacity of the GC column and cause peak widening and

shortening. These instrument-induced biases in the peak height measurement result in nonlinear variations in measured hydrocarbon pattern that can be erroneously construed as different chemical concentrations or patterns. Second, both peak area and height measurements typically are not derived from analyses accompanied by a multilevel calibration curve that demonstrates the linear performance range of the analysis. In this instance, the laboratory has no way of demonstrating that the measurements at different apparent concentrations were taken in a linear operating range of the instrument; such data negate the principal assumption of quantitative fingerprinting, that is, that relative differences in concentration are indicative of differences in chemical composition and source. Multilevel calibration curves are an absolute necessity for definitive chemical fingerprinting using numerical analysis beyond a "presence or absence" test. Third, laboratories that fail to demonstrate the linearity of the instrument also usually fail to run continuing calibration standards that demonstrate the reproducibility of the instrument on a daily basis. Instrument performance is known to change as more samples are analyzed and organic residues accumulate in the injector, column, and detector (Jennings, 1987). These changes can occur rapidly, especially when analyzing samples with high levels of natural and anthropogenic hydrocarbons. It is for this reason that EPA methods require the analysis of at least one continuing calibration standard per day (EPA, 1997a). Finally, unlike PAH concentration results with known precision and accuracy, raw peak area or peak height fingerprints vary over time and between individual laboratories due to the use of replacement columns, different instruments, and different analytical methods. Accordingly, there is little ability to demonstrate comparability among sample data collected over the course of a typical investigation or by different laboratories. In summary, these many uncontrolled factors related to raw peak area and height have been considered by EPA and resulted in its requirement that concentrations based on multilevel calibration be determined for all compounds of concern in order to generate data protective of public health (EPA, 1997). If peak height and area data (without initial calibrations, continuing calibrations, and limited batch QC samples) must be used, these data are considered estimated values and admissible for qualitative or semiquantitative analysis only. By contrast, quantitative analysis should be limited to concentration data only. The requirements of the regulatory compliance methods and the associated functional guidelines for data validation (EPA, 1995) should be construed as a minimum benchmark for demonstrating the applicability of forensic data for defensible conclusions about the source of anthropogenic chemicals in the environment.

Internal standards (IS) are added to every calibration standard sample extract immediately prior to analysis (EPA, 1997a). The IS standards are used

to calculate analyte concentrations using the internal standard quantitation method and the average response factors from the ICal. It should be noted that although each level of the calibration has a different concentration of each target analyte (used to generate an average response factor over the working analytical range), surrogates and internal standards contained in these levels are constant and equivalent to that amount spiked into authentic samples.

The calculation of an average response factor (RF) for each compound is as follows:

$$RF = (A_s \times C_{is})/(A_{is} \times C_s)$$

where:

A_s = Area of the target analyte to be measured
A_{is} = Area of the appropriate recovery internal standard
C_{is} = Amount of the recovery internal standard added to the extract (ng)
C_s = Amount of the target analyte to be measured

Based on the average response factors generated from the initial calibration set, concentrations for the target compounds are calculated using:

$$C_e = [(A_s \times I_s)/(A_{is} \times RF)]/V_s$$

where:

C_e = Sample extract concentration
A_s = Area of the target analyte
A_{is} = Area of the appropriate recovery internal standard
I_s = Amount of recovery internal standard added to the extract (ng)
RF = Average response factor for the target compound from initial calibration
V_s = Sample size (volume (L), g dry, wet or oil weight)

A summary of analytical QC parameters and associated data quality objectives (DQOs) is presented in Table 9.3. As described previously, the purpose of the DQOs is the initiation of a data quality review and subsequent corrective action to fix the problem and prevent it from reoccurring. Collectively, the QA/QC program serves as a system of checks and balances that promotes the generation of high quality data and motivates continuous procedural improvements in the forensic methods discussed in the balance of this chapter.

9.4 VOLATILE HYDROCARBON FINGERPRINTING METHODS

Light refined petroleum products (e.g., automotive gasoline, aviation gas, various refinery intermediates) and tar distillates (e.g., light oil) or the water-soluble fractions of many hydrocarbon mixtures are comprised largely of volatile hydrocarbons in the C_4 to C_{12} range (Kaplan *et al.*, 1997; Emsbo-Mattingly *et al.*, 2003a; Stout *et al.*, 2006). These products are composed of a wide range of compounds, the most common being those defined by the so-called PIANO groupings: *P*araffins (e.g., *n*-hexane), *I*soparaffins (e.g., 2,2,4-trimethylpentane (isooctane)), *A*romatics (e.g., benzene), *N*aphthenes (e.g., methylcyclohexane), and *O*lefins (e.g., cis-2-hexene). Fortunately, analytical standards containing many of the PIANO chemicals of concern (COC) to the forensic investigator are available commercially, and can be analyzed by conventional purge and trap (P&T) GC/MS methods such as EPA Method 8260B or as modified by Uhler *et al.* (2003). The major modification in the Uhler *et al.* (2003) method is the addition of a wider range of gasoline specific target analytes. Table 9.4 shows a typical inventory of volatile PIANO hydrocarbons—and related compounds—that can be useful in chemical fingerprinting of light products, particularly automotive gasoline and light solvents. Importantly, this extended list of target compounds includes the volatile compounds of regulatory concern—benzene, toluene, ethylbenzene, and *o*-, *m*-, and *p*-xylenes (BTEX). Thus, using methods and analyte lists such as those described here, allows the environmental forensic chemist to gather information simultaneously for both regulatory and forensic applications.

Automotive gasoline can, or had in the past, contained various additives or blending agents, for example, ether or alcohol oxygenates such as methyl-*tert*-butyl ether (MTBE), ethyl-*tert*-butyl ether (ETBE), *tert*-amyl-methyl ether (TAME), and *tert*-butyl alcohol (TBA), ethylene dibromide (EDB) and ethylene dichloride (EDC), methylcyclopentadienyl manganese tricarbonyl (MMT; Gibbs, 1990; Stout *et al.*, 2006) and alkylleads (Gibbs 1990, 1993). These volatile nonhydrocarbons also can be measured using the PIANO GC/MS methods discussed later, and should be included in any comprehensive target analyte list (see Table 9.4). In addition, with the evolving regulations on sulfur in gasoline (EPA, 1999c), measurement of volatile sulfur compounds (e.g., thiophenes) by GC/MS and total sulfur is appropriate and often useful to the forensic chemist for constraining the age of a gasoline product release.

Volatile hydrocarbon data can be used to identify product type and source, degree of weathering, and in some cases constrain the age of the release (Kaplan *et al.*, 1996; Stout *et al.*, 2006). Due to the labile nature of these

Table 9.4

Inventory of volatile hydrocarbon PIANO analytes.

Volatile Hydrocarbon PIANO Analyte	Peak ID	Group ID	Molecular Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
Pentane	C5	P	C_5H_{12}	43	2 μg/L	2 μg/Kg
Hexane	C6	P	C_6H_{14}	57	2 μg/L	2 μg/Kg
Heptane	C7	P	C_7H_{16}	43	2 μg/L	2 μg/Kg
Octane	C8	P	C_8H_{18}	43	2 μg/L	2 μg/Kg
Nonane	C9	P	C_9H_{20}	43	2 μg/L	2 μg/Kg
Decane	C10	P	$C_{10}H_{22}$	43	2 μg/L	2 μg/Kg
Undecane	C11	P	$C_{11}H_{24}$	57	2 μg/L	2 μg/Kg
Dodecane	C12	P	$C_{12}H_{26}$	43	2 μg/L	2 μg/Kg
Tridecane	C13	P	$C_{13}H_{28}$	57	2 μg/L	2 μg/Kg
Isopentane	IP	I	$C_{13}H_{28}$	43	2 μg/L	2 μg/Kg
2,3-Dimethylbutane	23DMB	I	C_5H_{12}	71	2 μg/L	2 μg/Kg
2-Methylpentane	2MP	I	C_6H_{14}	43	2 μg/L	2 μg/Kg
3-Methylpentane	3MP	I	C_6H_{14}	57	2 μg/L	2 μg/Kg
2,2-Dimethylpentane	22DMP	I	C_6H_{14}	57	2 μg/L	2 μg/Kg
2,4-Dimethylpentane	24DMP	I	C_6H_{14}	43	2 μg/L	2 μg/Kg
2-Methylhexane	2MH	I	C_7H_{16}	43	2 μg/L	2 μg/Kg
2,3-Dimethylpentane	23DMP	I	C_7H_{16}	56	2 μg/L	2 μg/Kg
3-Methylhexane	3MH	I	C_7H_{16}	43	2 μg/L	2 μg/Kg
Isooctane	ISO	I	C_8H_{18}	57	2 μg/L	2 μg/Kg
2,5-Dimethylhexane	25DMH	I	C_8H_{18}	57	2 μg/L	2 μg/Kg
2,4-Dimethylhexane	24DMH	I	C_8H_{18}	57	2 μg/L	2 μg/Kg
2,2,3-Trimethylpentane	223TMP	I	C_8H_{18}	57	2 μg/L	2 μg/Kg
2,3,4-Trimethylpentane	234TMP	I	C_8H_{18}	43	2 μg/L	2 μg/Kg
2,3,3-Trimethypentane	233TMP	I	C_8H_{18}	43	2 μg/L	2 μg/Kg
2,3-Dimethylhexane	23DMH	I	C_8H_{18}	43	2 μg/L	2 μg/Kg
3-Ethylhexane	3EH	I	C_8H_{18}	43	2 μg/L	2 μg/Kg
2-Methylheptane	2MHEP	I	C_8H_{18}	57	2 μg/L	2 μg/Kg
3-Methylheptane	3MHEP	I	C_8H_{18}	43	2 μg/L	2 μg/Kg
Benzene	B	A	C_6H_6	78	2 μg/L	2 μg/Kg
Toluene	T	A	C_7H_8	91	2 μg/L	2 μg/Kg
Ethylbenzene	EB	A	C_8H_{10}	91	2 μg/L	2 μg/Kg
p/m-Xylene	MPX	A	C_8H_{11}	91	2 μg/L	2 μg/Kg
Styrene	STY	A	C_8H_8	104	2 μg/L	2 μg/Kg
o-Xylene	OX	A	C_8H_{10}	91	2 μg/L	2 μg/Kg
Isopropylbenzene	IPB	A	C_9H_{12}	105	2 μg/L	2 μg/Kg
n-Propylbenzene	PROPB	A	C_9H_{12}	91	2 μg/L	2 μg/Kg
1-Methyl-3-ethylbenzene	1M3EB	A	C_9H_{12}	105	2 μg/L	2 μg/Kg
1,3,5-Trimethylbenzene	135TMB	A	C_9H_{12}	105	2 μg/L	2 μg/Kg
1-Methyl-2-ethylbenzene	1M2EB	A	C_9H_{12}	105	2 μg/L	2 μg/Kg
1,2,4-Trimethylbenzene	124TMB	A	C_9H_{12}	105	2 μg/L	2 μg/Kg
sec-Butylbenzene	SECBUT	A	$C_{10}H_{14}$	105	2 μg/L	2 μg/Kg
1-Methyl-3-isopropylbenzene	1M3IPB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1-Methyl-4-isopropylbenzene	1M4IPB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1-Methyl-2-isopropylbenzene	1M2IPB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
Indan	IN	A	C_9H_{10}	117	2 μg/L	2 μg/Kg
1-Methyl-4-propylbenzene	1M4PB	A	$C_{10}H_{14}$	105	2 μg/L	2 μg/Kg
n-Butylbenzene	BUTB	A	$C_{10}H_{14}$	91	2 μg/L	2 μg/Kg

Table 9.4
Continued

Volatile Hydrocarbon PIANO Analyte	Peak ID	Group ID	Molecular Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
1,2-Dimethyl-4-ethylbenzene	12DM4EB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,2-Diethylbenzene	12DEB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,4-Dimethyl-2-ethylbenzene	14DM2EB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,3-Dimethyl-4-ethylbenzene	13DM4EB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,3-Dimethyl-5-ethylbenzene	13DM5EB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,2-Dimethyl-3-ethylbenzene	12DM3EB	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
1,2,4,5-Tetramethylbenzene	1245TMP	A	$C_{10}H_{14}$	119	2 μg/L	2 μg/Kg
Pentylbenzene	PENTB	A	$C_{11}H_{16}$	91	2 μg/L	2 μg/Kg
Naphthalene	N	A	$C_{10}H_{8}$	128	2 μg/L	2 μg/Kg
2-Methylnaphthalene	2MN	A	$C_{11}H_{10}$	142	2 μg/L	2 μg/Kg
1-Methylnaphthalene	1MN	A	$C_{11}H_{10}$	142	2 μg/L	2 μg/Kg
Cyclopentane	CYP	N	$C_{5}H_{10}$	70	2 μg/L	2 μg/Kg
Methylcyclopentane	MCYP	N	$C_{6}H_{12}$	56	2 μg/L	2 μg/Kg
Cyclohexane	CH	N	$C_{6}H_{12}$	56	2 μg/L	2 μg/Kg
Methylcyclohexane	MCYH	N	$C_{7}H_{14}$	83	2 μg/L	2 μg/Kg
1-Pentene	1P	O	$C_{5}H_{10}$	42	2 μg/L	2 μg/Kg
2-Methy1-1-butene	2M1B	O	$C_{5}H_{10}$	55	2 μg/L	2 μg/Kg
2-Pentene (trans)	T2P	O	$C_{5}H_{10}$	55	2 μg/L	2 μg/Kg
2-Pentene (cis)	C2P	O	$C_{5}H_{10}$	55	2 μg/L	2 μg/Kg
1-Hexene	1HEX	O	$C_{6}H_{12}$	56	2 μg/L	2 μg/Kg
1-Octene	1O	O	$C_{8}H_{16}$	55	2 μg/L	2 μg/Kg
1-Nonene	1N	O	$C_{9}H_{18}$	56	2 μg/L	2 μg/Kg
1-Decene	1D	O	$C_{10}H_{20}$	41	2 μg/L	2 μg/Kg
1,2-Dichloroethane	12DCA	ADD	$C_{2}H_{4}Cl_{2}$	62	2 μg/L	2 μg/Kg
1,2-Dibromoethane	12DBE	ADD	$C_{2}H_{4}Br_{2}$	107	2 μg/L	2 μg/Kg
MMT	MMT	ADD	$CH_{3}C_{5}H_{4}Mn(CO)_{3}$	134	2 μg/L	2 μg/Kg
Tertiary butanol	TBA	OX	$C_{4}H_{10}O$	59	2 μg/L	2 μg/Kg
MTBE	MTBE	OX	$C_{5}H_{12}O$	73	2 μg/L	2 μg/Kg
Diisopropyl Ether (DIPE)	DIPE	OX	$C_{6}H_{14}O$	45	2 μg/L	2 μg/Kg
Ethyl Tertiary Butyl Ether (ETBE)	ETBE	OX	$C_{6}H_{14}O$	95	2 μg/L	2 μg/Kg
TAME	TAME	OX	$C_{6}H_{14}O$	73	2 μg/L	2 μg/Kg
Thiophene	THIO	S	$C_{4}H_{4}S$	84	2 μg/L	2 μg/Kg
2-Methylthiophene	2MTHIO	S	$C_{5}H_{6}S$	97	2 μg/L	2 μg/Kg
3-Methylthiophene	3MTHIO	S	$C_{5}H_{6}S$	97	2 μg/L	2 μg/Kg
2-Ethylthiophene	2ETHIO	S	$C_{6}H_{8}S$	97	2 μg/L	2 μg/Kg
Benzothiophene	BT0	S	$C_{8}H_{6}S$	134	2 μg/L	2 μg/Kg

Group IDs:
P Paraffins
I Isoparaffins
A Aromatics
N Naphthenes
O Olefins
ADD Gasoline Additives
OX Gasoline Oxygenates
S Thiophenes

compounds, the impact of weathering on the quantity and distribution of these compounds in environmental samples must be included within the interpretation strategy. In this section we describe several analytical methods currently used for the forensic analysis of volatile organic compounds (VOCs) in soils, water, air, and NAPL and provide examples of how this important class of compounds may be used for forensic investigations.

9.4.1 VOLATILE HYDROCARBONS IN AIR AND VAPOR BY GC/MS

The volatile components that emanate from complex hydrocarbon mixtures such as petroleum fuels or tar-derived wastes contain dozens of volatile organic compounds that, if measured, can be used to identify or fingerprint the composition of air- or vapor-borne contamination. Whether measured in samples of subsurface vapor, indoor-, or ambient air, the ability to fingerprint hydrocarbon contamination of any sort is strongly dependent on the number of hydrocarbon chemicals that are measured in samples, because as the number of compounds measured in a sample increases, so does the ability to develop a source-specific chemical fingerprint. In order to more accurately characterize subsurface vapors, indoor-, or outdoor air, we have been measuring a more extensive list of volatile organic compounds using a modified version of EPA Method TO-15, *Determination of Volatile Organic Compounds in Air Collected in Specially Prepared Canisters and Analyzed by Gas Chromatography/Mass Spectrometry (GC/MS)* (EPA, 1999d). This method holds tremendous promise as an important chemical fingerprinting tool in vapor intrusion studies, whether to confirm or refute a potential subsurface migration pathway, to understand the complexity and varied sources of indoor air, or simply to better characterize outdoor ambient air so that difference between indoor and outdoor air can be quantified (EPA, 2002b; McCarthy *et al.*, 2006).

The primary modification to the TO-15 method is an expanded analyte list that provides far greater specificity in the characterization of hydrocarbons in air. The modified method incorporates many of the same source diagnostic target analytes presented in Table 9.4 for the forensic analysis of air samples collected in conventional SUMMA canisters. In the field, air or subsurface vapor is sampled by introduction into stainless steel SUMMA canisters, prepared following procedures outlined in EPA TO-15 (EPA, 1997). A sample of air is drawn through a sampling train comprised of components that regulate the rate and duration of sampling into the preevacuated canister. Sample holding times of up to 30 days have been demonstrated for many of the target compounds of interest (Kelly and Holdren, 1995), and are so recommended by EPA in Method TO-15. Analytically, samples are analyzed by introducing a known volume of sample from the canister through a multisorbent concentra-

tor. After concentration and drying, the target compounds are thermally desorbed from the concentrator using an inert carrier gas, and then concentrated by trapping the residues on a reduced temperature trap. The sample is then thermally desorbed and carried onto a gas chromatographic column for separation, followed by mass spectrometry detection (GC/MS).

Standards, quality control samples, and authentic field samples are quantified using the method of internal standards. EPA Method TO-15 recommends using an internal standard spiking mixture that contains bromochloromethane, chlorobenzene-d_5, and 1,4-difluorobenzene at 10 parts-per-million by volume (ppmv) each, although other suitable internal standards that cover the approximately C_4 to C_{15} carbon range may be used. A 500 μL aliquot of this mixture is spiked into 500 mL of sample, resulting in an effective concentration of 10 parts-per-billion by volume (ppbv). In practice, the internal standard is introduced into the trap during the collection time for all calibration, blank, and sample analyses.

Calibration standards containing target compounds of interest such as those listed in Table 9.4 are prepared from certified pure, commercially available reference standards (e.g., Supelco, Inc. or Restek, Inc.). The neat standards are used to prepare working standards using the high pressure cylinder technique, or equivalent, described in EPA TO-15.

Gas chromatography is carried out with an instrument capable of cryogenic operation, and capable of temperature programming to a maximum temperature of 300˚C. The system must include or be interfaced to a concentrator and have all required accessories including analytical columns and gases. Carrier gas and transfer lines should be stainless steel or copper tubing. The GC should be equipped with a nonpolar stationary phase fused silica capillary column capable of separating the target compounds of interest. Nominally, a RTX-1 PONA©, 50 m, 0.32 ID, 0.5 μm film thickness fused silica capillary column (Restek, Inc.) has been found to be an effective column for this analysis. The GC/MS analytical conditions are listed in Table 9.5. The GC is

Table 9.5
The GC/MS can be operated under conditions recommended in TO-15: Equivalent conditions may be used if data quality guidelines are achieved.

GC Parameter	Program Settings
Initial Temperature:	–50°C
Initial Hold Time:	2 minutes
Ramp Rate:	8°C/minute
Final Temperature:	200°C
Final Hold Time:	60 minutes
Detector Temperature:	280°C
Flow Rate (Helium):	1–3 mL/min
MS Scan Parameters:	Full Scan Mode; range 35–350 *m/z*

Table 9.6
BFB mass spectrometer tuning criteria.

Mass	Ion Abundance Criteria
50	8.0 to 40.0% of m/e 95
75	30.0 to 66.0% of m/e 95
95	Base Peak, 100% Relative Abundance
96	5.0 to 9.0% of m/e 95
173	Less than 2.0% of m/e 174
174	50.0 to 120.0% of m/e 95
175	4.0 to 9.0% of m/e 174
176	93.0 to 101.0% of m/e 174
177	5.0 to 9.0% of m/e 176

interfaced to a quadrupole or similarly performing 70 ev positive ionization mass spectrometer that can scan from 35 to 300 amu every one second or less. The mass spectrometer is tuned using 50 ng or less of p-bromofluorobenzene (BFB) calibrant, to the criteria listed in Table 9.6.

Prior to analysis of authentic samples, the mass spectrometer must be tuned with BFB, and meet prescribed criteria. The tune of the mass spectrometer should be checked every 24 hours. The gas chromatograph should be calibrated with an initial five-level calibration spanning a concentration range of approximately 1 to 100 ppbv per component. The GC calibration should be checked every 24 hours with a mid-level calibration mixture. The calculated percent relative standard deviation (%RSD) for the relative response factors (RRF) for each compound in the calibration table should be less than 30% with at most two exceptions up to a limit of 40%. Adjustments to this criteria may be made on a case-specific basis, depending on the stability and chromatographic performance of certain analytes. The relative retention time (RRT) for each target compound should be within 0.06 RRT units of the mean RRT for the compound from the initial calibration. The area response for each compound must be within 40% of the mean area response over the initial calibration range for each internal standard. The retention time shift for each of the internal standards at each calibration level should be within 20 s of the mean retention time over the initial calibration range for each internal standard. Routine quality control samples should be run with each batch of samples. At a minimum, a method blank and a sample duplicate should be analyzed with each batch of 20 or fewer samples. The data quality objectives for this method are listed in Table 9.3.

The modified TO-15 analytical method provides extremely useful qualitative and quantitative data for developing detailed characterization of volatile hydrocarbons in air and subsurface vapors. Qualitatively, carefully run TO-15 analyses will provide valuable gas chromatographic fingerprints that can be revealing about the gross compositional similarities and differences among

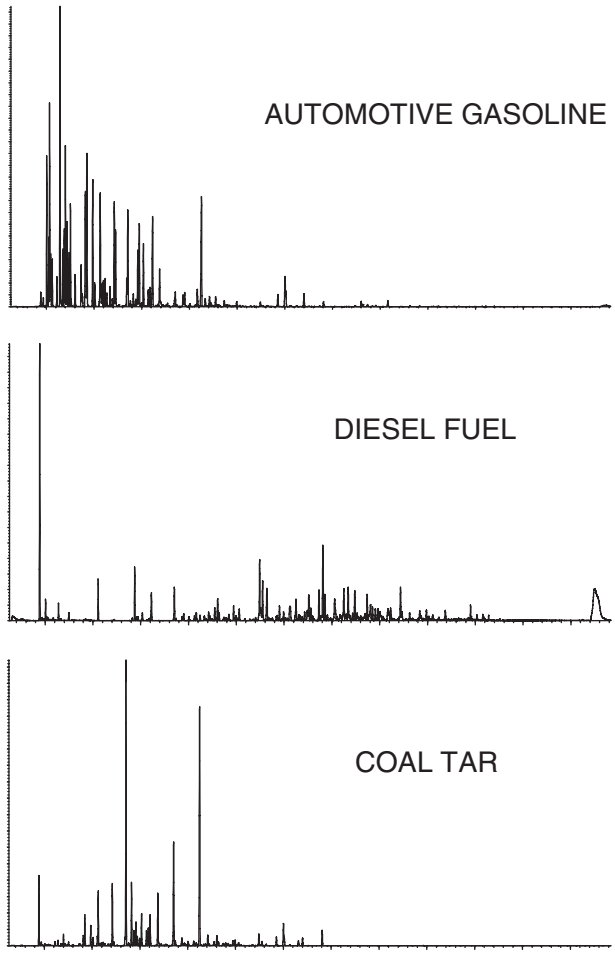


Figure 9.2
Comparison of TO-15 product chromatograms. Product standards prepared as headspace samples.

samples, or between samples and suspect sources. Consider Figure 9.2, which illustrates a series of TO-15 gas chromatograms for several common hydrocarbon products: gasoline, diesel fuel, and coal tar. The qualitative differences among these samples are striking and obviously distinguish each from one another. The classic features of automotive gasoline vapor—a complex mixture of straight and branched chained alkanes, monoaromatics, cyclic compounds, and low level olefins—are evident; in contrast, the vapor phase of diesel fuel is simpler, composed of higher molecular weight cyclic alkanes, branched alkanes, and C_1-C_3 alkyl substituted benzenes; coal tar liquid vapors are composed almost exclusively of monoaromatic compounds.

Quantitatively, because the target analyte list of the modified EPA TO-15 method contains almost 10 times the number of hydrocarbon target compounds than measured by conventional TO-15, analysis by the modified tech-

nique provides the forensic chemist the opportunity to carefully fingerprint and compare the chemical makeup of samples and suspected sources, for example, indoor air versus subslab vapor composition.

An example of the power of the modified EPA TO-15 method in a forensic investigation is provided in Figure 9.3, which shows comparative data from the analysis of the same subsurface soil gas taken above a NAPL plume composed of automotive gasoline by EPA TO-15 in both its standard and modified deployment. The top panel in Figure 9.3 shows concentration data for a the analysis of the soil gas using the modified EPA TO-15 method (truncated at 1,2,4-TMB for clarity); the bottom panel depicts the results of the analysis of the same vapor sample developed from the abbreviated target compound list using standard EPA Method TO-15. It is clear that data generated from the standard TO-15 analysis provides a limited (and largely nonspecific) fingerprint, whereas the more detailed analysis derived from the modified EPA TO-15 offers a much more detailed depiction of the make up of the vapor phase. Such data greatly increases the potential for confirming or refuting any hypothesis surrounding a completed vapor intrusion pathway between the subsurface and indoor air.

To further illustrate the potential utility of modified EPA TO-15 in vapor intrusion investigations, we offer data collected both subslab and within indoor air of a building situated above a subsurface gasoline NAPL plume. Figure 9.4 shows the histograms of the relative concentrations of C_4 to C_{10} gasoline range hydrocarbons measured using a modified EPA TO-15 method for subsurface vapor (lower panel) and indoor air (upper panel) of the building. The data reveal that, although there are some common chemicals found in both samples (e.g., varying concentrations of BTEX), there are also significant differences in the chemical compositions of the subsurface vapor and the indoor air. Notably, the indoor air contained elevated concentrations of ethanol and MTBE, benzene, and toluene relative to the subsurface vapor. Also evident is the significant amounts of C_7–C_{10} hydrocarbons—particularly alkylated benzenes—in the indoor air that was largely absent in the subsurface vapor. Such distinct chemical differences must be reconciled before a clear connection between subsurface contamination and indoor air-borne hydrocarbons can be made; that is, the data are insufficient to match or fingerprint the chemical profile of subslab contamination versus the profile of indoor contaminants. Such differences between these two samples would have gone unrecognized, had the samples been analyzed by conventional EPA TO-15.

The source of indoor air contamination often is related back to the underlying NAPL and or NAPL contaminated soil. The following section discusses the collection and analyses of those samples for volatile hydrocarbon contaminants.

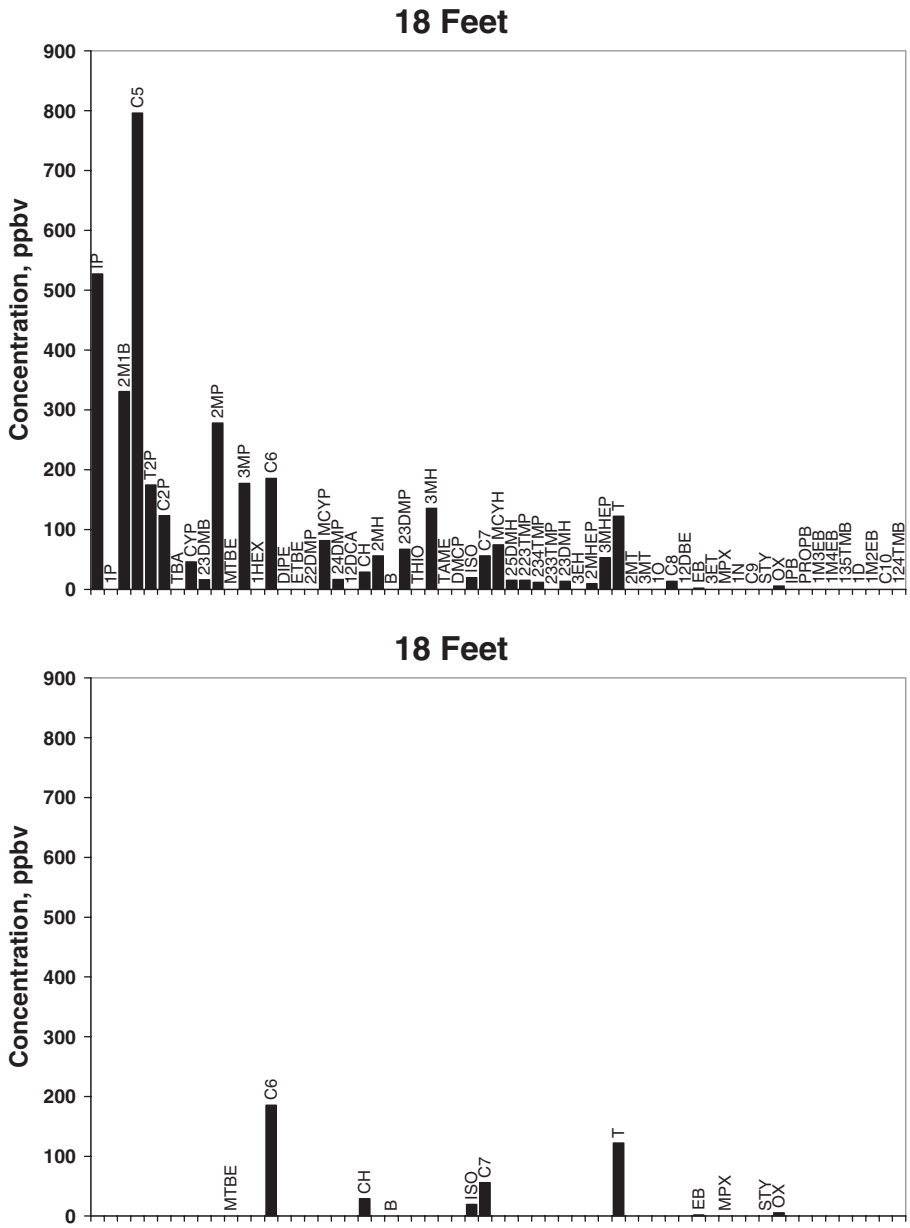


Figure 9.3
Histograms of a soil gas sample from above an automotive gasoline plume, standard TO-15 (bottom) and modified TO-15 with expanded target analyte list (top). Compand abbreviations found in Table 9.4.

9.4.2 COLLECTION OF SOLID AND LIQUID SAMPLES FOR VOLATILE HYDROCARBON ANALYSIS

Under most circumstances, once released into the environment, the chemical fingerprint of volatile hydrocarbon products (e.g., gasoline) begins to change almost immediately due to evaporation. Due to the labile nature of these

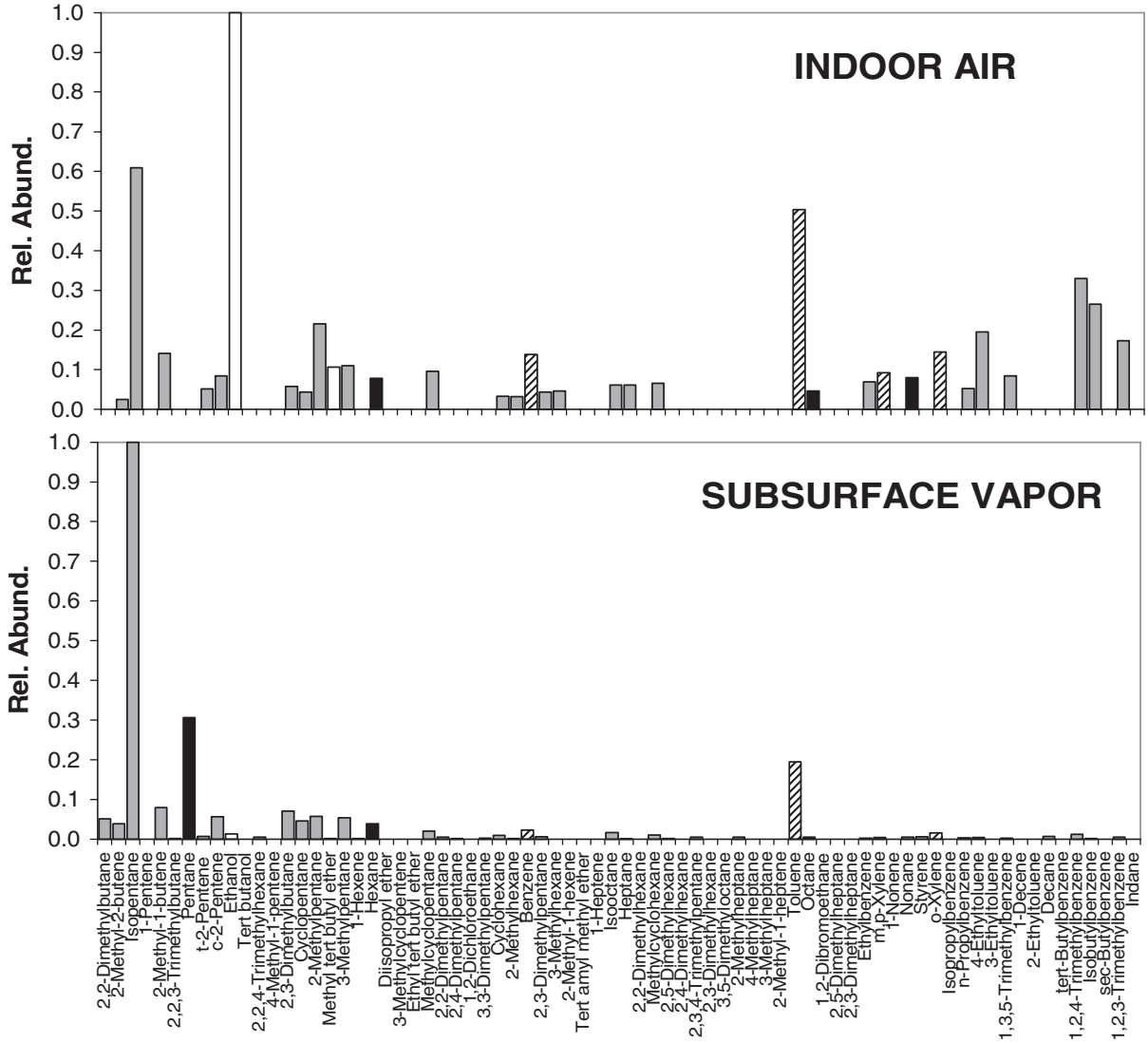


Figure 9.4

PIANO histograms of relative concentrations of C4-C10 gasoline range hydrocarbons and additives in subsurface vapor and indoor air of a residential building situated above gasoline-contaminated soil and groundwater.

volatile compounds it is important to collect samples that preserve the chemical signature of the sample as it exists at the time of collection. Therefore the manner by which volatile data are used is highly dependent on how the field sample is collected and stored. The following sections discuss field sample collection methods that are designed to minimize alteration of these volatile chemical signatures.

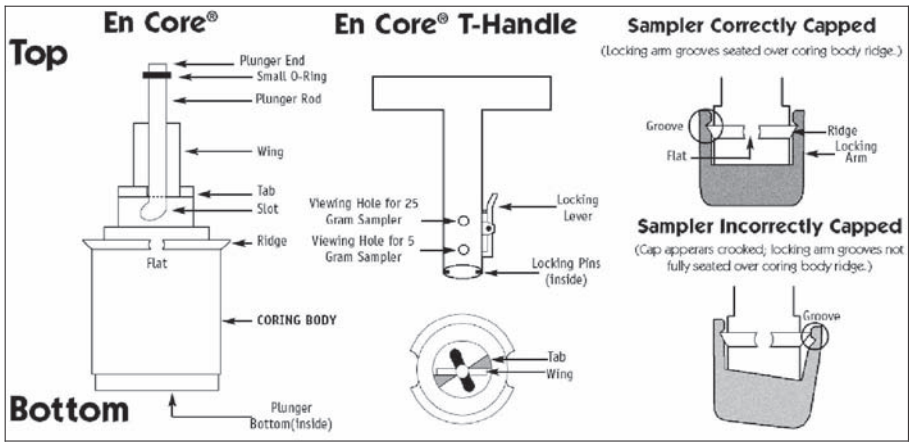


Figure 9.5
Diagram of the En Core® soil sampling system.

9.4.2.1 Soil Samples

The collection of soil samples for VOC analysis should be performed in a manner consistent with EPA methods (EPA, 1997a). Soil sample collection methods have evolved during the past decade in an attempt to minimize sample losses due to volatilization and biodegradation. EPA Method 5035A for example uses the En Core® sampler for the collection of soil samples for volatile hydrocarbon analysis. The En Core® sampler (see Figure 9.5) is a tube with a plunger and a cap to seal the end of the tube after the sample is collected. The sampler is attached to a pole or a specially designed insertion device (see Figure 9.5) and pushed into the soil sample. The plunger is flush with the soil surface as the beveled tube penetrates the soil. When the plunger reaches the end of the tube, the sample is withdrawn. The *o*-ring is then wiped with a clean Kim Wipe® (or equivalent) to remove any soil particles that may prevent a proper seal, and the end cap is then locked into place. The sampler is then placed back into its prelabeled foil container and shipped back to the laboratory where it must be processed into preweighed VOA vials within 48 hours. An alternative approach is to process the samples in the field therefore extending the holding times of the samples to 14 days. When field processing is performed, the En Core® sampler can be replaced by a 10 mL plastic syringe that has been cut and beveled at the proper length to collect a 5 gram soil sample. The 5 gram soil sample is placed into a pretared 40 mL septum capped VOA vial containing 5 mL of deionized water and a Teflon® stirbar. The soil/water slurry is then immediately frozen and may be analyzed within 14 days of sample collection. As an alternative to the En Core® sampler, packed jar samples with minimal headspace may also be used for moderately and highly contaminated soil samples. The holding times for soils collected in jars is 14 days.

9.4.2.2 Water Samples

Water samples are collected with precleaned Teflon bailers. The water is drained from the bottom of the bailer into a 40 mL septum capped VOA vial containing at least three drops to 1.0 mL of 12N HCl to achieve a pH of 2. The vial is filled to the top and sealed with the septum cap without any air bubbles in the vial. Samples are then stored at 4°C and shipped according to the EPA protocols. The holding time for water samples is 14 days preserved (as per Table 9.2).

9.4.2.3 NAPL Samples

NAPL samples are collected directly from product dispensers, precleaned Teflon® bailers, or other suitable collection equipment into 25 or 40 mL hardcap VOA vials with minimal headspace. The presence of a headspace in the sample vial is less critical for NAPL samples than for low level water samples. However, the headspace should be minimized to the degree possible. If there is a very small amount of NAPL available (e.g., <10 mL), it may be prudent to minimize the headspace by including groundwater from the same bailer, resulting in a small volume of NAPL floating atop a larger volume of associated groundwater within the VOA vial. The NAPL samples are then stored at 4°C and shipped to the laboratory where the NAPL can be transferred to a smaller vial for long-term storage. Samples shipped at ambient temperatures in sealed vials are also acceptable as there is no prescribed holding time for NAPL samples.

9.4.3 PIANO AND NONHYDROCARBON ANALYSIS BY PURGE-AND-TRAP GC/MS

GC/MS analysis of volatile hydrocarbons and related compounds in soils, sediments, NAPL, and water typically involves the use of purge-and-trap (P&T) as the means of sample introduction. This method is a straightforward adaptation of EPA Method 5035, *Closed-System Purge-and-Trap and Extraction for Volatile Organics in Soil and Waste Samples* and EPA Method 5030, *Purge-and-Trap for Aqueous Sample* for soils/sediments and waters, respectively.

Uhler *et al.* (2003) thoroughly described the adaptations to EPA Method 8260B for use in forensic investigations. The revised analyte list (see Table 9.4) is the primary modification to EPA Method 8260B, however other modifications involving instrumentation (or software) are minimal. Analysis is performed on a programmable gas chromatograph (GC) with flows controlled by an electronic pneumatic control (EPC) system and the detector used is a mass spectrometer (MS). Sample introduction is accomplished using a Tekmar Precept II robotic autosampler and Tekmar 3100 liquid concentrator

system (or equivalent). System control and data acquisition is performed by the integration of both Hewlett-Packard Enviroquant© software (Version 3.0) and Tekmar Teklink© software (Version 5.0). Data analysis and reduction is performed on the Hewlett-Packard Enviroquant© software.

9.4.3.1 Standard Preparation

Because of the complexity of gasoline and other light hydrocarbon materials (e.g., natural gas condensate), quantitative analysis of environmental samples requires accurate identification of each target analyte. Toward this end, it is imperative to incorporate the use of certified, standard grade reference materials and calibration solutions in the modified 8260B method described. These chemical standards can be assembled from various commercial vendors (e.g., Supelco and Restek). Preliminary analysis of standards should be performed to test their purgeability and to optimize chromatographic conditions to establish reproducible retention times based on the methods just described. A calibration solution containing the relevant target analytes listed in Table 9.4 should be prepared as a stock solution in methanol, at a concentration of approximately 5 μg/mL. This stock solution can then be serially diluted to generate a suite of (at least) five linear calibration solutions, with the desired range of analysis for each target analyte to be approximately 0.025 μg–4.0 μg. The use of a calibration solution that contains the gasoline specific target analytes is a significant modification from the EPA Method 8260B, which requires only selected analytes in the calibration (see Table 9.4).

Surrogate and recovery internal standard solutions necessary to monitor purging efficiency are prepared at a concentration of 400 μg/mL in methanol. Surrogate internal standard monitoring compounds for the Uhler *et al.* (2003) method include 1,4-difluorobenzene, chlorobenzene-d_5, and ethylbenzene-d_{10}, and recovery internal standards included benzene-d_6, toluene-d_8 and 1,4-dichlorobenzene-d_4. These internal standard compounds were chosen primarily since they span the full boiling range of the analysis and do not coelute with their nondeuterated analogue. Other internal standards and surrogates may be used as long as the method data quality objectives can be achieved (see Table 9.3).

9.4.3.2 Sample Preparation

Protocols for the handling and preparation of samples after they are received in the laboratory are critical to accurate and precise sample analysis. At a minimum, samples for detailed analysis of volatile constituents should be stored at 4°C from the time of their collection. Product, soil, sediment, and water samples require different preparations prior to analysis, which are described next.

NAPL and Product Samples: Upon receipt at the laboratory, subsamples of dispensed gasolines or NAPL samples are aliquoted and stabilized in a methanol solution. These samples are prepared in an area free of solvent fumes by rapidly weighing approximately 20 mg of the product into a tared 10 mL volumetric flask containing HPLC grade methanol. The resulting diluted extracts (at a concentration of approximately 2.0 mg/mL) are then transferred to, and stored in, 4 mL Teflon© lined screw-capped vials with no measurable headspace, and further protected against losses with Teflon© tape. The samples are then stored at 4˚C (±2˚C) in a dark refrigeration unit free of solvent fumes with the extracts being analyzed within 40 days of stabilizing the sample. In a sealed 40 mL VOA vial, exactly 20 mL of reagent water is spiked through a Teflon© septa with 50 μL of the sample/methanol extract. This vial then is placed onto the autosampler tray, with the instrument fortifying the sample with the necessary surrogate internal standard/recovery internal standard (RIS/SIS) solution.

Water Samples: For water samples, exactly 20 mL of aqueous sample is removed from a 40 mL VOA vial containing the water sample and transferred into a sealed 40 mL VOA vial. This vial then is placed onto the autosampler tray and fortified with the necessary RIS/SIS solution. This method is designed for water samples containing individual purgeable compounds at concentrations of approximately 25 μg/L or less.

Soil and Sediment Samples: Due to the variability in concentration of contaminants in soil and sediment matrices two types of samples most often are prepared in the field for analysis. For soil samples where minimal contamination is expected (e.g., based upon the absence of odor), a low-level method for samples is performed by purging a heated 5 gram soil/reagent water slurry containing both the surrogate and internal standards. If the low-level sample contains analytes that are above the highest calibration standard, a specified volume of a methanol extract of the sample is analyzed. If packed jar samples are collected in the field, the amount of soil in the sample may be adjusted based on the level of contamination. If the expected concentration for target analytes is less than 0.1 mg/kg (dry weight), a 5-g sample is prepared for analysis. If the expected concentrations are between 0.1 and 1 mg/kg (dry weight), a 1-g sample is prepared for analysis. A 40 mL VOA vial containing the appropriate amount of soil/sediment and a Teflon© coated magnetic stir bar is placed onto the autosampler tray. The autosampler then delivers 10 mL of reagent water and RIS/SIS solution to the vial. The spiked soil/sediment sample is then preheated to 40˚C and automatically stirred, forming an aqueous slurry.

If the concentration of the packed jar soil/sediment sample is expected to exceed 1 mg/kg (e.g., they exhibit an obvious hydrocarbon odor), samples are

prepared by the high level method. A 5 gram aliquot of the sample is added to 10 mL HPLC grade methanol and then gently mixed, allowing the methanol to extract target analytes from the soil/sediment samples. A 50–100 μL aliquot (depending on the expected level of contamination) then is removed from the methanol extract and added to a 40 mL VOA vial containing 10 mL of reagent water. This vial is then placed onto the autosampler tray, where the sample is fortified automatically with the necessary surrogate internal standard/recovery internal standards.

9.4.3.3 Sample Analysis

The GC is configured using the capillary split/splitless injection port with a direct interface to the liquid concentrator system for sample introduction. The injection port is sealed with a septum on top of a 2 mm splitless glass liner (Restek or equivalent). The GC is outfitted with a 50 m, RTX-1 PONA 50 m, 0.32 i.d., 0.5 μm film thickness fused silica capillary column (Restek or equivalent). The column provides best performance when plumbed directly to the mass spectrometer, rather than through a jet or other physical separator. The GC/MS conditions employed for volatile organic analysis are provided in Table 9.7.

The autosampler and liquid concentrator system are directly interfaced to the injection port of the GC. The autosampler and liquid concentrator system units are designed to quantitatively strip the sample of volatile analytes by purging with a fine stream of helium. The volatiles and purge gas are then passed through a VOCARB 3000 (Carbopack B/Carboxen 1000 & 1001) sorbent trap (or equivalent) within the liquid concentrator system, which is designed to absorb a broad range of volatile compounds. The instrumental conditions for the liquid concentrator system are listed in Table 9.7.

After the samples are fortified with surrogate and internal standard, the analytical sequence is initiated and includes the automated stirring and purging, trapping, and desorbing of the sample onto the GC column being performed. The samples are then analyzed under the same conditions as the calibration standards. If during the course of the initial analyses, a sample contains a target analyte at a concentration that exceeds the working range of the initial calibration, the sample must be reanalyzed at a more dilute concentration. Experience with purge-and-trap systems has shown that analysis of highly contaminated samples can potentially contaminate the transfer system; after analysis of such a sample, a reagent water blank should be analyzed to confirm that the system is free from interferences and cross-contamination. If the reagent water blank analysis is not free of interferences, the system must be decontaminated by either a bake-out procedure or through more detailed instrument maintenance. Sample analysis should not

Table 9.7
GC/MS and P&T analytical run conditions for the volatile hydrocarbon method. Equivalent conditions may be used if data quality guidelines are achieved.

GC/MS Program Settings		Liquid Concentrator Conditions	
Initial column temperature:	35°C	Purge Temperature:	35°C
Initial hold time:	15 minutes	Purge Time:	11 min.
Program rate 1:	2°C/minute	Purge Flow:	40 mL/min.
Final temp.:	120°C	Desorb Flow:	40 mL/min.
Hold time:	0 minutes	Sample Preheat:	3 min.
Program rate 2:	15°C/minute	Sample Temperature:	40°C
Final temp.:	220°C	Dry Purge:	5 min.
Final hold time:	10 minutes	Desorb Preheat:	245°C
Injector temperature:	220°C	Desorb Time:	2 min.
Column flow rate:	~1 mL/min. (helium)	Desorb Temperature:	250°C
Detector temperature:	280°C	Bake Time:	10 min.
MS scan parameters:	Full Scan Mode; range 35–350 m/z	Bake Temperature:	285°C

resume until a reagent water blank demonstrates that the system is free of interferences.

The modified EPA 8260B described herein employs a calibration solution containing all the target analytes, allowing for the development of compound-specific internal standard-based response factors using the most appropriate ion for each target compound (USEPA, 1997a). Each level of the multilevel calibration contains target analytes and internal standards and surrogates. It should be noted that although each level of the calibration has a different concentration of each target analyte (used to generate an average response factor over the working analytical range), surrogates and internal standards contained in these levels are constant and generally equivalent to that amount spiked into authentic samples. The quality control guidelines are discussed in Section 9.3, and are similar to those provided in EPA Method 8260B. Figure 9.6 shows examples of the raw and processed outputs from this method. The diagnostic compounds (see Table 9.4) measured from the total ion current (TIC) are quantified and represented as a bar graph, which provides a convenient means to compare and contrast the chemical composition or fingerprint among samples (Stout *et al.*, 2006).

The raw TIC chromatograms also provide a visual means to observe weathering differences in light hydrocarbon products such as gasoline (see Figure 9.7).

The P&T GC/MS PIANO analysis provides excellent forensic data for most light product applications. However, for NAPL samples, or for samples containing highly polar organic compounds such as ethanol, direct injection GC/MS may offer some advantages.

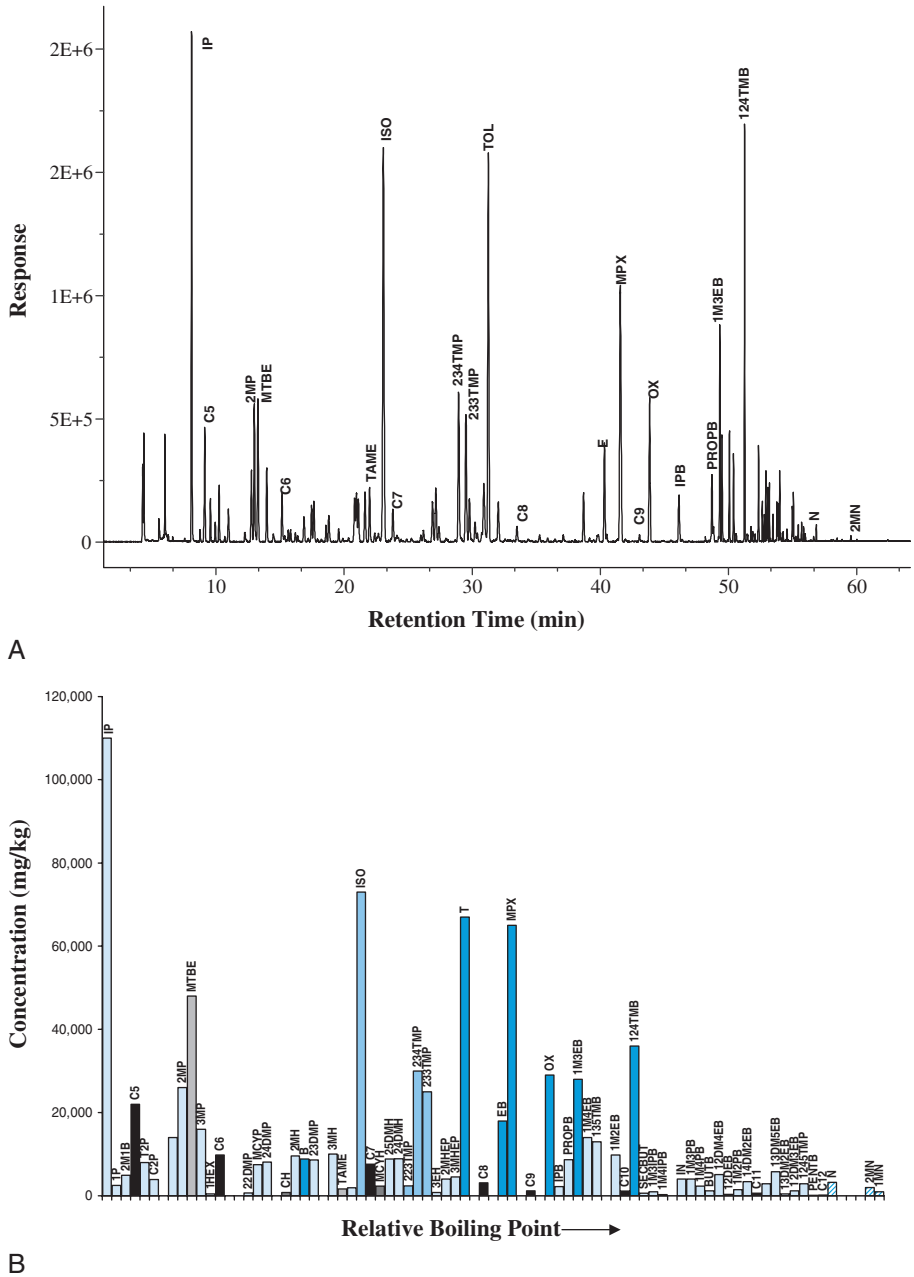


Figure 9.6

A. *The total ion chromatogram and concentration histogram for a premium gasoline sample using the purge and trap GC/MS method.* ***B.*** *The associate compound concentration bar plot. Compound abbreviations as per Table 9.4.*

9.4.4 PIANO AND NONHYDROCABON ANALYSIS BY DIRECT INJECTION GC/MS

Analysis of PIANO compounds, other volatile hydrocarbons, and some semi-volatile compounds in NAPL and water samples can be analyzed without any form of sample preconcentration by direct injection GC/FID (Bruce, 1993)

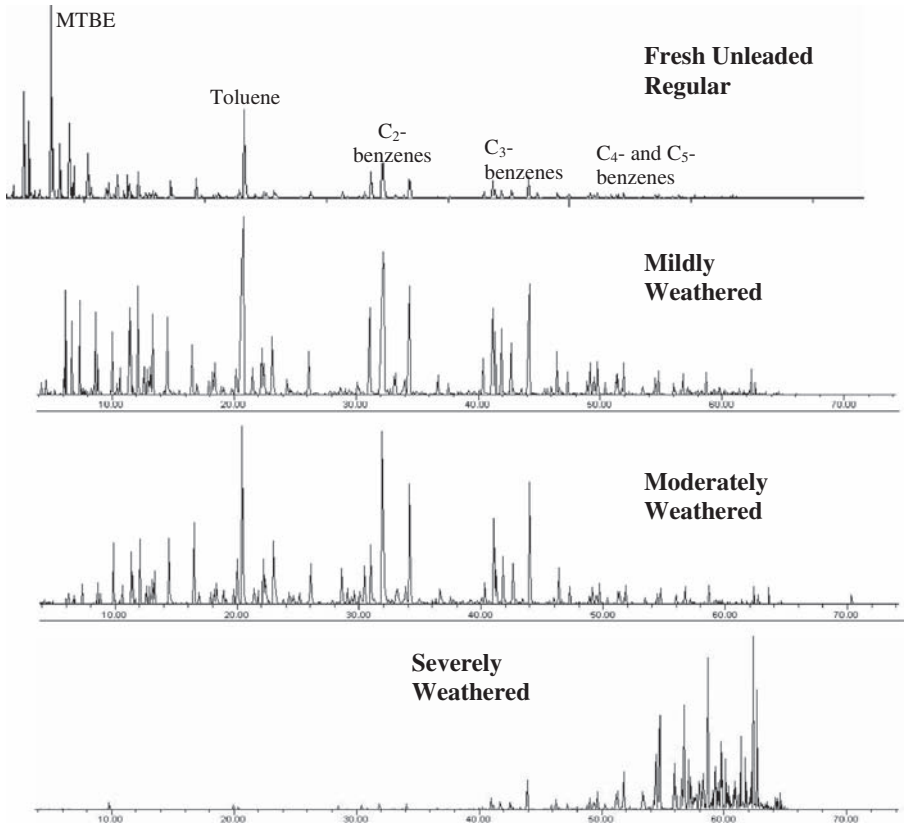


Figure 9.7
Purge and Trap GC/MS TIC patterns showing progressive evaporative weathering of gasoline impacted soils.

and GC/MS techniques. The advantages of this analytical approach relative to the purge-and-trap method discussed earlier are (1) compound purge efficiency is not a performance factor (e.g., ether and alcohol oxygenates), thereby improving analysis of polar chemicals such as alcohols (e.g., ethanol); (2) a broader range of volatile and semivolatile compounds can be measured within a single analysis (e.g., C_{12}+ compounds, diamondoids); and (3) there is less potential for sample carryover (concentrator cross-contamination) when analyzing highly contaminated samples. The main disadvantage of the direct injection method is that the reporting limits for low concentration analytes in water and soil samples is often higher than what can be achieved by purge-and-trap methods. Method sensitivity is generally not a problem in NAPL samples but may be an issue with water samples.

In some instances, microscale solvent extraction of moderately to heavily contaminated water (EPA Method 3511) and soil (EPA Method 3570) samples can help extract and analyze volatile and semivolatile organic compounds (Mauro, 1999, 2006).

Target compounds for this direct injection analysis of NAPLs includes the list of PIANO compounds in Table 9.4, as well as butanes, C_{12} through C_{21} hydrocarbons, diamondoids, alkyllead, and oxygenates in NAPL. Water samples can be analyzed for dissolved-phase chemicals such as alcohols, ethers, and monoaromatics at reporting limits of approximately 0.5 mg/L (Koester, 1999).

9.4.4.1 Standard Preparation

Calibration standards containing PIANO and other target compounds should be prepared by mixing commercially available standards. Standard solutions purchased from a chemical supply house as neat or diluted materials in ampoulated glass vials may be retained for two years from the preparation date, unless the manufacturer recommends a shorter time period. If internal standard response factors are desired, ethylbenzene-d_{10} may be added to the calibration standards and NAPL samples to achieve a 1% v/v concentration.

9.4.4.2 Sample Preparation

When performing direct injection analysis of NAPL, the density of the liquid must be determined so that the target compound concentrations can be reported on an oil weight basis with typical reporting units of parts per million (mg/kg). To determine the density of a sample, 1 mL of NAPL is weighed in a tared vial on a laboratory balance accurate to 0.001 g at approximately 20˚C (ASTM Density Method). Qualitative analysis of soil samples can be carried out by the direct injection method if a 2-g soil sample is microextracted with 10 mL methylene chloride (EPA, 2002d). A 1 μL aliquot of this extract can be analyzed by direct injection.

9.4.4.3 Sample Analysis

Small amounts of NAPL (e.g., 1 μL) or calibration standard (0.5 μL, 1 μL and 2 μL) are injected directly into a gas chromatograph interfaced to a mass spectrometer that is operated in the full scan mode. Variable injection volumes of the calibration standards are used to determine target compound response factors by the method of external standards. If butane, isobutene, and isopentane are not target analytes then diluted (e.g., methylene chloride) calibration standards may also be used at a 1 μL injection volume. Calculation of response factors by the internal standard method (EPA, 1997a) may be performed if ethylbenzene-d_{10} is added to the calibration and NAPL samples prior to injection. Due to the chemical complexity of light hydrocarbon products and therefore the potential for compound coelution, a high resolution capillary GC column is used to ensure optimal chemical compound resolution. A 100 m $\times$ 0.25 mm ID $\times$ 0.25 μm film thickness *Petrocol* DH capillary column (see

Table 9.8
DI GC/MS instrument conditions used for product, NAPL, and microextracted soil samples. Equivalent conditions may be used if data quality guidelines are achieved.

GC Parameter	Method A Program Settings	Method B Program Settings
Column	100 m × 0.25 mm ID × 0.25 μm film thickness *Petrocol* DH capillary column	30 m × 0.25 mm ID × 0.25 μm film thickness DB-5
Initial column temperature:	35°C	35°C
Initial hold time:	5 minutes	2 minutes
Program rate 1:	3°C/minute	4°C/minute
Final temp.:	230°C	310°C
Hold time:	15 minutes	10 minutes
Program rate 2:	4°C/minute	NA
Final temp.:	260°C	NA
Final hold time:	10 minutes	NA
Injector temperature:	250°C	280°C
Detector temperature:	290°C	310°C
Column flow rate:	~1 mL/min. (helium)	~1 mL/min. (helium)
Injector:	Split	Split
Split Ratio:	350:1 NAPL 15:1 Soil Microextraction	400:1
Sample volume:	0.5 μL, 1.0 μL, 2.0 μL	0.5 μL, 1.0 μL, 2.0 μL
MS scan parameters:	Full Scan Mode; range 35–350 m/z	
Electron energy:	70 eV	70 eV
Mass range:	35–350 amu	35–350 amu

Table 9.8, Method A) is recommended. Other GC columns may be used, (e.g., 30 m × 0.25 mm ID × 0.25 μm film thickness DB-5; Table 9.8 Method B). GC conditions for this method that provide suitable compound resolution of volatile hydrocarbons are provided in Table 9.8.

Calibration of the direct injection method is performed by analyzing a standard solution of target hydrocarbons at least three concentration levels (see Table 9.4) by external or internal standard calibration. The response factors (RF) for each target compound are determined from the response of the primary characteristic quantitation ions found in Table 9.4. For target analytes that are not in the PIANO mix, the response factor from an isomer or closest (RT) structurally related compound can be used. Quality control guidelines for this method are provided in Section 9.3.

Examples of GC/MS fingerprints obtained using the direct injection techniques for a dispensed gasoline and weathered gasoline NAPL are presented in Figure 9.8. The associated quantitative concentration bar plots are provided in Figure 9.9.

Both the total ion chromatograms and the associated NAPL concentration histograms allow the chemist to distinguish compositional similarities and differences among samples. Detailed forensic interpretation of direct injec-

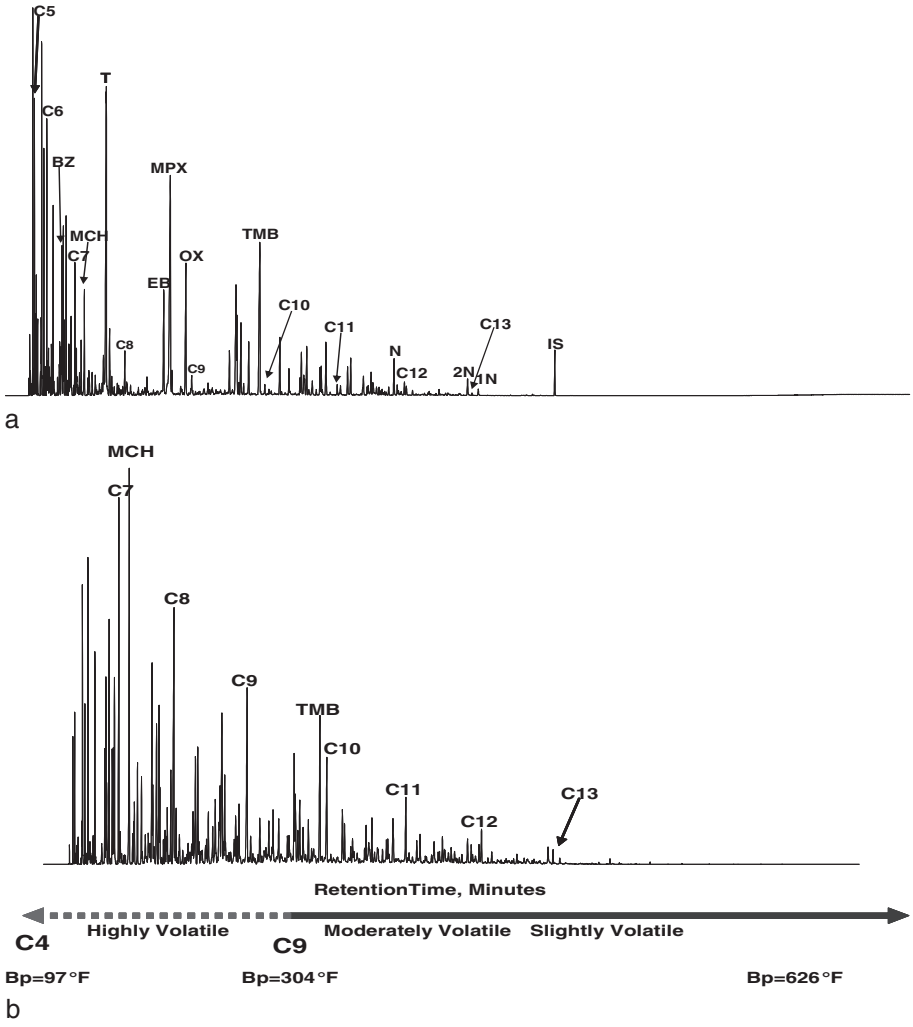


Figure 9.8
*PIANO analysis Total ion chromatograms of (**a**) modern reformate based gasoline and (**b**) weathered leaded gasoline NAPL released prior to 1980. Comparison of the two PIANO analyses highlights the chemical changes that occur in gasoline NAPL due to water washing and evaporation.*

tion GC/MS data is discussed extensively in Beall *et al.* (2002), Stout and Douglas (2004) and Stout *et al.* (2006). In addition to PIANO and oxygenate compounds, with a modification to the GC/MS operating parameters (e.g., selected ion monitoring), direct injection GC/MS methods may also be used to reliably measure gasoline additives such as tetraethyllead (TEL) in dispensed product and NAPL samples.

9.4.5 ALKYL LEAD GASOLINE ADDITIVES BY DIRECT INJECTION GC/MS

Tetraethyllead (TEL) was first introduced into automotive gasoline as an antiknock agent in 1923 (Gibbs, 1990). It was the lone lead alkyl compound

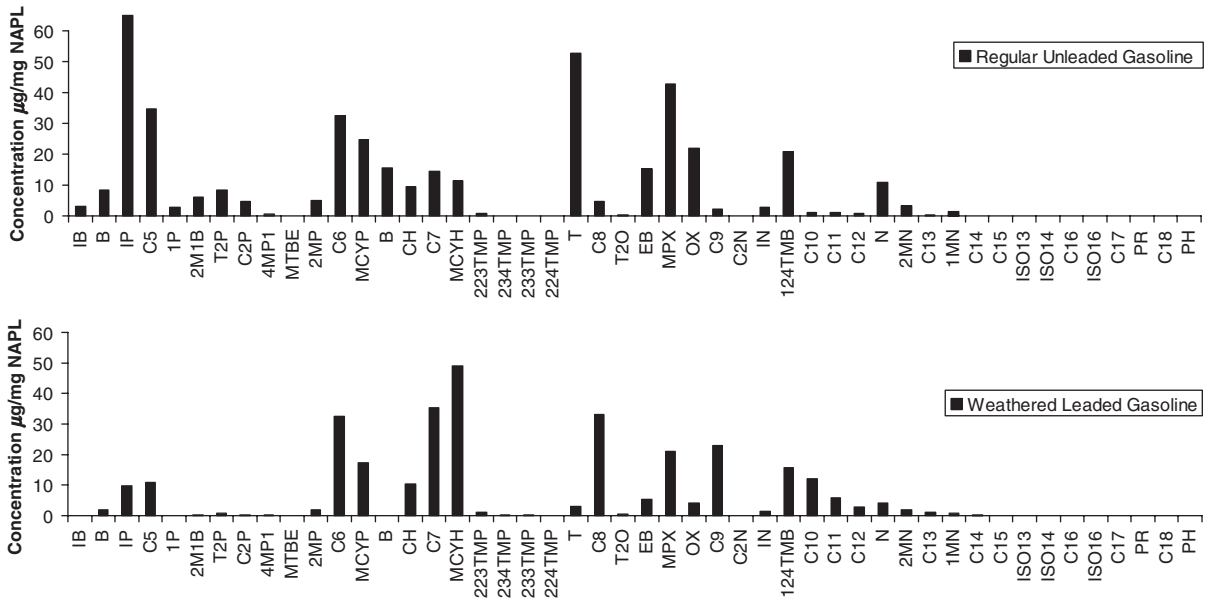


Figure 9.9

Direct injection GC/MS quantitative data for dispensed and weathered gasoline NAPL samples. The loss of soluble aromatics (e.g., toluene) and olefins (trans-2-pentene), combined with the relative enrichment of alkanes (e.g., n-C7) and cycloalkanes (e.g., methylcyclohexane) document the high degree of water washing in the NAPL sample. Compound abbreviations as per Table 9.4.

Table 9.9

Alkyllead compounds found in gasoline.

Analyte	Abbreviation	Primary m/z	Secondary m/z
Tetramethyllead	TML	253	223
Trimethylethyllead	TMEL	253	223
Dimethyldiethyllead	DMDL	267	223
Methyltriethyllead	MTEL	281	223
Tetraethyllead	TEL	295	237

added to automotive gasoline up until the early 1960s. Gibbs (1990) reports that the first commercial use of the other lead alkyls in gasoline occurred in 1960. The five individual alkyl lead compounds that were contained in alkyl lead additive packages (Gibbs, 1990) and their associated quantitation ions are listed in Table 9.9.

The compounds are measured by GC/MS in the selected ion monitoring mode (SIM) following adaptations of EPA Method 8270C. In the analysis of these compounds, free product can be diluted to 10 mg/mL in solvent, fortified with conventional Method 8270C internal standards, and analyzed by GC/MS using standard splitless injection techniques. A multipoint calibration curve containing authentic standards of each of the alkyl lead target com-

GC Parameter	Program Settings
Column – High resolution capillary	DB-5, 0.25 mm × 60 m × 0.25 μm
Initial column temperature:	45°C
Initial hold time:	1 minute
Program rate 1:	6°C/minute
Final temp.:	150°C
Hold time:	0 minutes
Program rate 2:	20°C/minute
Final temp.:	310°C
Final hold time:	10 minutes
Injector temperature:	250°C
Detector temperature:	325°C
Column flow rate:	~1 mL/min. (helium)
Injector:	Splitless
Sample volume:	1.0 μL
MS Scan parameters:	Selected Ion Mode; m/z 223, 253, 237, 267, 281, 295
Electron energy:	70 eV
Mass range:	35–350 amu

Table 9.10
Example of GC/MS conditions for direct injection analysis of alkyllead compounds in NAPL. Equivalent conditions may be used if data quality guidelines are achieved.

pounds should be developed prior to the analysis of field samples covering the concentration range of 5 μg/mL to 500 μg/mL. The gas chromatographic operating conditions used to separate and measure these target compounds are presented in Table 9.10. An example of a total ion chromatogram of an alkyllead standard containing all five of the alkyllead species used in gasoline is presented in Figure 9.10.

The results of such an analysis—typically reported in units of micrograms of individual alkyl lead compound per milliliter of gasoline (ppm)—can be easily converted to units of grams lead per gallon (glpg) if the density of the product is measured. The reporting limit for this method is approximately 0.1 to 1 ppm TEL.

In the United States, pre-1996 (pre-1992 in California) automotive gasoline could contain varying amounts of alkyl lead compounds (Gibbs, 1990). The measurement of these compounds can be exploited to estimate the total lead burden in a fuel and hence, potentially constrain the time of release (Johnson and Morrison, 1996; Kaplan *et al.*, 1997; Stout *et al.*, 1999a) or deduce the source of the gasoline based on the distribution of the individual alkyl lead compounds (Kaplan and Galperin, 1996). The quantitative analysis of alkyl lead compounds is currently restricted to free products, since laboratory and field experiments have demonstrated that, in the absence of free phase gasoline, alkyl leads are strongly adsorbed to soils, precluding their efficient extraction and analysis (Kaplan *et al.*, 1997; Mulroy and Ou, 1998).

Figure 9.11 represents the alkyl lead distribution in two NAPL samples. Both samples contain reactive mixtures of the alkyl leads, thereby constrain-

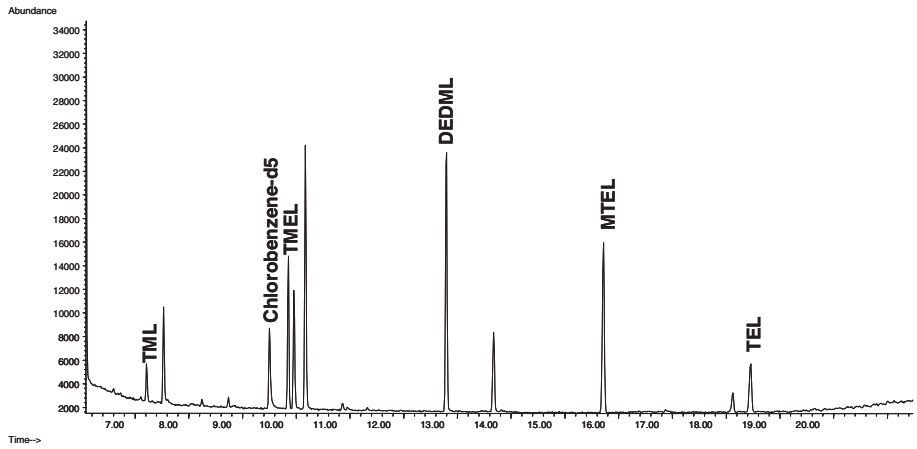


Figure 9.10

Distribution of five alkylated lead compounds (and internal standards) measured by GC/MS (m/z 117 + 223 + 253 + 267 + 281 + 295).

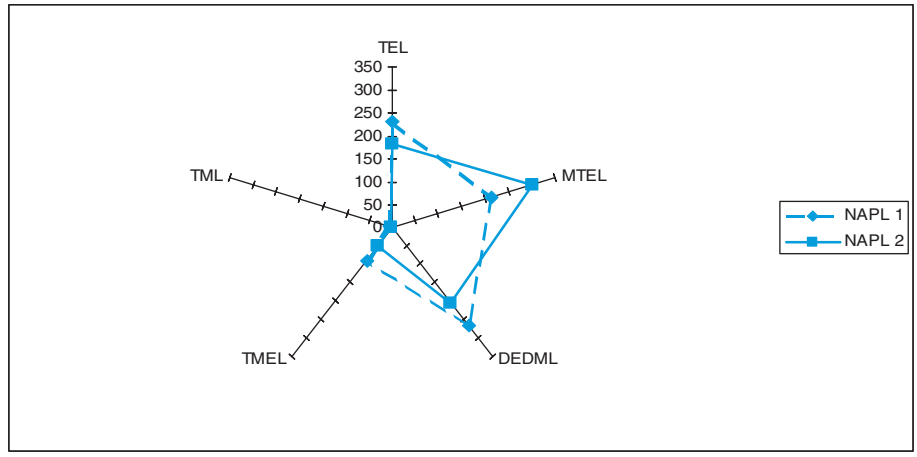


Figure 9.11

NAPL alkyllead speciation analysis by direct injection GC/MS provides data that can be used to constrain the age of the release and resolve chemical differences within NAPL samples.

ing the NAPL release date to between approximately 1960 and mid 1980s (Wakim *et al.*, 1990; Stout *et al.*, 2006). The distribution indicates that the organic lead mixtures within the NAPL samples are derived from different sources and that the samples are chemically distinct.

9.4.6 DIAMONDOID ANALYSIS BY DIRECT INJECTION GC/MS

Diamondoids are volatile saturated hydrocarbons that consist of three or more fused cyclohexane rings, which results in a cage-like or diamond-like structure (Fort and Schleyer, 1964; Fort, 1976)—hence, their class name (see Figure 9.12). The addition of various alkyl side chains (e.g., methyl- and ethyl-groups) yields a series of substituted diamondoids within each parent diamondoid group. For example, addition of a methyl group in a bridgehead or secondary position on adamantane (see Figure 9.12) produces 1-methyladamantane or

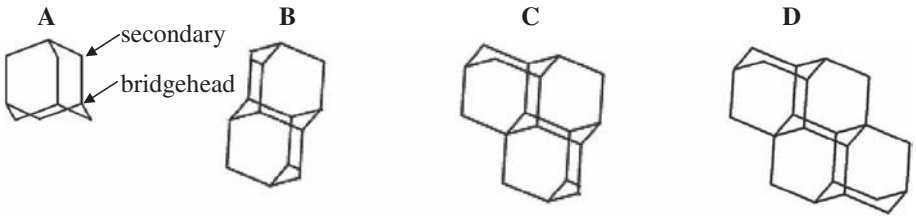


Figure 9.12
Molecular structures for four parent (nonsubstituted) diamondoids; (A) adamantane, (B) diamantane, (C) triamantane, and (D) iso-tetramantane. Hydrogen atoms omitted for clarity.

2-methyladamantane, respectively. Diamondoids have been demonstrated to be quite resistant to even severe levels of biodegradation of reservoired petroleum over geologic time scales (Williams *et al.*, 1986; Grice *et al.*, 2000). By extension, their resistance to biodegradation would be anticipated for diamondoids found in petroleum released (or spilled) in the near-surface environment. As such, the relative abundance and chemical distribution of diamondoids in light petroleum products in the environment offers significant promise as a fingerprinting tool.

The presence and distribution of diamondoids in petroleum generally has been observed using gas chromatograph-mass spectrometry (GC/MS) on the whole oil or aliphatic hydrocarbon fraction of oil. Examples of GC/MS programs used to analyze for these compounds are provided in Table 9.8. The overall stability of these compounds (see earlier) results in predictable (electron impact, 70 eV) mass spectra for parent diamondoids that exhibit a strong molecular ion (M^+) base peak (a feature also characteristic of polycyclic aromatic hydrocarbons), whereas methyl- and ethyl-substituted diamondoids exhibit base peaks of M^+-15 and M^+-29, respectively (Wingert, 1992; Chen *et al.*, 1996). The elution order of the commonly recognized parent and alkyl-substituted adamantanes and diamantanes on nonpolar GC columns is quite well established (e.g., Wingert, 1992; Chen *et al.*, 1996). On a nonpolar capillary column, adamantane is expected to elute just prior to *n*-C_{11} (see Figures 9.13 and 9.14) and diamantane is expected to elute between *n*-C_{15} and *n*-C_{16} (see Figures 9.13 and 9.15). Based upon their elution relative to *n*-alkanes, the boiling points (BP) of adamantane and diamantane are estimated to be 190˚C and 272˚C, respectively (Wingert, 1992). Although chemical standards are available for some diamondoids (*Chiron AS*, Norway), the predictable mass spectral and chromatographic behavior of adamantanes and diamantanes generally allows for confident identification of these compounds under most GC/MS conditions. NAPL and product samples have been analyzed using both direct injection methods discussed earlier with satisfactory results.

The identification of individual adamantanes and diamantanes is determined based upon comparison to published (relative) retention times and full scan mass spectra (e.g., Wingert, 1992). This same information is used to

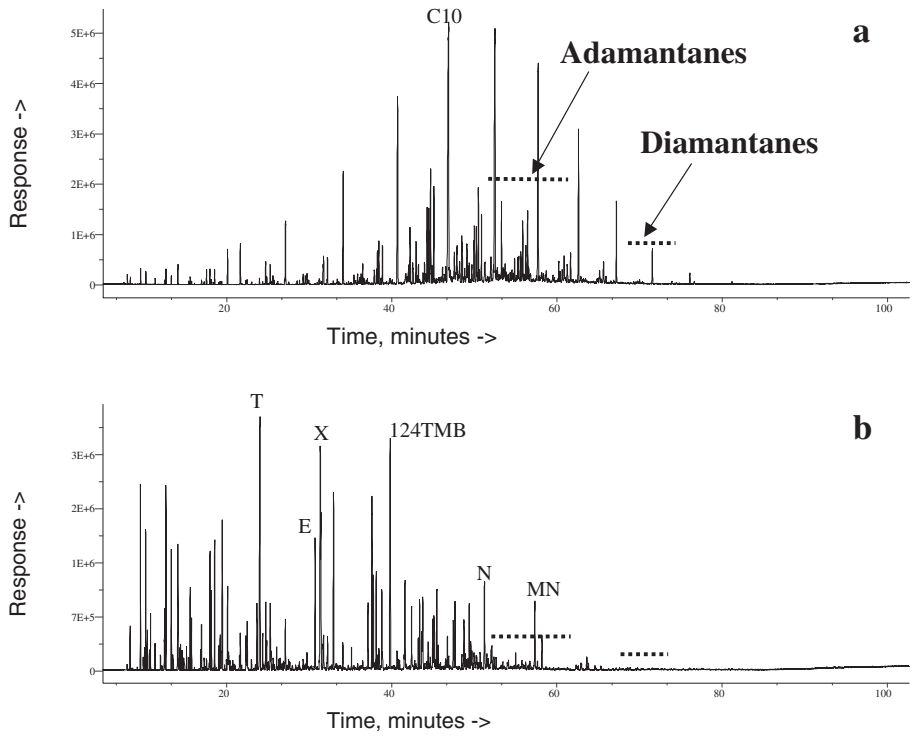


Figure 9.13a, b
*Total ion chromatograms of (**a**) a natural gas condensate and (**b**) gasoline-derived NAPL showing intervals in which adamantanes and diamantanes elute (dashed lines). C10-n-decane, T-toluene, E-ethylbenzene, X-xylenes, 124TMB-1,2,4-trimethylbenzene, N-naphthalene, MN-methylnaphthalenes.*

develop a data processing method that utilized various extracted ion chromatograms to determine relative concentrations (based upon manually integrated peak areas) of the target compounds (or compound groups) listed in Table 9.11. Each target compound's base peak area counts are used to determine the relative concentration.

The retention time range of adamantanes and diamantanes in natural gas condensate (see Figure 9.13a) and in gasoline (see Figure 9.13b) can also be determined by the techniques described here (Stout and Douglas, 2004). This group of chemicals are among the least impacted of the volatile range hydrocarbons by weathering processes (e.g., evaporation, volatilization, biodegradation). An example of extracted ion plots for C0-C4 alkylated adamantanes and C0-C2 alkylated diamantanes relative to the *n*-alkane distribution are provided in Figures 9.14 and 9.15, respectively. These compounds elute in the *n*-C_{11} through *n*-C_{15} carbon range, many outside the analytical range of the P&T GC/MS method.

9.4.7 OXYGENATE ANALYSIS BY DIRECT INJECTION GC/MS IN NAPL

As the EPA mandated the lowering and eventual removal of lead from automotive gasoline (Gibbs, 1990), new classes of octane-boosting additives, with fewer

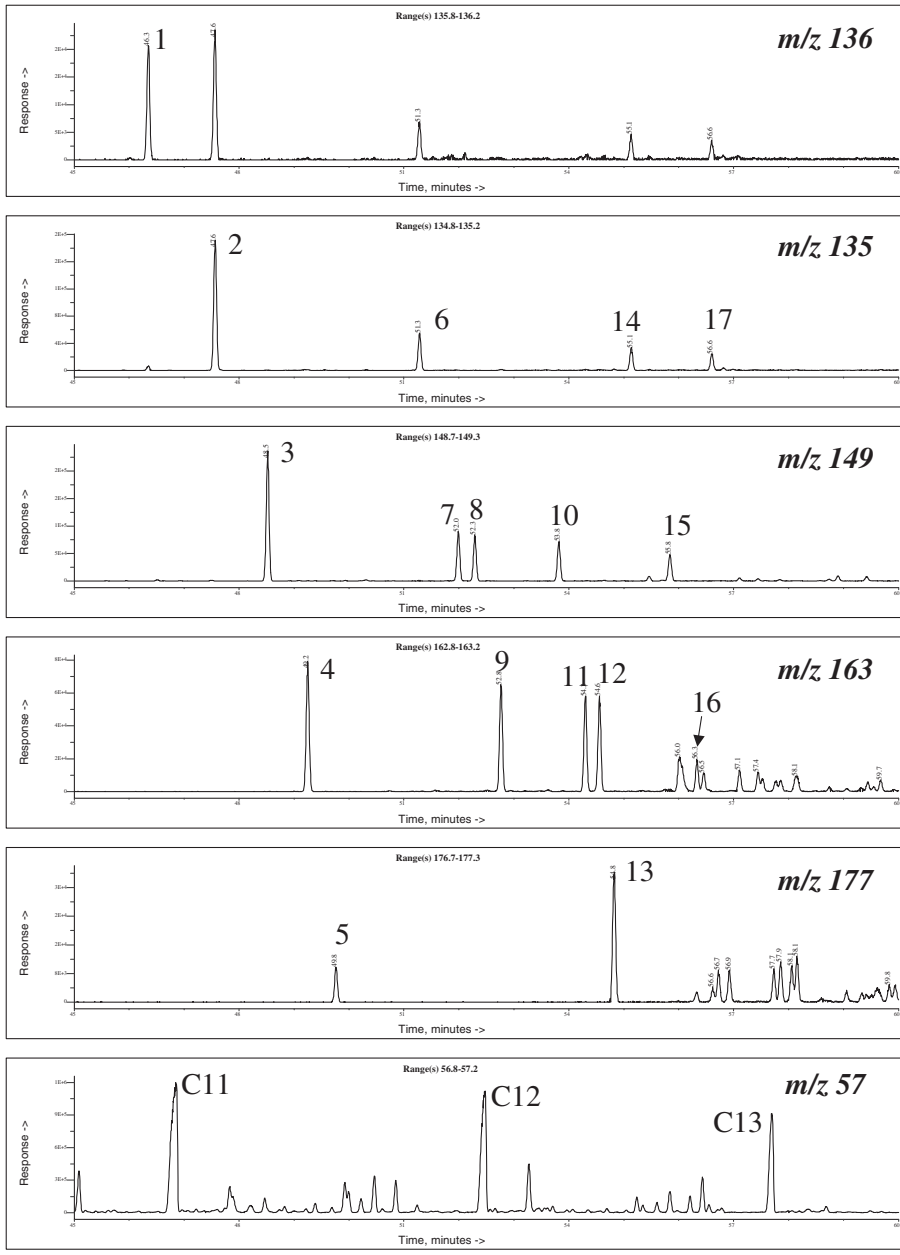


Figure 9.14
Extracted ion profiles of a natural gas condensate showing the C0 to C4 alkylated adamantanes eluting in the n-C11 to n-C13+ range. For peak identifications refer to Table 9.11.

(presumed) negative impacts than lead alkyl, were being developed and used in increasing frequency to augment gasoline performance. The additives included a variety of oxygen-containing additives such as alcohols and ethers.

Alcohols are a broad class of organic compounds containing a hydroxyl (-OH) functional group. Alcohols can be obtained from plant matter or synthetically from petroleum derivatives. Ethers are a class of organic compounds

Table 9.11
Inventory of adamantane, diamantane, n-alkane and alkylbenzene target analytes and key masses used in their identification and (relative) quantification via GC/MS. Peak numbers are used in the figures.

Compound	Pk. #	Formula	M^+	Base Peak
adamantane	1	$C_{10}H_{16}$	136	136
1-methyladamantane	2	$C_{11}H_{18}$	150	135
1,3-dimethyladamantane	3	$C_{12}H_{20}$	164	149
1,3,5-trimethyladamantane	4	$C_{13}H_{22}$	178	163
1,3,5,7-tetramethyladamantane	5	$C_{14}H_{24}$	192	177
2-methyladamantane	6	$C_{11}H_{18}$	150	135
1,4-dimethyladamantane, *cis*	7	$C_{12}H_{20}$	164	149
1,4-dimethyladamantane, *trans*	8	$C_{12}H_{20}$	164	149
1,3,6-trimethyladamantane	9	$C_{13}H_{22}$	178	163
1,2-dimethyladamantane	10	$C_{12}H_{20}$	164	149
1,3,4-trimethyladamantane, *cis*	11	$C_{13}H_{22}$	178	163
1,3,4-trimethyladamantane, *trans*	12	$C_{13}H_{22}$	178	163
1,2,5,7-tetramethyladamantane	13	$C_{14}H_{24}$	192	177
1-ethyladamantane	14	$C_{12}H_{20}$	164	135
1-ethyl-3-methyladamantane	15	$C_{13}H_{22}$	178	149
1-ethyl-3,5-dimethyladamantane	16	$C_{14}H_{24}$	192	163
2-ethyladamantane	17	$C_{12}H_{20}$	164	135
diamantane	18	$C_{14}H_{20}$	188	188
4-methyldiamantane	19	$C_{15}H_{22}$	202	187
4,9-dimethyldiamantane	20	$C_{16}H_{24}$	216	201
1-methyldiamantane	21	$C_{15}H_{22}$	202	187
1,4- and 2,4-dimethyldiamantane	22	$C_{16}H_{24}$	216	201
4,8-dimethyldiamantane	23	$C_{16}H_{24}$	216	201
trimethyldiamantane	24	$C_{17}H_{26}$	230	215
3-methyldiamantane	25	$C_{15}H_{22}$	202	187
3,4-dimethyldiamantane	26	$C_{16}H_{24}$	216	201
undecane	C11	$C_{11}H_{24}$	156	57
dodecane	C12	$C_{12}H_{26}$	170	57
tridecane	C13	$C_{13}H_{28}$	184	57
pentadecane	C15	$C_{15}H_{32}$	196	57
Total C5-alkylbenzenes			148	133

in which an oxygen atom is interposed between two carbon atoms: C-O-C. They both can be made from petroleum derivatives and are widely used as industrial solvents. Alcohols, predominantly *tert*-butyl alcohol (TBA) and ethanol were introduced as a gasoline additive/supplement in the late 1960s and late 1970s, respectively, the latter being primarily in response to the Arab oil embargo. Throughout the 1970s a variety of alcohol blends and ether-based additives were developed, which further extended the gasoline supply, and boosted octane (Guetens *et al.*, 1982). The most volumetrically important of the ethers, MTBE, was first added to commercially available gasolines in the United States in 1979 (Squillace *et al.*, 1995). An inventory of the more common oxygen-containing additives that have been used in gasoline is given in Table 9.12. Volumetrically, MTBE and ethanol are the most commonly used and are the focus of the remainder of this section.

MTBE has been the primary oxygen-containing additive used throughout most of the United States, except in the Midwest where tax advantages have favored the use of ethanol. Reviews of the geographic distribution of the types and concentrations of oxygen-containing additives are available in various

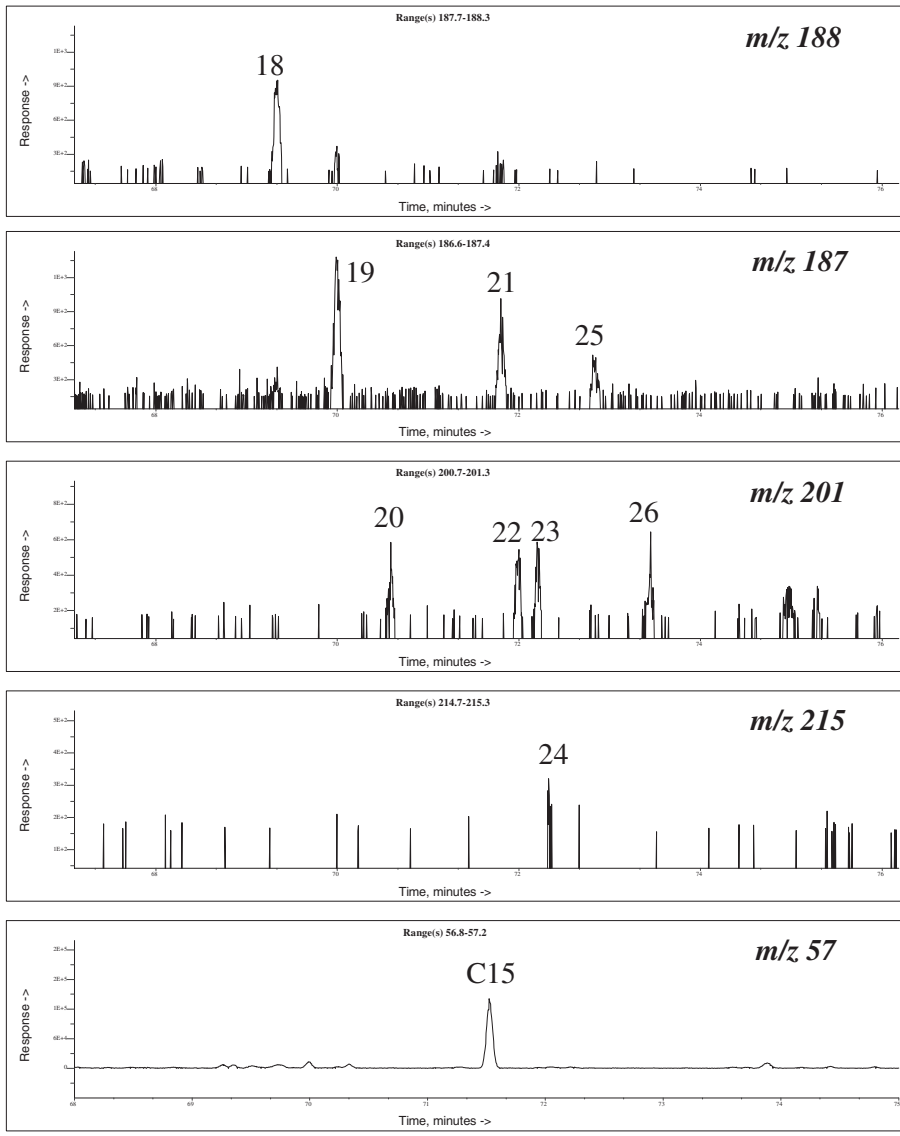


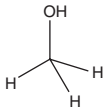
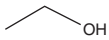
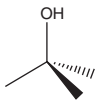
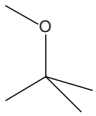
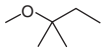
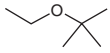
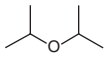
Figure 9.15
Extracted ion profiles of a natural gas condensate showing the C0 to C2 alkylated diamantanes eluting in the n-C15 range. The low signal results from low, but detectable, concentrations of diamantanes in this condensate. For peak identifications refer to Table 9.11.

gasoline databases (e.g., NIPER and U.S. EPA). A review of these survey data has demonstrated that high concentrations (8–11 vol%) of ethanol were preferentially used in most metropolitan areas requiring Oxyfuels (Moran *et al.*, 2000). Similarly, there was good agreement between cities requiring RFG and the use of gasoline containing high concentrations of MTBE (averaging around 9–11 vol%). Conventional gasoline from areas not requiring Oxyfuel or RFG contained lower concentrations of MTBE <2 wt% (on average), most likely solely as an octane booster.

Ethers are analyzed with the PIANO compounds by both P&T and direct injection GC/MS. Tertiary-butyl-alcohol often is measured with the PIANO analytes but exhibits substantial peak broadening and higher reporting limits

Table 9.12

Inventory of oxygen-containing alcohols and ethers and their physical-chemical properties commonly used in gasoline.

Chemical	Formula	Mol. Wt.	R + M /2	Blending RVP	Spec. Grav.	Boil. Pt.	Oxygen wt. %	Oxygen Vol %	H_2O solubility mg/L @ 20°C
ALCOHOLS									
Methanol 	CH_4O	32.0	99.5	4.6 psi	0.792	65°C	50	3.7	miscible
Ethanol 	C_2H_6O	46.0	113	17.6 psi	0.789	78°C	34.8	5.4	miscible
tert-butyl alcohol 	$C_4H_{10}O$	74.1	100	6.8 psi	0.786	71°C	21.6	8.7	soluble
ETHERS									
MTBE 	$C_5H_{12}O$	88.2	109.5	7.8 psi	0.740	55°C	18.1	11	43,000–54,300
TAME 	$C_6H_{14}O$	102.2	104	1.4 psi	0.770	71°C	15.7	12.2	5,500–20,000
ETBE 	$C_6H_{14}O$	102.2	110	4.0 psi	0.742	72°C	15.7	12.7	7,650–26,000
DIPE 	$C_6H_{14}O$	102.2	105	5.4 psi	0.726	68°C	15.7	13.0	9,000–11,250

(approximately 10×) than the ethers. Analysis of alcohols, however, is more difficult using the P&T method due to the miscible nature of ethanol and methanol in water.

The analysis of ethanol and other alcohols in NAPL and dispensed gasoline can also be achieved by direct injection of the product onto a gas chromatograph, followed by flame ionization detection following adaptations of ASTM D 4815, *Standard Test Method for Determination of MTBE, ETBE, TAME, DIPE, tertiary-Amyl Alcohol and C_1 to C_4 Alcohols in Gasoline by Gas Chromatography*. Preconcentration steps can achieve detection limits of approximately 500 ppb (Koester, 1999). GC/MS detection may also provide adequate sensitivity for

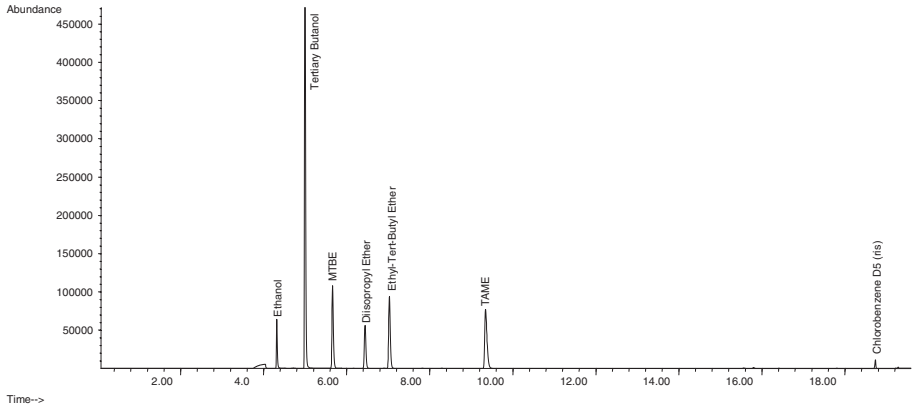


Figure 9.16

Total ion chromatogram of the oxygenate standard (low level) used in the quantification of alcohols and ethers used in automotive gasoline.

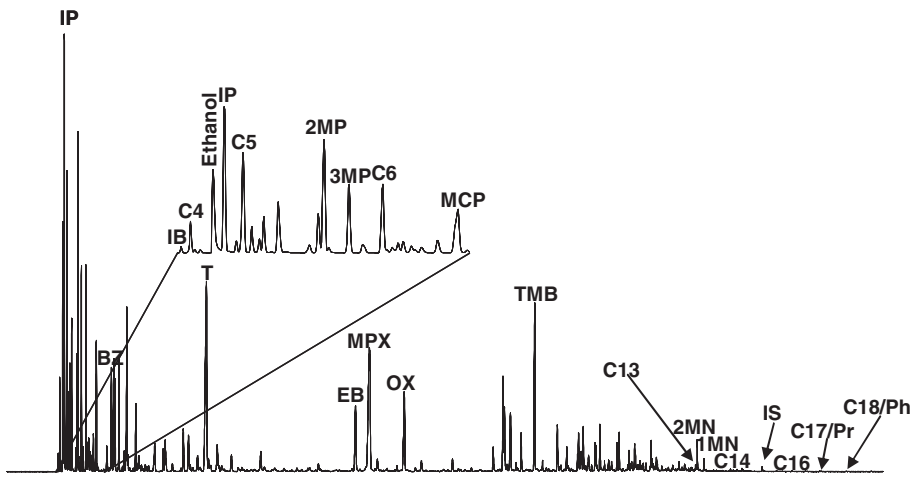


Figure 9.17

Direct injection PIANO and ethanol analysis by GC/MS. Compound abbreviations as per Table 9.4.

this analysis. Figure 9.16 is the TIC for an oxygenated standard analyzed by DI GC/MS; Figure 9.17 is the analysis of an ethanol in gasoline by direct injection GC/MS.

Because of their water solubility, the presence of oxygenates in NAPL may be short lived due to the effects of water washing (see Figure 9.18). Therefore, the presence/concentration of the oxygenate (e.g., MTBE or ethanol) in water underlying NAPL accumulations may provide more useful forensic information relating to the source or, under some circumstances, to the time of release of the gasoline (Stout *et al.*, 2006).

As with the NAPL analysis, conventional purge-and-trap GC/MS methods provide reasonable recovery and analytical sensitivity for the ethers and TBA. Ethanol however, because of its significant solubility, is particularly difficult to analyze in water. Common strategies for the analysis of ethanol in groundwater include heated headspace extraction with GC/FID detection (Watts

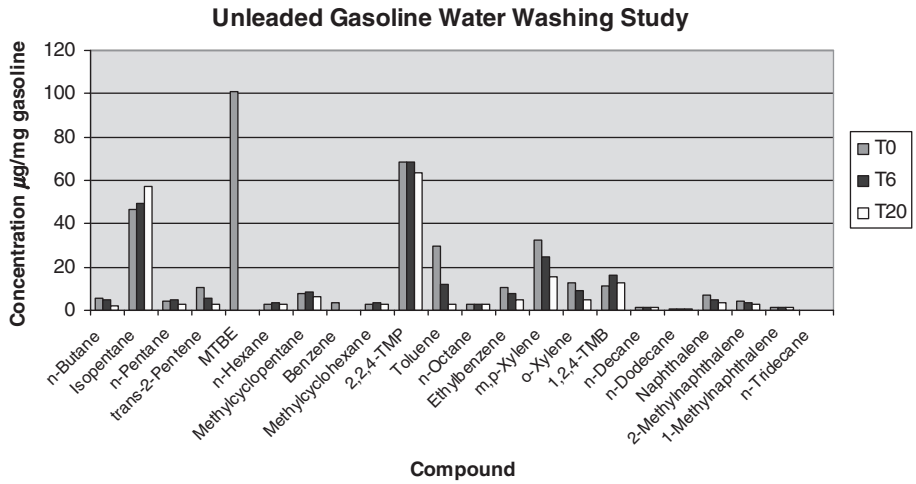


Figure 9.18
PIANO concentration plot for laboratory water washing study documenting the rapid loss of MTBE, benzene, and toluene from a modern gasoline NAPL (Stout et al. *2006). T0 = start of experiment, T6 = 6 months, T20 = 20 months.*

et al., 1990; Macchia *et al.*, 1995; McCarver-May and Durisin, 1997), direct injection with GC/FID detection (EPA, 1997; Tangerman, 1997; Corseuil, 1998), or two-dimensional GC/FID (ASTM, 1997). Additional sample concentration by solid-phase microextraction followed by GC/MS detection has also been used for the analysis of ethanol in water (Koester, 2001). The use of routine contract analytical laboratory methods for the measurement of ethanol in water samples was evaluated by Koester (2001). In their study, a round-robin study of five laboratories was performed with four laboratories using P&T GC/MS and one laboratory using direct injection GC/FID. The ethanol method detection limits for the five laboratory samples varied from a low of 0.005 mg/L (purge-and-trap GC/MS) to a high of 5 mg/L (DI GC/FID). For the working range of the study (0.072 to 11.2 mg/L) the author concluded the good analytical laboratories were capable of achieving ethanol detection limits of between 0.05 to 0.5 mg/L in clean water with purge-and-trap GC/MS. These detection limits were expected to increase in groundwater contaminated with gasoline components.

Although time consuming and costly techniques such as azeotropic distillation (EPA Method 5031) are also candidate techniques for measurement of ethanol and other alcohols in water, direct aqueous injection (DAI), followed by routine GC/FID or GC/MS, can more readily and cost effectively achieve detection limits of approximately 1 mg/L (American Society for Testing and Materials, 1999) to as low as 0.02 mg/L (EPA, 1997a). This same gas chromatographic approach, using a mass spectrometer detector, should give the investigator even more reliable data than the simple GC/FID approach. Emerging techniques, such as solid phase micro extraction (EPRI, 1995) followed by GC analysis, hold promise for even lower detection limits of ethanol and other alcohols in water samples.

9.4.8 ANALYSIS OF OTHER POLAR FUEL ADDITIVES

Many finished automotive gasolines and diesel fuels have been augmented by retailers with pigmented dyes to differentiate among fuel grades or to identify fuels for intended end use (e.g., home heating fuel *versus* automotive diesel Henry *et al*, 1988). There are commercially used dyes—red (azobenzen-4-azo-2-napthol), orange (benzene-azo-2-napthol), yellow (*p*-diethyl aminoazobenzene), and blue (1,4-diisopropyl aminoanthraquinone). In free products, these dyes can be identified quickly by thin layer chromatography (Kaplan and Galperin, 1996; Kaplan *et al.*, 1997). Identifying the presence, absence, or mixture of these dyes can be of forensic utility in cases where one needs to differentiate between or among potentially similar fugitive petroleum products. One potential limitation of reliance on fuel dye data for forensic purposes is the fact that these dye compounds are reportedly unstable under most environmental circumstances (Kaplan *et al.*, 1997). Thus, the absence of dyes (or oxygenates) in a fugitive fuel does not necessarily mean that the compound(s) were not initially present in the fresh fuel.

9.4.9 FULL RANGE WHOLE OIL ANALYSIS

Many petroleum products overlap between the volatile and semivolatile analytical ranges. For example, the direct injection volatile method discussed earlier will detect compounds between *n*-C_4 and *n*-C_{20}, allowing for the detection of a broader range of products beyond the gasoline range (e.g., C_{12}+) such as kerosene, jet fuels, and condensate. However, broader boiling or heavier products such as diesel fuel, fuel oils, crude oils, lubricating oils, coal tars, or asphalt may be only partially detected (e.g., diesel fuel) or completely missed using analytical methods for volatile organics (<*n*-C_{20}). The analytical gap between volatile and semivolatile methods can be closed relatively easily for NAPL samples using a full range, direct injection whole oil analysis described herein. For this analysis a 1 μL volume of the NAPL is injected into a GC/FID or GC/MS using conditions similar to those identified in Table 9.6, Method B. A 30 m × 0.25 mm ID × 0.25 μm film thickness DB-5 (Supleco) chromatography column provides adequate separation for hydrocarbons ranging from *n*-C_4 to *n*-C_{44} (see Figure 9.19) and will detect a full range of petroleum products from gasoline to asphalt.

This method often is performed qualitatively and an analyte specific calibration is not performed. Under these conditions the qualitative data can be used to identify product types and relative concentrations in the samples, for example, by conducting integrations over specific boiling ranges. If a mass spectrometer detector is used the results can be further evaluated by

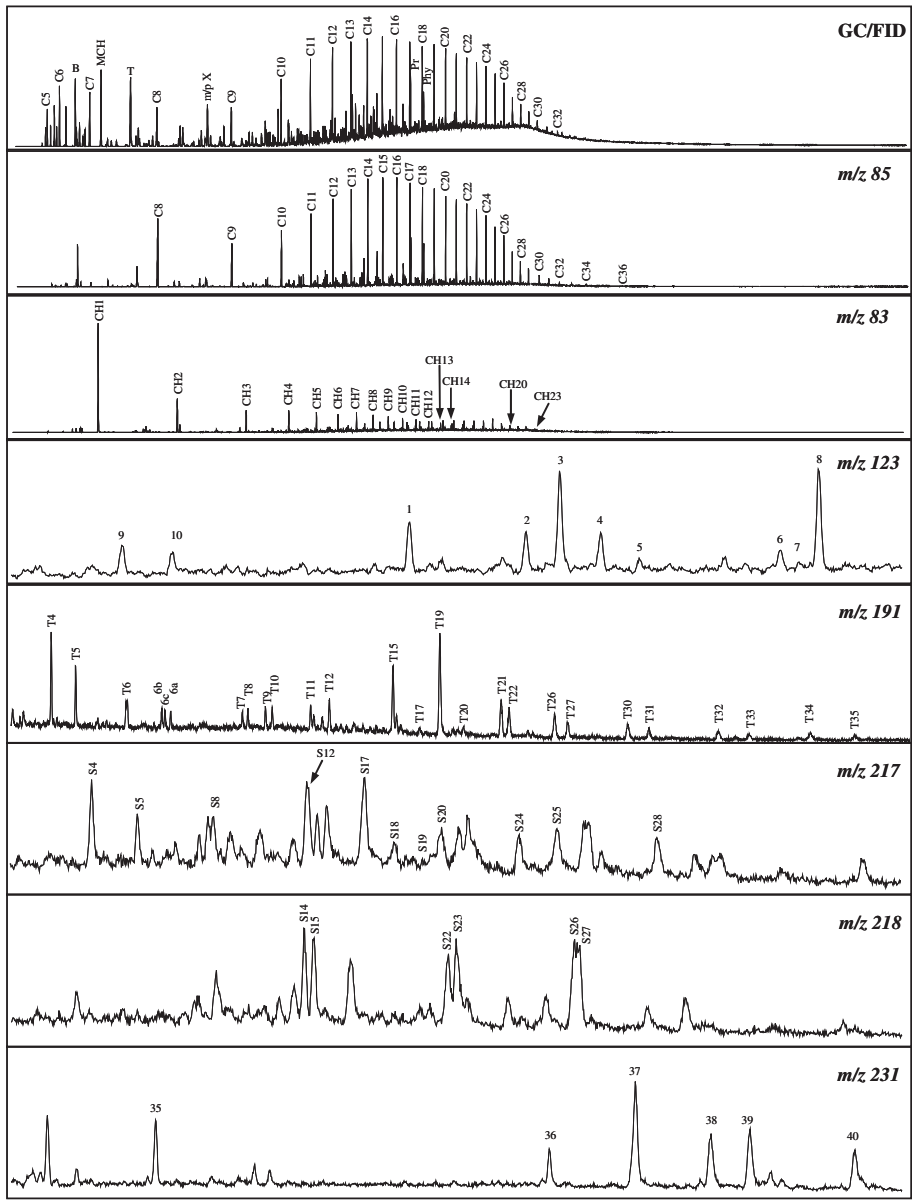


Figure 9.19

Full range (C_4 to C_{44}) whole oil analysis of North Slope crude oil and associate diagnostic extracted ion profiles. m/z 85 = n-alkane and iso-alkane distribution, m/z 83 = n-alkylcyclohexane distribution, m/z 123 = sesquiterpane distributions, m/z 191 = diterpane and triterpane distribution, m/z 217 and 218 = sterane distribution, m/z 231 = triaromatic sterane distribution. The m/z 123, 191, 217, 218, 231 distributions are presented as expanded retention time scales for clarity. These distributions are commonly used to identify product type, source and degree of weathering in petroleum contaminated samples.

examining certain diagnostic ions (e.g., Kaplan *et al.*, 1997; *n*-alkylcyclohexanes *m/z* 83) whose boiling distributions are characteristic of the product type(s), and are represented in Figure 9.19. The major compound classes of interest include the normal and iso-alkanes (*m/z* 85, 113), n-alkylcyclohexanes (*m/z* 83), sesquiterpanes (Stout *et al.*, 2005, *m/z* 123), di-, tri-, tetra-, and pentacyclic triterpanes (Peters *et al.*, 1992; *m/z* 191), regular and rearranged steranes (Peters *et al.*, 1993; *m/z* 217 and 218), and triaromatic steranes (Peters *et al.*, 2005; *m/z* 231; Fig. 9.19). Detailed interpretive applications of these

extracted ion plots are discussed in Kaplan *et al.* (1997), Stout *et al.* (1999b, 2005), Wang *et al.* (1994), Waples and Machihara (1990), Wang *et al.* (2006), Stout *et al.* (2002), and references therein.

Although most commonly conducted on NAPL samples, full range whole oil analysis can be conducted on soil and water extracts if micro-extraction sample preparation methods are used (EPA, 2002c and 2002d, respectively). The analytical protocols for this type of sample preparation and analysis are defined in Mauro and Emsbo-Mattingly (1999) and Mauro (2000). As discussed earlier, the micro-extraction preparation procedures are best suited for heavily contaminated samples due to the relatively high reporting limits associated with the method.

9.5 SEMIVOLATILE HYDROCARBON FINGERPRINTING METHODS

There is no formal EPA definition for semivolatile compounds; however, from a functional standpoint EPA suggests that this class of compounds generally contains chemicals with boiling points above about 150˚C and with vapor pressures of less than 0.1 mm Hg (EPA, 2006). In this section, we consider semivolatile hydrocarbons as those that can be resolved by GC techniques that separate and quantify hydrocarbons that elute between approximately *n*-octane (*n*-C_8) and *n*-tetratetracontane (*n*-C_{44}). Because the boundary between volatile and semivolatile characteristics are simply functional, it is not surprising that there is some overlap in the boiling ranges of volatile and semivolatile chemicals. Thus, some volatile chemicals can be detected by GC techniques that are principally designed for semivolatile analysis such as the whole oil analysis discussed in Section 9.4.9. This is advantageous in hydrocarbon fingerprinting, because lower molecular weight volatile chemicals present in certain hydrocarbon products can be recognized in the early part of their respective semivolatile gas chromatograms (e.g., the alkylbenzene distribution in weathered gasoline; Stout *et al.*, 2006).

Generally, the semivolatile methods of analysis are suitable for characterization of a wide range of hydrocarbon materials such as crude oil, refinery intermediates, and petroleum products like kerosene, diesel, and residual fuel oils. Pyrogenic materials such as coal tars, oil tars, wood tars, and their derivatives such as creosote and pitch are readily characterized by semivolatile GC methods (Novotny *et al.*, 1980). Many other materials are detected by semivolatile GC methods, like some plant waxes, sewage constituents, and residual products of petroleum and tar refining (e.g., road pavement and sealer). The semivolatile range is bracketed on the high molecular weight end of the spectrum by organic constituents, generally heavier than *n*-C_{44}, that are difficult to chromatograph under typical GC conditions. Characterization of very high

molecular weight materials are best accomplished using high temperature gas chromatographic techniques (discussed later in Section 9.5.3.2), or solid state characterization methods such as pyrolysis in combination with gas chromatography (Behar and Pelet, 1984), and/or mass spectrometry, infrared spectroscopy, or nuclear magnetic resonance spectroscopy, which are not the focus of this chapter. The remainder of this section provides a review of semivolatile hydrocarbon fingerprinting methods that are particularly useful in environmental forensic investigations.

9.5.1 SAMPLE PREPARATION

Environmental matrices that most commonly require chemical fingerprinting include free phase products (e.g., oils and tars), soils, sediments, waters, air, wipes, and biological tissues. Depending upon the chemicals of concern—particularly those that are volatile or prone to sequestration in the matrix—the sample preparation methods are a critical component for fingerprinting the sample. Obviously, the robustness of the sample preparation procedure will govern the types of compounds that are removed from these matrices and made detectable during subsequent instrumental analysis. In this section, we provide guidance on the sample preparation techniques according to a tiered analysis of extractable semivolatile hydrocarbons. The specific target analytes are described in the analytical section that follows the sample extraction and cleanup methods.

9.5.1.1 Sample Extractions

With the exception of NAPL samples, chemicals of environmental concern typically are not found in the environment in a form or concentration that can be directly analyzed or detected by modern analytical equipment. Over the last 20 years, analytical chemists have developed specialized techniques for the isolation and concentration of organic chemicals in a range of environmental media. Following are descriptions of proven methods for the preparation of environmental and quality control samples for the analysis of semivolatile hydrocarbons.

9.5.1.1.1 Water Samples

Samples of water are extracted for low-level hydrocarbon analysis using solvent extraction techniques consistent with EPA Method 3510, *Separatory Funnel Liquid-Liquid Extraction* (EPA, 1997). Typically, 1 L of water is added to a 2 L separatory funnel and fortified with known amounts of surrogate QC compounds. The water sample is then serially extracted three times with approximately 60 mL aliquots of dichloromethane (DCM). The DCM extracts are

combined, residual water removed with anhydrous sodium sulfate, and concentrated to 0.5–1.0 mL. The Kuderna Danish (KD) apparatus is used to reduce large volumes of solvent (generally 20 to 200 mL) to a more concentrated volume, and a nitrogen gas evaporation (NEvap) technique is used to slowly reduce smaller volumes of solvent (generally less than 20 mL) to instrumental preinjection volumes typically of 1 mL or less. Regardless of the concentration technique, that final concentration of solvent should be carried out at ambient temperatures in order to minimize the loss of the most volatile compounds in the extract (e.g., naphthalene).

Following final concentration, most water extracts are spiked with an appropriate internal standard and analyzed by appropriate chromatographic techniques. If necessary, extracts with high levels of nontarget analytes that interfere with the chemical analysis can be further purified with specialized sample cleanup methods (EPA, 1997).

9.5.1.1.2 Soil and Sediment Samples

Sediments and soils are primarily inorganic mineral particulates with adhered, entrained, or adsorbed natural and anthropogenic organic materials. Sediments and soils are grouped with other types of particulate matrices (e.g., concrete, stone, asphalt, brick, glass, plastic, metal, and others) classified as solids. In our experience, solid samples are best prepared for analysis using solvent extraction and cleanup techniques promulgated by the NOAA National Status and Trends Program (NOAA, 1993, 1998, 2002). These methods were specifically developed for the efficient extraction of highly sequestered organics in complex media, like sediments and tissue, that require active particle motion and long solvent exposure conditions. Extraction efficiency for semivolatile compounds using these methods is generally good for labile and residual compounds present in the semivolatile hydrocarbon range.

Large solid samples (e.g., concrete, asphalt) are subjected to particle reduction (e.g., crushing, drilling, or cutting). Free liquid typically is decanted out of the sample jar. Extraneous plant debris (twigs, large roots, branches, and leaves) and large stones are removed from the sample. The remaining solid sample is homogenized and an aliquot (5 to 10 g wet weight) is removed for the determination of the sample dry weight. Approximately 30 g (wet weight) of sample is transferred to an extraction vessel (e.g., precleaned and disposable 8 oz jars with Teflon cap liner). A smaller sample weight can be used if gross hydrocarbon contamination is present or a limited sample is available. Surrogate standards are spiked into the sample for the purpose of monitoring the extraction and sample processing efficiency and should be adjusted to the extent possible for the estimated level of contamination in the sample (low, moderate, high).

Anhydrous sodium sulfate is mixed into the solid sample to absorb excess water from the sample and facilitate sample extraction with the organic solvent. It is recommended that activated copper be added to sediment samples for the removal of excess elemental sulfur in marine sediment samples. An organic extraction solvent, such as dichloromethane (DCM) is added to the extraction vessel before it is sealed and serially extracted three times (30–60 mL). In conjunction with the motion of the particles, the duration of the extraction is important because it facilitates the mobilization of tightly bound compounds. The first of the three serial extractions is shaken or tumbled for a minimum of 12 hours (1X), whereas the second and third extractions are carried out for at least one hour. This high energy extraction technique agitates the particles over a long period of time in order to efficiently extract target analytes sequestered in the recess of the sediment matrix. The sample is centrifuged between the extractions to optimize solvent recovery and decanted and collected into a precleaned Erlenmeyer flask. The combined extract is then filtered and dried through a glass fiber filter containing anhydrous sodium sulfate to remove particles and water from the extract. The KD apparatus is used to remove large volumes of solvent (generally 20 to 200 mL), and the NEvap technique is used to remove smaller volumes (generally less than 20 mL) of solvent at a slow evaporation rate and at ambient temperature or less to the final preinjection volume (e.g., 1 mL). At this point solid sample extracts other than soils and sediments are spiked with an appropriate internal standard (IS) and analyzed. Complex sample extracts (e.g., soil and sediments) with high levels of nontarget analytes can be further purified (e.g. GPC) or fractionated (e.g., aliphatic and aromatic hydrocarbons) with specialized sample cleanup methods to improve method detection limits.

9.5.1.1.3 Tissue Samples

Tissue samples are prepared for analysis using solvent extraction and cleanup techniques followed by the NOAA National Status and Trends Program (NOAA, 1998). These methods were developed specifically for the efficient extraction of highly sequestered organics in complex media, like sediments and tissue, that require active particle motion and long solvent exposure conditions.

Tissue samples are collected (e.g., fillet, shucked, cut into pieces) and transferred into a precleaned extraction vessel container and thoroughly homogenized using a Tekmar Tissuemizer. Fish may be filleted and composited prior to extraction (EPA, 1993b). Approximately 30 g (wet weight) of homogenized tissue is used for solvent extraction and an additional 5 g to 10 g (wet weight) is used for dry weight determination.

Anhydrous sodium sulfate is added to the homogenate to absorb water and facilitate extraction. The tissue homogenate is mixed with DCM and macer-

ated three times, each for two minutes, with a Tissuemizer. The sample is centrifuged between extractions to optimize solvent recovery and the solvent is decanted and collected in a precleaned Erlenmeyer flask. The combined extract is filtered and dried through a glass fiber filter containing anhydrous sodium sulfate. A KD apparatus is used to concentrate the large volume of extraction solvent (generally 20 to 200 mL) to approximately 10 mL, followed by further concentration by NEvap to a precleanup volume of approximately 1 mL. At this point the tissue extracts are subjected to sample cleanup procedures to remove coextracted biogenic materials that interfere with instrumental analysis.

9.5.1.1.4 Nonaqueous Phase Liquid (NAPL) Samples

Nonaqueous phase liquid samples such as petroleum or tar products can be prepared for analysis following an adaptation of EPA Method 3580, *Waste Dilution* (EPA, 1997). The NAPL samples are generally liquid, but can be solidified hydrocarbon residues that dissolve readily in DCM. An aliquot of the NAPL sample is quantitatively transferred into a dilution vessel (e.g., 10 mL scintillation vial with Teflon lined cap). Dichloromethane is added to the volumetric and the sample is mixed or vortexed until the NAPL dissolves. The NAPL extract is filtered through a glass fiber filter into a 10 mL volumetric flask. The NAPL extract is spiked with surrogate and brought up to the calibrated volume. An aliquot of the extract is removed to determine the sample mass or oil weight. Liquid hydrocarbon samples (e.g., diesel fuel, crude oil) may also be analyzed directly without sample dilution and cleanup using direct injection GC/MS methods, also known as whole oil analysis (see Section 9.4.9). For these analyses the density of the NAPL must be determined to calculate the weight of the sample that is injected on the instrument. Internal standards may also be added to the NAPL prior to instrumental analysis.

9.5.1.1.5 Wipe Samples

Sorbent wipes can be used to effectively collect organic residues off various kinds of environmental or structural surfaces, or from floating "sheens" of contaminants from water. The best wipe material for trace level measurements is prerinsed and dried Teflon netting or glass fiber filter. This material is used routinely by the U.S. Coast Guard to collect oil sheen samples on surface waters (Plourde *et al.*, 1995). During emergency situations, when it is better to collect a sample than not, other materials such as cotton pads or paper towels have been used for the collection of grossly contaminated surfaces with the understanding that the use of these materials may contaminate the sample with mineral oil, antibiotic agents, and other interferences. Thus, blank wipes should be analyzed in order to appreciate the potential contribution of the sampling material to the authentic sample(s). Wipes are a very

useful replacement for equipment rinse blanks after decontamination in the field because they do not require cumbersome bottles of purified water and they can be processed as solid samples in the laboratory.

Surface wipes usually employ a Teflon swab (or equivalent) to wipe a specific surface area (e.g., 10 cm × 10 cm) (EPA, 1998). The analyte concentration can be reported in concentration per unit area, for example, $\mu g/100\ cm^2$. Sheen samples on water are collected with a Teflon swab or net that is placed or drawn through a floating organic layer. Wipes containing adsorbed, but nonflowing NAPL, can be spiked with surrogate standards and the serially extracted with organic solvent as described earlier for solid samples. In the case of wipes that contain free flowing NAPL, remove an aliquot of the hydrocarbon liquid, and process as a NAPL (described earlier).

The raw Teflon material used to construct wipe samplers usually contain low levels of hydrocarbon oligomers, that if used untreated, can interfere in GC analysis of hydrocarbons extracted from such materials. All Teflon® wipe sampling material must be rigorously precleaned with methylene chloride, stored in precleaned glass jars, and blank tested prior to use. It is not recommended that these devices be stored for long periods of time prior to use. In addition, a field blank of the wipe should be provided to the laboratory during each sampling event.

9.5.1.2 Extract Cleanup and Purification

Complex environmental samples like highly contaminated or organic-rich soils, sediments, and biological tissues contain numerous compounds that coextract with target analytes that can interfere with the chemical analysis. Without aggressive sample cleanup procedures to remove these coextracted, nontarget compounds, it is often impossible to detect and quantify target compounds at environmentally significant concentrations. Thus, over the last 20 years, supplemental sample extract cleanup techniques have been developed to remove substances that obscure or bias the measurement of target analytes. This section describes some effective techniques for purifying the sample extracts and improving the quality of the target analyte measurement in complex media. These cleanup steps are direct adaptations of those described by EPA in SW-846 (EPA, 1997) and the National Oceanic and Atmospheric Administration (NOAA) in the National Status and Trends Program Analytical Methods (NOAA, 1998) and The Total Petroleum Hydrocarbon Working Group (TPHCWG) methods (Douglas *et al.*, 2001).

9.5.1.2.1 Alumina Solid Phase Adsorbent

Alumina cleanup can be accomplished following adaptations of EPA Method 3611 (EPA, 1997). Alumina is a solid sorbent used to separate (and thereby

purify) target chemicals of interest from organic materials that are coextracted from various kinds of environmental samples. The alumina cleanup step is recommended for all tissue samples and most sediment samples that contain extractable organic matter (i.e., with organic carbon loadings greater than 0.5%). This cleanup step is also useful for water sample extract cleanup if high levels of polar organic matter (e.g., organic acids) are present.

Alumina is a porous and granular form of aluminum oxide that generally is prepared at neutral pH and used in its dry-activated state for the purification of hydrocarbon extracts. In practice, a chromatography column fitted with a polytetrafluoroethylene (PTFE) stopcock is prepared with a glass wool plug below a 10-g bed of alumina and topped with 1 cm plug of anhydrous sodium sulfate powder. The column is rinsed with DCM, loaded with the extract, and eluted with 30 mL of DCM. For hydrocarbon analysis, a nominal 10 mg of extractable hydrocarbon is the proper charge for optimal alumina column cleanup. The DCM elutriate contains both aliphatic (F1) and aromatic (F2) hydrocarbons. The resulting 30 mL elutriate is concentrated by NEvap to a preinjection volume of approximately 0.5–1.0 mL for instrumental analysis.

Sediment and tissue extracts typically contain higher proportions of lipid and moderately polar nontarget chemicals relative to water and soil samples. As such, a larger amount of alumina can be used to facilitate cleanup of sediment and tissue extracts. In these cases, a gravimetric measurement of the organic residue in the crude extracts is determined to ensure that no more than 10 mg of material are processed through a 20-g alumina column in order to obtain a purified aliphatic (F1) plus aromatic-unsaturated (F2) hydrocarbon fraction. The larger column is eluted with 100 mL of DCM. The elutriate is concentrated by KD and NEvap technique as described in the sample preparation method. If needed, the resulting alumina-purified extract can be subjected to additional cleanup by gel permeation chromatography or other cleanup prior to instrumental analysis. Calibration of the appropriate volumes of eluant for separation of aliphatic, aromatic, and polar compounds should be performed whenever a new batch of alumina is used for sample cleanup.

9.5.1.2.2 Gel Permeation Chromatography (GPC)

Sediment and tissue extracts benefit from an additional cleanup using high performance size exclusion gel permeation chromatography (GPC) following procedures described by NOAA's National Status and Trends Program methods (NOAA, 1998). The GPC cleanup step is recommended for all tissue and most heavily contaminated sediment samples. GPC is a size exclusion technique that separates macromolecules (e.g., biogenic compounds like lipids, polymers, proteins, resins, and cellulose) from the lower molecular weight chemicals like anthropogenic contaminants (e.g., hydrocarbons, synthetic pesticides,

and PCBs). The liquid chromatography system is fitted with a guard column (7.8 mm ID × 5.0 cm) and two tandem columns (22.5-mm ID × 25 cm) packed with Phenogel 10-μm particles of 100 Å pore size (Phenomenex, Torrance). The column is eluted isocratically (constant mobile phase) with DCM. The system response is monitored with a 253 nm ultraviolet detector. The system is precalibrated with a mixture of corn oil and perylene; the anthropogenic target chemicals of interest elute between these two marker compounds. The resulting elutriate from the GPC cleanup is concentrated by KD and NEvap to a final volume of 0.5–1 mL for instrumental analysis.

9.5.1.2.3 Sulfur Removal via Copper Granules

Sediments and some soil samples contain significant amounts of extractable elemental sulfur and/or sulfide compounds that interfere in the gas chromatographic analysis of semivolatile and halogenated organic compounds. Sulfur can be removed from sample extracts by treatment with activated copper granules following adaptations of EPA Method 3660 (EPA, 1997a) or NOAA National Status and Trends Program methods (NOAA, 1998).

Granular copper (99% pure, or equivalent) is activated by treatment with 6N hydrochloric acid. The sediment or soil extract, processed through adsorbent or GPC column cleanup, is added to a glass vial containing approximately 2 g of copper powder and shaken for 15 minutes. The mixture is allowed to stand for a minimum of two hours. A small amount of freshly activated copper is then added to the extract to determine if the reaction is complete. If the copper does not change from a bright orange to dull black, the extract is quantitatively separated from the copper powder and concentrated by NEvap to a final volume of 0.5–1 mL for instrumental analysis.

9.5.1.2.4 Silica Gel Solid Phase Adsorbent

In some specialized situations, such as certain biomarker analyses, it may be advantageous to separate (i.e., fractionate) aliphatic hydrocarbons from aromatic hydrocarbons in sample extracts prior to instrumental analysis. In these cases, such separation can be accomplished using silica gel solid phase chromatography following adaptations of EPA Method 3630 (EPA, 1997). Silica gel is an adsorbent made from sodium silicate and sulfuric acid. The sorbent material (100/200 mesh silica gel) is prepared by extraction with DCM and oven drying for a minimum of 16 hours at 130°C. It is advantageous to store the prepared silica gel in an oven at approximately 105°C prior to use.

A chromatography column fitted with a polytetrafluoroethylene (PTFE) stopcock is prepared with a glass wool plug below a 5-g bed of silica and topped with 1 cm plug of anhydrous sodium sulfate powder. The sample extracts are solvent exchanged to hexane. The prepared fractionation column is rinsed

with pentane, loaded with the extract, and eluted with 20 mL of pentane to isolate the aliphatic hydrocarbons (F1 fraction). Optionally, the column can be eluted with 25 ml of DCM to isolate the aromatic hydrocarbons (F2 fraction). The specific elution volumes can be adjusted based on analyte elution experiments. For hydrocarbon analysis, it is our experience that a nominal 10 mg of extractable hydrocarbon is the proper charge for optimal silica column cleanup. Higher masses can be used if aromatic breakthrough can be tolerated. The resulting elutriate is concentrated by NEvap to a preinjection volume of approximately 0.5–1 mL for instrumental analysis. A fractionation calibration standard composed of aliphatic and aromatic hydrocarbons should be analyzed with each analytical batch to monitor column separation efficiency (Douglas *et al.*, 2001).

9.5.2 SAMPLE ANALYSIS

Environmental samples that are the subject of environmental forensics investigations are routinely analyzed by several determinative methods in order to provide a comprehensive understanding of the complex hydrocarbon assemblages that compose the samples. In general, high resolution gas chromatography methods using flame ionization or mass spectrometer detectors are used to develop synoptic fingerprints of the dominant hydrocarbons and product type(s) that compose a sample. Quantitative gas chromatography/mass spectrophotometry is used to quantitatively measure important hydrocarbon chemicals that make up petroleum, tar, and combustion-derived residues in environmental samples (Stout *et al.*, 2002; Emsbo-Mattingly *et al.*, 2003a, 2003c). Each of these methods is described in greater detail in this section.

9.5.2.1 High Resolution Hydrocarbon Fingerprinting by GC/FID

The production of gas chromatographic fingerprints and measurements of total petroleum hydrocarbons (TPH) and individual hydrocarbon chemicals of samples generally is accomplished using a gas chromatograph equipped with a flame ionization detector (GC/FID). This method is designed to resolve a broad range of hydrocarbons (i.e., *n*-C_8 to approximately *n*-C_{44}) such that fugitive hydrocarbon products can be recognized (see Figure 9.20). Considerable effort is required to generate high resolution hydrocarbon fingerprints with an adequate degree of resolution to differentiate clearly anthropogenic products from each other and from naturally occurring hydrocarbons. The method should also include QC samples that document that mass discrimination (see later) across the target hydrocarbon range of interest has been adequately controlled.

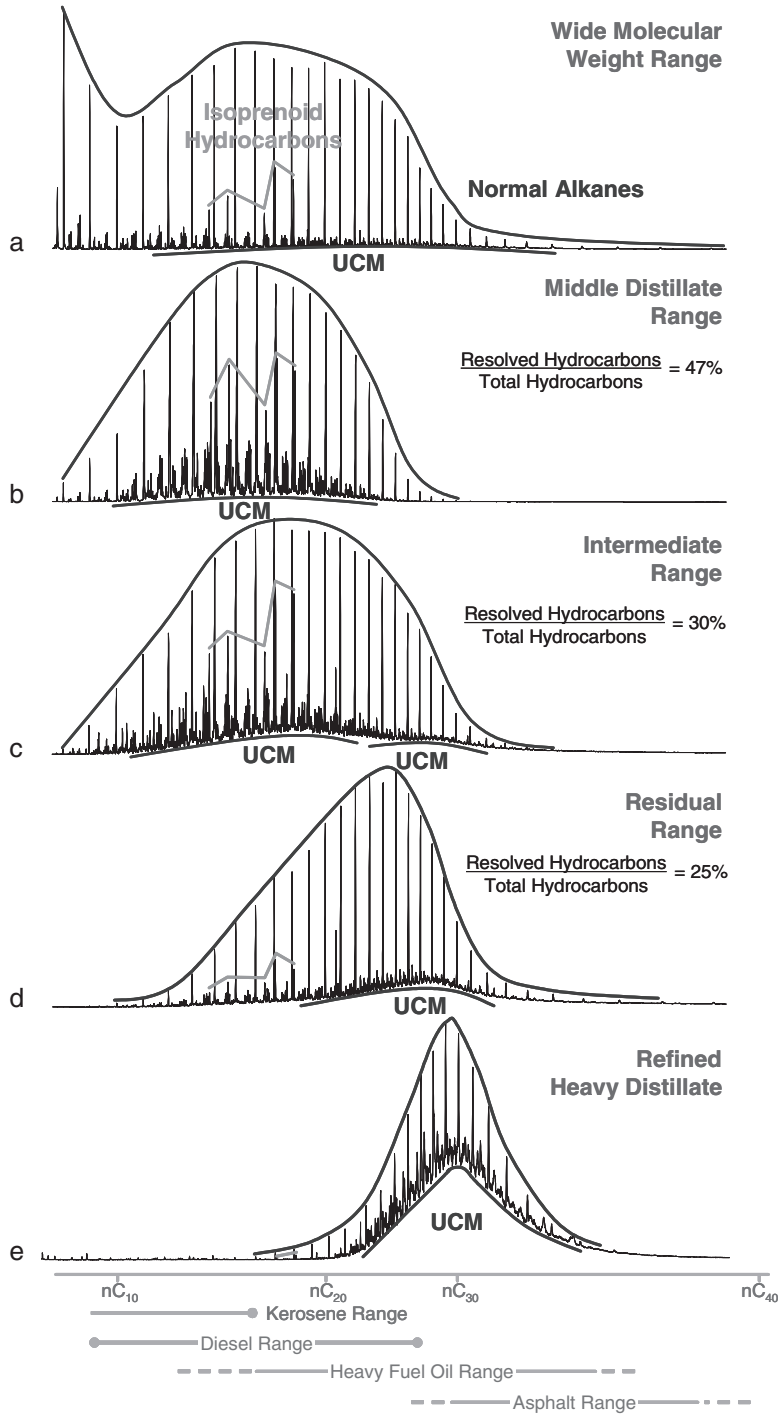


Figure 9.20
Selected high resolution hydrocarbon fingerprints:
a. *crude oil*
Crude oil is the primary feedstock for a petroleum refinery. This complex hydrocarbon mixture is the source of many petroleum derived materials.
b. *#2 fuel oil*
After distilling light hydrocarbons for gasoline, middle distillates, like this #2 fuel oil, are distilled from the crude oil feedstock.
c. *#4 fuel oil*
Intermediate fuel oils are commonly blended to include middle distillates and heavy fuel oils.
d. *#6 fuel oil*
Heavy fuel oils fall into the residual range of crude oil. They may be composed of heavy distillates that are generated after middle distillates are removed.
e. *Lubricating oils are manufactured from heavy distillates that are subsequently purified and fortified with performance enhancing agents.*

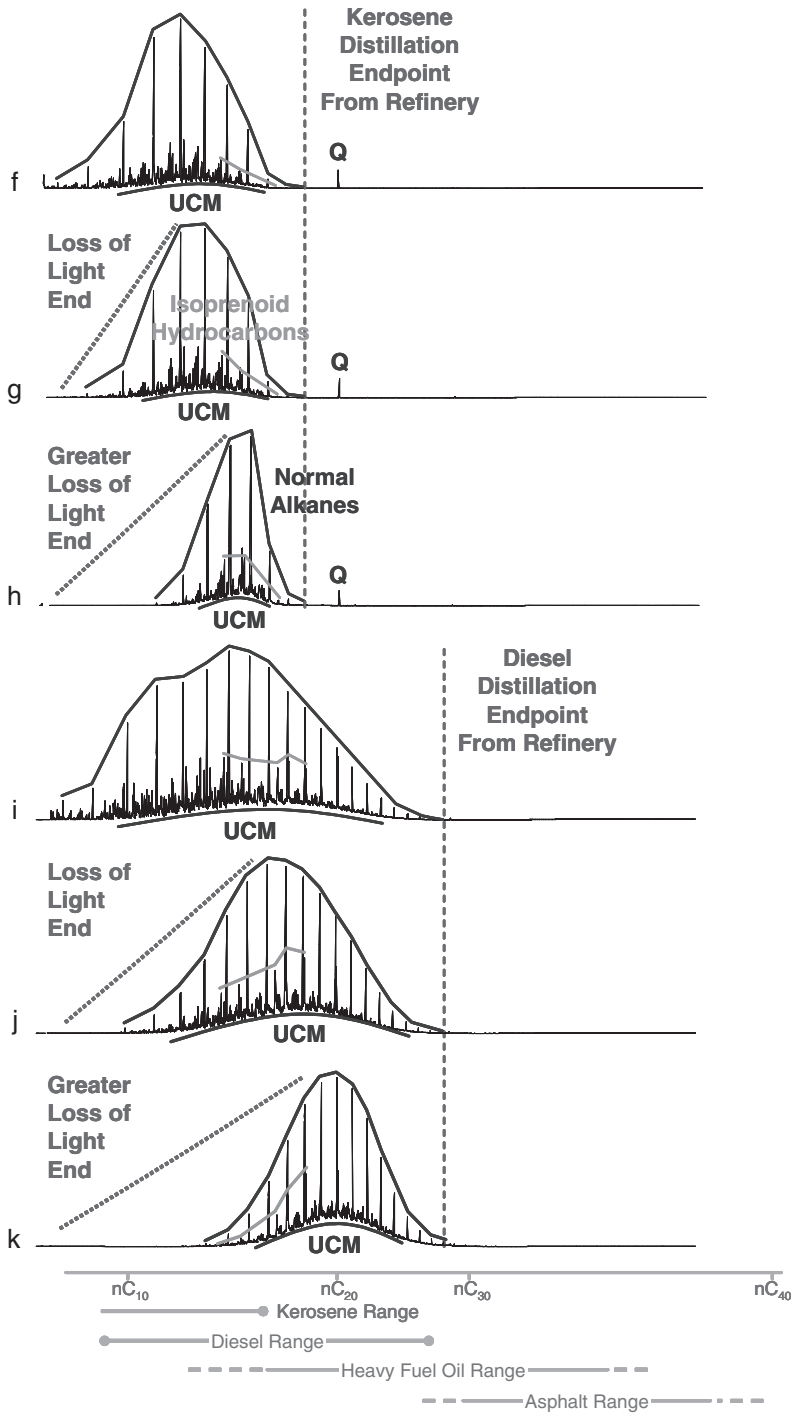


Figure 9.20—cont'd

f. *Kerosene*
This is a light to middle petroleum distillate. Notice how the heavy end profile does not change significantly during combustion.

g. *Kerosene*
50% combusted
Combustion results in the loss of lighter hydrocarbons as they volatilize during a fire.

h. *Kerosene*
Combustion Residue
When consumed and extinguished, traces of kerosene remain. Notice the absence of significant pyrogenic PAHs.

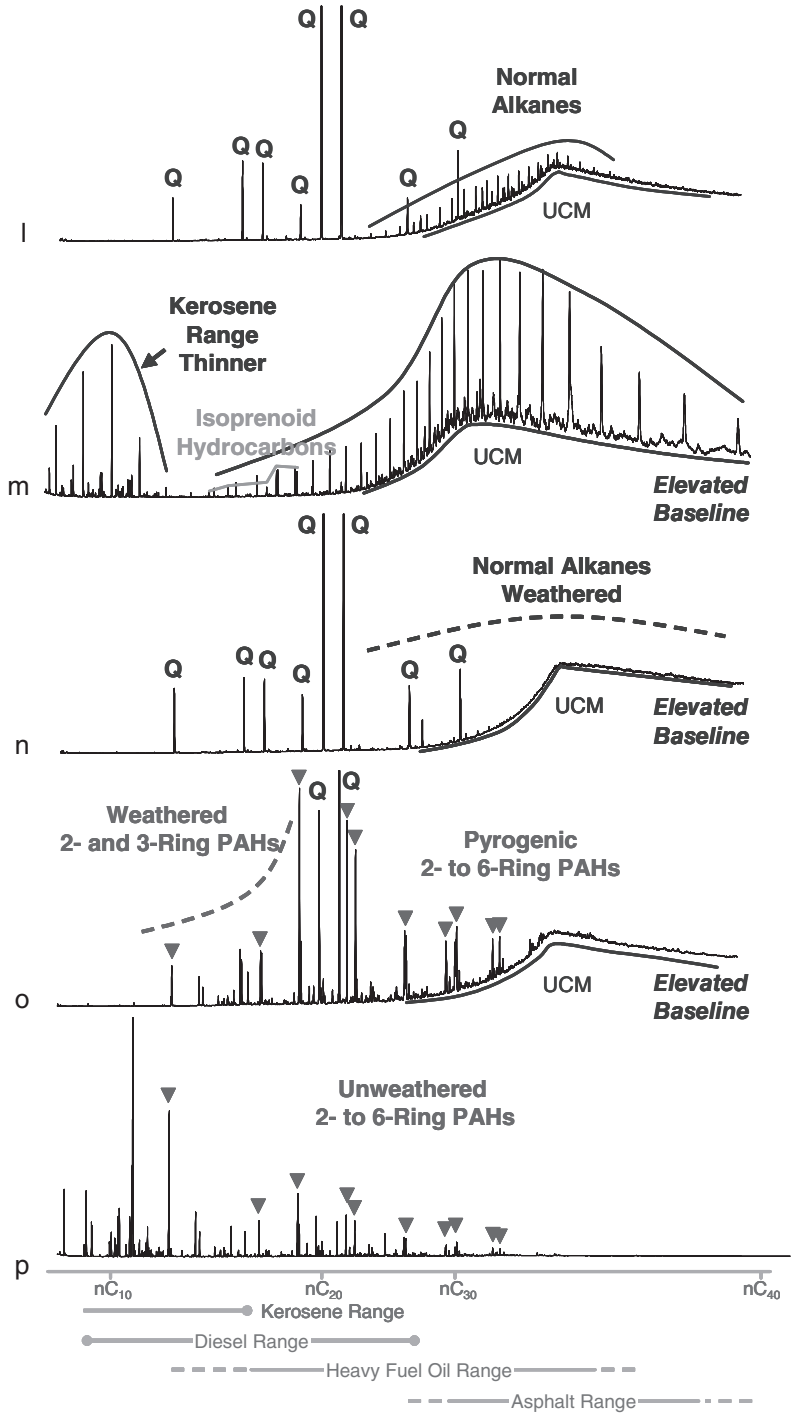
i. *Diesel*
This middle petroleum distillate contains a wider boiling point range of hydrocarbons. Again the heavy end profile does not change significantly during combustion.

j. *Diesel*
50% Combusted
Combustion results in the loss of lighter hydrocarbons as they volatilize during a fire.

k. *Diesel*
Combustion residue.
When consumed and extinguished, traces of diesel remain. Notice the absence of significant pyrogenic PAHs.

Figure 9.20—cont'd

l. *Hot patch*
This paving material commonly is used to patch roadways. Instead of kerosene, this residual range petroleum is heated to make it more malleable during its application.
m. *Roadway pavement (c. 2005)*
The dominant fraction of this material is residual range oil that is solid during the summer temperatures. Kerosene is sometimes added to improve fluidity during its application.
n. *Roadway pavement (c. 1940s)*
This pavement clearly contains residual range petroleum as evidenced by the late eluting UCM and elevated baseline. The normal alkanes typically degrade in old pavement.
o. *Roadway pavement (c. 1940s)*
The active bonding agent of roadway pavement and road oils was often tar, especially before the 1950s. Some of the pavement produced in this period also contained varying levels of residual petroleum.
p. *Coal tar oil*
Unweathered pyrogenic PAHs dominate the tar fingerprint.



Q = Quality Control compounds.
▼ = EPA priority pollutant PAHs.

The FID responds similarly for many classes of hydrocarbons (e.g., aliphatic and aromatic). When the GC/FID is operated under optimal peak resolution conditions, the method is used to generate reliable concentration measurements of individual hydrocarbons (e.g., normal alkanes, selected acyclic isoprenoids, pristane, phytane, and even prominent PAHs) (see Table 9.13). It can also be used to measure the concentration of other method-defined parameters, such as total extractable material (TEM) and total petroleum hydrocarbons (TPH). The difference between TEM and TPH is that the former still contains polar (nonhydrocarbon) material, whereas the latter has been treated using alumina or silica, thereby removing or minimizing the presence of any polar compounds. For the record, the conventional reference to TPH is misleading because this measurement also will include hydrocarbons derived from nonpetroleum sources, such as coal, tars, leaf waxes, and sewage constituents. Thus, a TPH value measured in a sample can result from the presence of petroleum, nonpetroleum hydrocarbons, or some combination thereof. As such, it is not uncommon to refer to the TPH fraction as the total extractable hydrocarbons (TEH or THC) so as to not imply any specific origin to the hydrocarbons.

The GC/FID method most often used in fingerprinting investigations is an adaptation of EPA Method 8015B, *Non-Halogenated Organics Using GC/FID* (EPA, 1997). The most appropriate chromatography columns for forensics analyses are narrow bore (0.25 mm ID) fused silica coated with a 0.25-μm 100% methylsilicone cross-linked stationary phase (or equivalent). Minimum column length for forensics purposes is 30 meters, although longer columns (e.g., 60 meters) will afford even better hydrocarbon resolution (see Table 9.14). The keys to the successful application of this method are:

- Optimal operation of the GC inlet (Douglas *et al.*, 1994; McCarthy and Uhler, 1994), which will minimize mass discrimination of heavier hydrocarbons (e.g., $>C_{25}$)
- Development of optimal GC operating conditions to facilitate hydrocarbon resolution of close-eluting compounds (e.g., baseline resolution of *n*-C_{18}/phytane)
- Calibration of the *n*-C_8 to *n*-C_{40} range with a homologous series of alkanes that are used for quantifying individual alkanes and TPH (Douglas *et al.*, 1994)

Mass discrimination in GC/FID analysis is the loss of hydrocarbon response due to suboptimal instrument performance. Mass discrimination most often is noted as a dramatic loss in instrument sensitivity above about *n*-C_{20}. Specifically, it is caused by a number of potential factors including the buildup of organic residues in the GC injection port, incompatible column type, column cut, injection temperature, liner type, and position of the capillary column in

Table 9.13

Normal alkane, acyclic isoprenoids, and range hydrocarbon analytes by GC/FID and GC/MS. Quant ion applies only to the GC/MS analysis.

Normal and Isoprenoid Hydrocarbon Analyte	Peak ID	Group ID	Molecular Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
n-Octane	nC8	P	C_8H_{18}	85	1.0 μg/L	100 μg/Kg
n-Nonane	nC9	P	C_9H_{20}	85	1.0 μg/L	100 μg/Kg
n-Decane	nC10	P	$C_{10}H_{22}$	85	1.0 μg/L	100 μg/Kg
n-Undecane	nC11	P	$C_{11}H_{24}$	85	1.0 μg/L	100 μg/Kg
n-Dodecane	nC12	P	$C_{12}H_{26}$	85	1.0 μg/L	100 μg/Kg
n-Tridecane	nC13	P	$C_{13}H_{28}$	85	1.0 μg/L	100 μg/Kg
n-Tetradecane	nC14	P	$C_{14}H_{30}$	85	1.0 μg/L	100 μg/Kg
n-Pentadecane	nC15	P	$C_{15}H_{32}$	85	1.0 μg/L	100 μg/Kg
n-Hexadecane	nC16	P	$C_{16}H_{34}$	85	1.0 μg/L	100 μg/Kg
n-Heptadecane	nC17	P	$C_{17}H_{36}$	85	1.0 μg/L	100 μg/Kg
n-Octadecane	nC18	P	$C_{18}H_{38}$	85	1.0 μg/L	100 μg/Kg
n-Nonadecane	nC19	P	$C_{19}H_{40}$	85	1.0 μg/L	100 μg/Kg
n-Eicosane	nC20	P	$C_{20}H_{42}$	85	1.0 μg/L	100 μg/Kg
n-Heneicosane	nC21	P	$C_{21}H_{44}$	85	1.0 μg/L	100 μg/Kg
n-Docosane	nC22	P	$C_{22}H_{46}$	85	1.0 μg/L	100 μg/Kg
n-Tricosane	nC23	P	$C_{23}H_{48}$	85	1.0 μg/L	100 μg/Kg
n-Tetracosane	nC24	P	$C_{24}H_{50}$	85	1.0 μg/L	100 μg/Kg
n-Pentacosane	nC25	P	$C_{25}H_{52}$	85	1.0 μg/L	100 μg/Kg
n-Hexacosane	nC26	P	$C_{26}H_{54}$	85	1.0 μg/L	100 μg/Kg
n-Heptacosane	nC27	P	$C_{27}H_{56}$	85	1.0 μg/L	100 μg/Kg
n-Octacosane	nC28	P	$C_{28}H_{58}$	85	1.0 μg/L	100 μg/Kg
n-Nonacosane	nC29	P	$C_{29}H_{60}$	85	1.0 μg/L	100 μg/Kg
n-Triacontane	nC30	P	$C_{30}H_{62}$	85	1.0 μg/L	100 μg/Kg
n-Hentriacontane	nC31	P	$C_{31}H_{64}$	85	1.0 μg/L	100 μg/Kg
n-Dotriacontane	nC32	P	$C_{32}H_{66}$	85	1.0 μg/L	100 μg/Kg
n-Tritriacontane	nC33	P	$C_{33}H_{68}$	85	1.0 μg/L	100 μg/Kg
n-Tetratriacontane	nC34	P	$C_{34}H_{70}$	85	1.0 μg/L	100 μg/Kg
n-Pentatriacontane	nC35	P	$C_{35}H_{72}$	85	1.0 μg/L	100 μg/Kg
n-Hexatriacontane	nC36	P	$C_{36}H_{74}$	85	1.0 μg/L	100 μg/Kg
n-Heptatriacontane	nC37	P	$C_{37}H_{76}$	85	1.0 μg/L	100 μg/Kg
n-Octatriacontane	nC38	P	$C_{38}H_{78}$	85	1.0 μg/L	100 μg/Kg
n-Nonatriacontane[1]	nC39	P	$C_{39}H_{80}$	85	1.0 μg/L	100 μg/Kg
n-Tetracontane	nC40	P	$C_{40}H_{82}$	85	1.0 μg/L	100 μg/Kg
2,6,10 Trimethyldodecane[1]	1380	I	$C_{16}H_{34}$	85	1.0 μg/L	100 μg/Kg
2,6,10 Trimethyltridecane[1]	1470	I	$C_{17}H_{36}$	85	1.0 μg/L	100 μg/Kg
Norpristane[1]	1650	I	$C_{18}H_{38}$	85	1.0 μg/L	100 μg/Kg
Pristane	Pr	I	$C_{19}H_{40}$	85	1.0 μg/L	100 μg/Kg
Phytane	Ph	I	$C_{20}H_{42}$	85	1.0 μg/L	100 μg/Kg
TEM or TPH (C_9-C_{40})[2]	TPH	R		85	33 μg/L	3300 μg/Kg

Notes:
[1]Analyte not in calibration. Quantified using RF of following Alkane.
[2]Quantified using average RF from C_9–C_{40} alkanes.

Group ID:
P Paraffins.
I Acyclic Isoparaffins.
R Range Analyte—Total Extractable Material (TEM) and Total Petroleum Hydrocarbons (TPH). TEM can contain polar compounds while TPH extracts were purified on an alumina column.

GC Parameter	Program Settings
Column—High resolution capillary	DB-5, 0.25 mm × 60 m × 0.25 μm (or equivalent)
Initial column temperature:	40°C
Initial hold time:	1 minute
Program rate 1:	6°C/minute
Final temp.:	315°C
Final hold time:	30 minutes
Injector temperature:	280°C
Detector temperature:	325°C
Column flow rate:	~1 mL/min. (helium)
Injector:	Splitless
Sample volume	1.0 μL

Table 9.14
Example of GC/FID conditions for the analysis of semivolatile hydrocarbons. Equivalent conditions may be used if data quality guidelines are achieved.

the injector (see Figure 9.21). Mass discrimination most severely affects the heavy molecular weight range with a negative bias that progressively increases in the pentacosane (n-C_{25}) to n-C_{44} range. Examples of mass discrimination are common in environmental laboratories; however, production chemists frequently overlook it, because control of mass discrimination is not explicitly stipulated in EPA methodology. However, in forensic investigations the potential for mass discrimination is an important consideration especially when comparing the chemical fingerprints of crude oils, residual fuels, lubricating oils, or other high boiling residues of tar. Mass discrimination can also obscure inflection points in mixed distillate fuel profiles, proportional differences in light versus heavy product fractions, and changes in any diagnostic features that rely upon compounds that elute at widely different retention times.

Fortunately, it is possible to monitor and correct for mass discrimination if the laboratory institutes early detection systems and performs proactive instrument maintenance. Mass discrimination can be efficiently monitored by analyzing an n-C_8 through n-C_{40} alkane standard before and after each analytical batch, and comparing the response factors of the higher molecular weight alkanes versus representative mid-range alkanes. Under normal operating conditions the response factor of n-C_{40} relative to the response factor of n-C_{20} should be greater than 0.85. Following the guidance of Douglas *et al.* (1994), instrument maintenance should be performed in order to maintain the highest quality chromatographic conditions when the mass discrimination exceeds this threshold.

Maximizing hydrocarbon resolution is a function of many parameters associated with the operation of the GC instrument. Principally, a slow GC oven heating program greatly improves the forensic value of the high resolution hydrocarbon fingerprints. These programs typically have initial temperatures

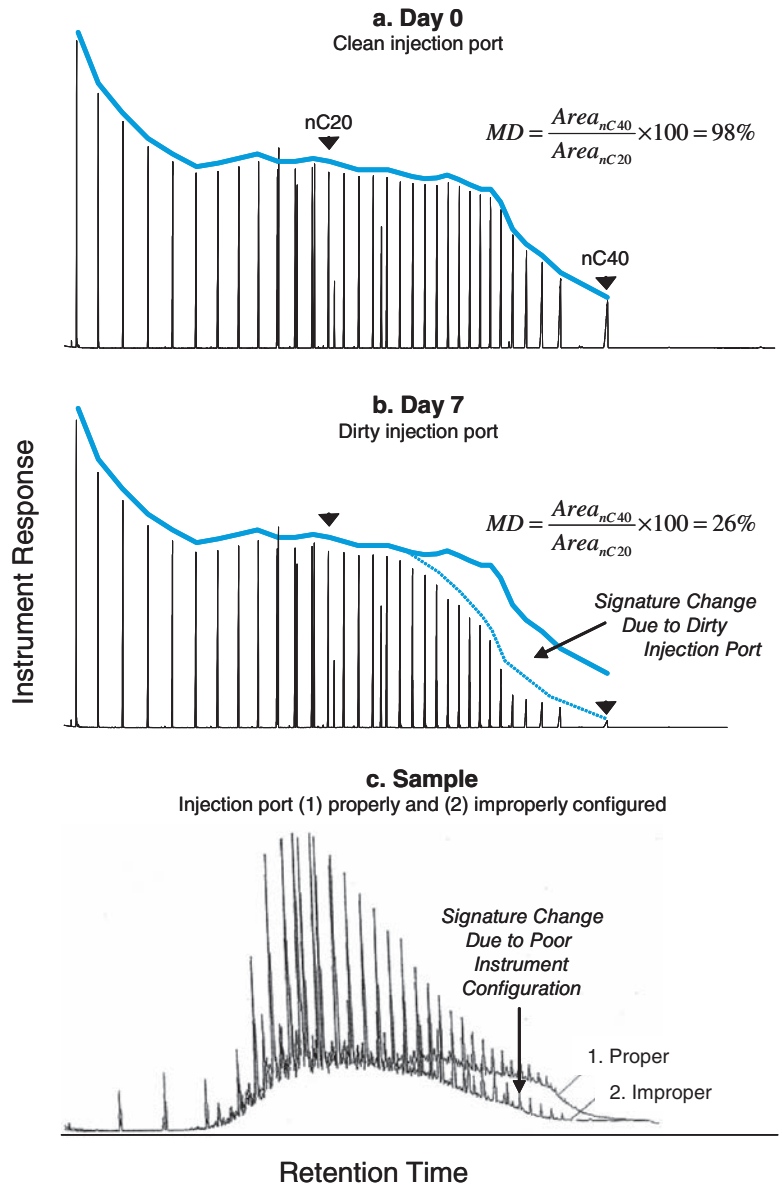


Figure 9.21
*Mass discrimination in normal alkane standard over one week (**a** and **b**). The effect of mass discrimination in a wide range petroleum product (**c**) run on an instrument with proper (1) and improper (2) injection port configurations.*

of 30˚ to 35˚C, oven ramp rates of 6˚C/min or less, and total run times of about one hour (see Table 9.14). Hydrocarbon resolution is also improved by using hydrogen as the carrier gas. These conditions stand in stark contrast to the implementation of EPA Method 8015 at most commercial laboratories, where GC run times are commonly 20 minutes or less (see Figure 9.22). While these rapid GC run times may not affect the accuracy of TEM or TPH measurements (assuming no mass discrimination), the resulting chromatographic

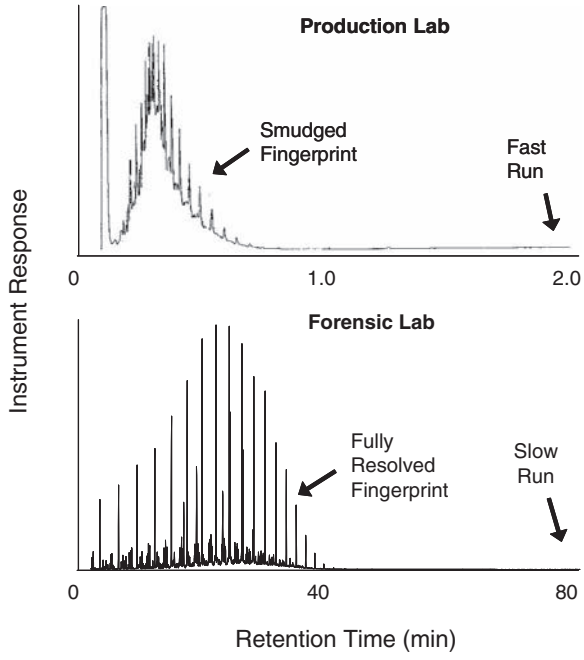


Figure 9.22
Diesel fuel oil run rapidly by a production laboratory compared to a high resolution hydrocarbon fingerprint run by a forensic laboratory.

fingerprints are compressed, poorly resolved analogs of the GC traces that are produced under optimal conditions.

Instrument sensitivity is also an important factor for detecting hydrocarbon signatures in low concentration samples. Appropriate calibration standard concentrations for quantitative GC/FID analysis range from approximately 1 μg/mL to 100 μg/mL per individual component. The calibration standard is composed of *n*-alkanes between *n*-C_8 through *n*-$C_{40,}$ pristane, phytane, surrogates, and internal standards. The individual hydrocarbon concentrations are measured using an internal standard quantitation method and the RF for the individual hydrocarbons in the calibration standard. TEM and TPH can be calculated as the resolved and unresolved hydrocarbons eluting between the *n*-C_8 and *n*-C_{40} using the IS quantitation method. These wider range analytes use total baseline-corrected areas integrated over the specific hydrocarbon range interval (Douglas *et al.*, 1994). The mean RRF of all the hydrocarbon compounds in the calibration standard serves as the RRF for the TEM and TPH measurement. Using this methodology, method detection limits (MDLs) for individual components are approximately 1 μg/L and 0.1 mg/kg (dry weight) for water and soils, respectively (Douglas *et al.*, 2004). The MDLs for TEM and TPH are approximately 100 μg/L and 5 mg/kg (dry weight) for water and soil/sediment, respectively.

Once appropriate instrumental conditions are established, high resolution hydrocarbon fingerprints for authentic environmental samples can be com-

pared to reference samples (see Figure 9.20) analyzed under the same conditions for source identification purposes. For example, crude oil contains a wide range of hydrocarbons dominated by normal alkanes eluting from before *n*-octane (*n*-C_8) to *n*-tetratriacontane (*n*-C_{40}) and beyond with lesser amounts of isoprenoid hydrocarbons (see Figure 9.20a). Petroleum refining isolates different boiling point fractions of the parent crude oil feedstock and introduces a secondary signature on the hydrocarbon fingerprint of daughter refined products (see Figure 9.20b, c, d, and e). The modified signatures of petroleum products can retain certain features of the parent crude oil, particularly those that are directly transferred from the distillation of the crude oil into the corresponding distillation range of the refined product (Morrison, 2000). Residual range petroleum products inherit at least the heavy molecular weight features of the parent crude oil feedstock and can retain them even under the influence of environmental weathering (Kennicutt 1988; McCarthy *et al.*, 1998; Douglas *et al.*, 2002) and combustion processes (see Figures 9.20f, g, h, i, j, and k—evaporation and combustion exhibit similar changes in the hydrocarbon fingerprint). The most recalcitrant source-related features recognized in the high resolution GC/FID fingerprint include the isoprenoid hydrocarbons and unresolved complex mixture (UCM) because these constituents resist biodegradation. The recalcitrant nature of heavy, residual range petroleum largely explains its preferential use of petroleum residuum for roadway pavement and road repair products (see UCM in Figures 9.20l, m, n, and o). The characteristic fingerprint of petroleum (see homologous series of normal alkanes plus isoprenoids and UCM in Figures 9.20a to 9.20n) contrasts with other commonly encountered hydrocarbon sources, such as tar-derived products (see the dominant parent PAHs with no petroleum features in Figure 9.20p) or naturally occurring hydrocarbons.

The high resolution hydrocarbon signatures of reference materials are useful in the evaluation of the compositional features of authentic field samples. An initial screening of field samples using high resolution hydrocarbon fingerprinting helps determine the range of likely products in the field sample and the need for additional tiers of analysis based on the project objectives (Stout *et al.*, 2004); for example, the presence of volatile hydrocarbons are evident in the early elution range of the high resolution hydrocarbon fingerprints and might indicate the need for PIANO analysis to characterize the source of release. This systematic evaluation employs multiple lines of evidence to identify and track source signatures of fugitive hydrocarbons in the environment. The lines of evidence include

- The molecular (boiling) range(s) of the hydrocarbon types (e.g., UCM shape)
- The relative abundances of any source specific markers (e.g., isoprenoids)
- Weathering effects

The examples provided in this section and those presented in the associated literature citations provide a basic framework for exploring the likely origins of hydrocarbon residues in many environmental forensic investigations. High resolution hydrocarbon fingerprints reveal well the overall characteristics of the predominant hydrocarbons present in field samples. However, other lines of evidence are helpful for tracking a specific source signature as it commingles with off-site hydrocarbons or "background" hydrocarbons (Bradley *et al.*, 1994; Stout *et al.*, 2000b; Stout and Uhler, 2003), and falls prey to environmental weathering (Atlas *et al.*, 1981). These lines of evidence often require quantitative measurements of diagnostic compounds that are often obscured to varying degrees in even the highest resolution GC/FID chromatograms. Fortunately, complementary Tier 2 methods using GC/MS are available for resolving these hydrocarbon signatures more definitively. Several of the more effective methods focus on selected assemblages of PAHs, saturated hydrocarbons, and geochemical biomarkers.

9.5.2.2 Polycyclic Aromatic Hydrocarbons by GC/MS/SIM

The majority of petrogenic and pyrogenic hydrocarbon sources encountered in the environment contain aromatic hydrocarbons in the *n*-C_9 to *n*-C_{40} carbon range. The homo- and heteroatomic polycyclic aromatic hydrocarbons (PAHs) constitute one of the most notable classes of compounds in this range (see Table 9.15). In crude oil and most refined petroleum products, the normal and branched alkanes comprise the majority of the hydrocarbon mixture, allowing them to be detected easily by the GC/FID fingerprinting techniques described earlier (see Figure 9.20). However, the PAHs in petroleum are usually present at much lower concentrations than the normal alkanes and require techniques such as GC/MS for their identification and measurement (Boehm *et al.*, 2000). By contrast, tar products are predominantly comprised of parent (nonalkylated) PAHs that can be readily identified in the GC/FID fingerprint (see Figure 9.20p). However, the tar derived PAHs with 2, 3, and 4 alkyl-groups are present at concentrations less than the parent PAHs recognized by GC/FID; therefore, the full PAH signature requires supplemental techniques such as GC/MS for full resolution. In summary, the high resolution hydrocarbon fingerprint method (GC/FID) describes the dominant hydrocarbons and the complementary PAH method (GC/MS/SIM) isolates source patterns within the PAH assemblage.

The diagnostic compounds in Table 9.15 includes parent and alkyl-substituted PAHs (Youngblood and Blumer, 1975) with demonstrated diagnostic utility in forensics investigations (Boehm and Farrington, 1984; Colombo *et al.*, 1989; Sauer and Uhler, 1994; Stout *et al.*, 2002). The GC/MS procedures for the optimal detection, separation, and measurement of PAHs are based on EPA Method 8270C, *Semivolatile Organic Compounds by Gas*

Table 9.15

PAH target analytes measured by GC/MS/SIM.

PAH Analyte	Peak ID	Ring No.	Molecular Formula	Quant[1] Ion (m/z)	Approximate Detection Limit		PAH Group				
					Water	Solid	PP	Total	Par	Alk	Dia
Naphthalene	N0	2	$C_{10}H_8$	128	10 ng/L	2 μg/Kg	X	X	X		
C1-Naphthalenes	N1	2	$C_{11}H_{10}$	142	10 ng/L	2 μg/Kg		X		X	
C2-Naphthalenes	N2	2	$C_{12}H_{12}$	156	10 ng/L	2 μg/Kg		X		X	
C3-Naphthalenes	N3	2	$C_{13}H_{14}$	170	10 ng/L	2 μg/Kg		X		X	
C4-Naphthalenes	N4	2	$C_{14}H_{16}$	184	10 ng/L	2 μg/Kg		X		X	
Biphenyl	B	2	$C_{12}H_{10}$	154	10 ng/L	2 μg/Kg		X	X		
Dibenzofuran	DF	3	$C_{12}H_8O$	168	10 ng/L	2 μg/Kg		X	X		
Acenaphthylene	AY	3	$C_{12}H_8$	152	10 ng/L	2 μg/Kg	X	X	X		
Acenaphthene	AE	3	$C_{12}H_{10}$	153	10 ng/L	2 μg/Kg	X	X	X		
Fluorene	F0	3	$C_{13}H_{10}$	166	10 ng/L	2 μg/Kg	X	X	X		
C1-Fluorenes	F1	3	$C_{14}H_{12}$	180	10 ng/L	2 μg/Kg		X		X	
C2-Fluorenes	F2	3	$C_{15}H_{14}$	194	10 ng/L	2 μg/Kg		X		X	
C3-Fluorenes	F3	3	$C_{16}H_{16}$	208	10 ng/L	2 μg/Kg		X		X	
Dibenzothiophene	DBT0	3	$C_{12}H_8S$	184	10 ng/L	2 μg/Kg		X	X		
C1-Dibenzothiophenes	DBT1	3	$C_{13}H_{10}S$	198	10 ng/L	2 μg/Kg		X		X	
C2-Dibenzothiophenes	DBT2	3	$C_{14}H_{12}S$	212	10 ng/L	2 μg/Kg		X		X	
C3-Dibenzothiophenes	DBT3	3	$C_{15}H_{14}S$	226	10 ng/L	2 μg/Kg		X		X	
C4-Dibenzothiophenes	DBT4	3	$C_{16}H_{16}S$	240	10 ng/L	2 μg/Kg		X		X	
Phenanthrene	P0	3	$C_{14}H_{10}$	178	10 ng/L	2 μg/Kg	X	X	X		
Anthracene	A0	3	$C_{14}H_{10}$	178	10 ng/L	2 μg/Kg	X	X	X		
C1-Phenanthrenes/Anthracenes	PA1	3	$C_{15}H_{12}$	192	10 ng/L	2 μg/Kg		X		X	
C2-Phenanthrenes/Anthracenes	PA2	3	$C_{16}H_{14}$	206	10 ng/L	2 μg/Kg		X		X	
C3-Phenanthrenes/Anthracenes	PA3	3	$C_{17}H_{10}$	220	10 ng/L	2 μg/Kg		X		X	
C4-Phenanthrenes/Anthracenes	PA4	3	$C_{18}H_{18}$	234	10 ng/L	2 μg/Kg		X		X	
Retene	RET	3	$C_{18}H_{18}$	234	10 ng/L	2 μg/Kg					X
Benzo(b)fluorene	BF	4	$C_{17}H_{12}$	216	10 ng/L	2 μg/Kg		X	X		

Table 9.15

Continued

PAH Analyte	Peak ID	Ring No.	Molecular Formula	Quant[1] Ion (m/z)	Approximate Detection Limit		PAH Group				
					Water	Solid	PP	Total	Par	Alk	Dia
Fluoranthene	FL0	4	$C_{16}H_{10}$	202	10 ng/L	2 μg/Kg	X	X	X		
Pyrene	PY0	4	$C_{16}H_{10}$	202	10 ng/L	2 μg/Kg	X	X	X		
C1-Fluoranthenes/Pyrenes	FP1	4	$C_{17}H_{12}$	216	10 ng/L	2 μg/Kg		X		X	
C2-Fluoranthenes/Pyrenes	FP2	4	$C_{18}H_{14}$	230	10 ng/L	2 μg/Kg		X		X	
C3-Fluoranthenes/Pyrenes	FP3	4	$C_{19}H_{16}$	244	10 ng/L	2 μg/Kg		X		X	
C4-Fluoranthenes/Pyrenes	FP4	4	$C_{20}H_{18}$	258	10 ng/L	2 μg/Kg		X		X	
Naphthobenzothiophenes	NBT0	4	$C_{16}H_{10}S$	234	10 ng/L	2 μg/Kg		X	X		
C1-Naphthobenzothiophenes	NBT1	4	$C_{17}H_{12}S$	248	10 ng/L	2 μg/Kg		X		X	
C2-Naphthobenzothiophenes	NBT2	4	$C_{18}H_{14}S$	262	10 ng/L	2 μg/Kg		X		X	
C3-Naphthobenzothiophenes	NBT3	4	$C_{19}H_{16}S$	276	10 ng/L	2 μg/Kg		X		X	
C4-Naphthobenzothiophenes	NBT4	4	$C_{20}H_{18}S$	290	10 ng/L	2 μg/Kg		X		X	
Benz[a]anthracene	BA0	4	$C_{18}H_{12}$	228	10 ng/L	2 μg/Kg	X	X	X		
Chrysene/Triphenylene	C0	4	$C_{18}H_{12}$	228	10 ng/L	2 μg/Kg	X	X	X		
C1-Chrysenes	BC1	4	$C_{19}H_{14}$	242	10 ng/L	2 μg/Kg		X		X	
C2-Chrysenes	BC2	4	$C_{20}H_{16}$	256	10 ng/L	2 μg/Kg		X		X	
C3-Chrysenes	BC3	4	$C_{21}H_{18}$	270	10 ng/L	2 μg/Kg		X		X	
C4-Chrysenes	BC4	4	$C_{22}H_{20}$	284	10 ng/L	2 μg/Kg		X		X	
Benzo[b]fluoranthene	BBF	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg	X	X	X		
Benzo[j,k]fluoranthene	BKF	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg	X	X	X		
Benzo[a]fluoranthene	BAF	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg		X	X		
Benzo[e]pyrene	BEP	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg		X	X		
Benzo[a]pyrene	BAP	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg	X	X	X		
Perylene	PER	5	$C_{20}H_{12}$	252	10 ng/L	2 μg/Kg		X			X
Dibenz[a,h]anthracene	DA	5	$C_{22}H_{14}$	278	10 ng/L	2 μg/Kg	X	X	X		
Indeno[1,2,3-cd]pyrene	IND	6	$C_{22}H_{12}$	276	10 ng/L	2 μg/Kg	X	X	X		
Benzo[g,h,i]perylene	GHI	6	$C_{22}H_{12}$	276	10 ng/L	2 μg/Kg	X	X	X		

Note:
[1]The alkyl homolog groups are assigned the RF of the parent compound.

PAH Groups:
PP EPA Priority Pollutant PAHs (n = 16).
Total Total PAHs do not include retene because it is included in the PA4 analyte.
Par Parent PAHs have no alkylated functional groups.
Alk Alkylated PAHs have one or more methyl functional group.
Dia Diagentic PAHs can be formed naturally in the environment.

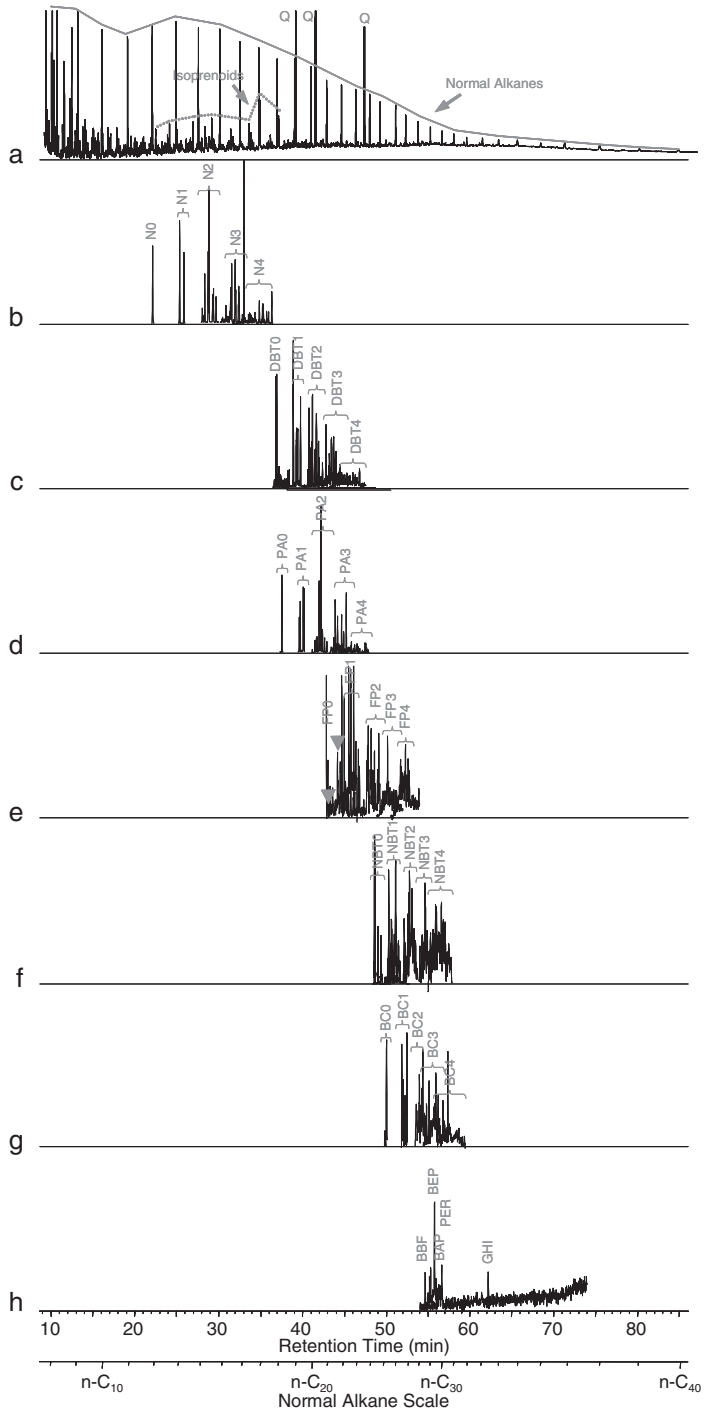


Figure 9.23
Crude oil signatures revealed by high resolution hydrocarbon fingerprints (GC/FID) and assemblages of PAH analytes (GC/MS). ***a.*** *Crude oil (GC/FID) Dominated by normal alkanes.* ***b.*** *Naphthalenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***c.*** *Dibenzothiophenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***d.*** *Phenanthrenes and anthracenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***e.*** *Fluoranthenes and pyrenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***f.*** *Naphthabenzothiophenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***g.*** *Benz(a)anthracenes and chrysenes (GC/MS/SIM) Petrogenic profile evident by low parent relative to alkylated PAHs.* ***h.*** *5- and 6-ring PAHs Trace heavy PAHs dominated by BEP and PER. Compound abbreviations as per Table 9.15.*

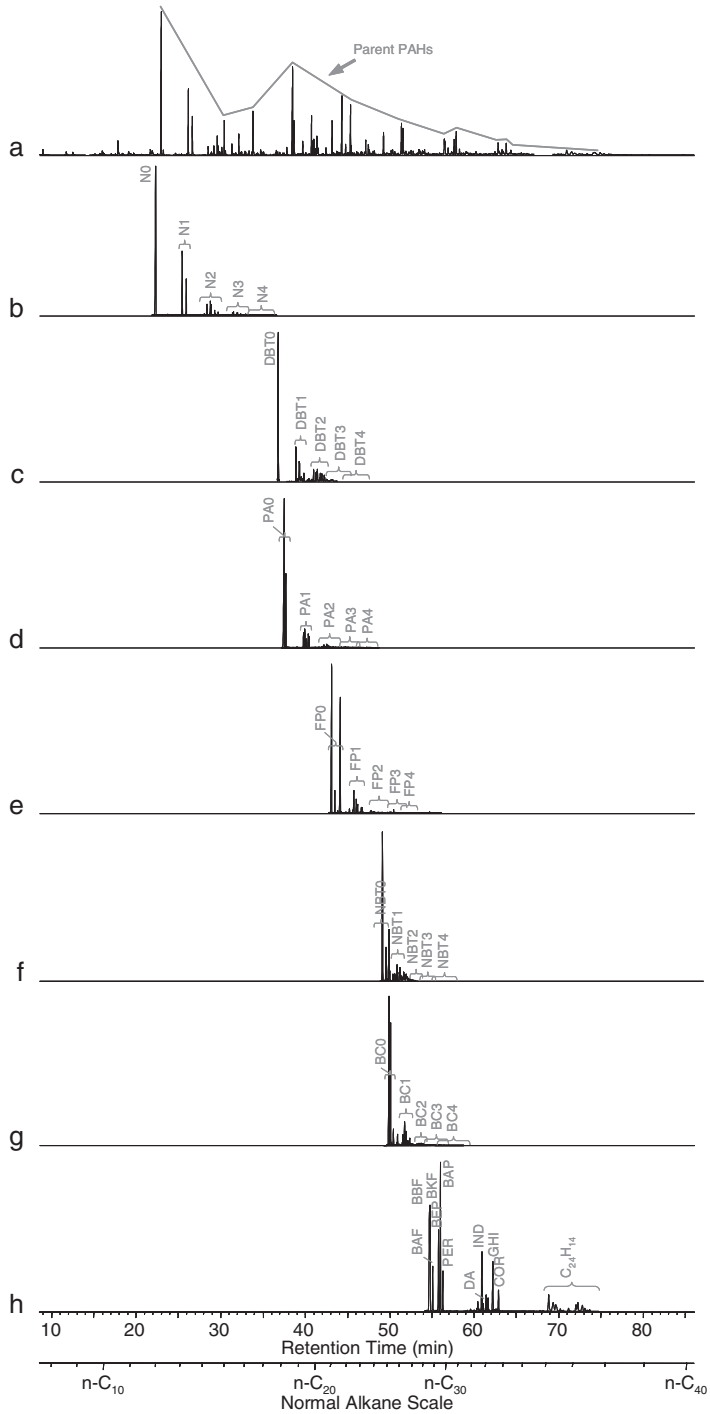


Figure 9.24
Coal tar signatures revealed by high resolution hydrocarbon fingerprints (GC/FID) and assemblages of PAH analytes (GC/MS).
a. *Coal tar (GC/FID) Dominated by parent PAHs.*
b. *Naphthalenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
c. *Dibenzothiophenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
d. *Phenanthrenes and anthracenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
e. *Fluoranthenes and pyrenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
f. *Naphthabenzothiophenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
g. *Benz(a)anthracenes and chrysenes (GC/MS/SIM) Pyrogenic profile evident by declining abundance with increasing alkylation.*
h. *5- and 6-ring PAHs Full range of heavy PAHs. All thermogenic except PER. Compound abbreviations as per Table 9.15.*

Chromatography/Mass Spectrometry (GC/MS) (EPA, 1997). As discussed previously, the GC and MS operating conditions need to be optimized for separation and sensitivity. A capillary column should be at least 30 m long with a nonpolar stationary phase (DB-5 or equivalent) with a 0.25 mm inner diameter (i.d.) and a 0.25 μm film thickness of stationary phase (5% phenyl-95% methyl- silicone, or equivalent). Optimal operation of the GC should be established to minimize mass discrimination of heavier hydrocarbons (e.g., $>C_{30}$) during sample injection (McCarthy and Uhler, 1994). A slow GC oven temperature program (see high resolution hydrocarbon fingerprinting method) should also be used to ensure adequate resolution of close-eluting compounds and alkylated PAH isomer patterns. The appropriate GC oven program typically starts around 35˚C and ramps at 6˚C/min during the elution of most of the compounds. Total GC run times range between one and two hours. As a practical matter, the specific GC operating parameters for the high resolution hydrocarbon fingerprints via GC/FID (described earlier) also work well for the analysis of PAHs (see Table 9.14).

As part of the method and development of an alkylated PAH method, the chromatographic retention times for each target compound must be established under the appropriate GC/MS operating conditions. The diagnostic and confirmatory ions for each target compound must be verified from GC/MS analyses of authentic standards. Since only a handful of authentic standards exist for the literally hundreds of C_1 to C_4 alkylated isomers of PAHs, the characteristic ions, retention time windows, and homolog group patterns must be constructed based on carefully documented PAH and alkyl PAH chromatographic retention indices (Lee *et al.*, 1979) and then redocumented by the analysis of a well-characterized reference petroleum, such as Alaska North Slope Crude Oil (see Figure 9.24) (Wang *et al.*, 1994). Alkyl PAH homologs typically are quantified using the straight baseline integration method and response factors assigned from the parent PAH compound (Sauer and Boehm, 1991; Douglas *et al.*, 1994; Federal Register, 1994; Sauer and Boehm, 1995). Although the actual RFs of the parent PAH and all of the alkylated equivalent may differ slightly, it is impractical to change the parent RF calibration procedure due to the nonexistence of standards for all the alkylated PAH isomers. It is more practical to use the parent PAH RFs for the calculation of alkylated PAH concentrations to provide a uniform and precise quantitation system that is particularly well suited for qualitative and quantitative comparisons among authentic and reference samples that are generated using the same calibration method.

Alkylated PAH data frequently are presented as histograms that graphically represent the absolute or relative target analyte concentrations for the purpose of qualitative fingerprinting. Using the procedures described in this section,

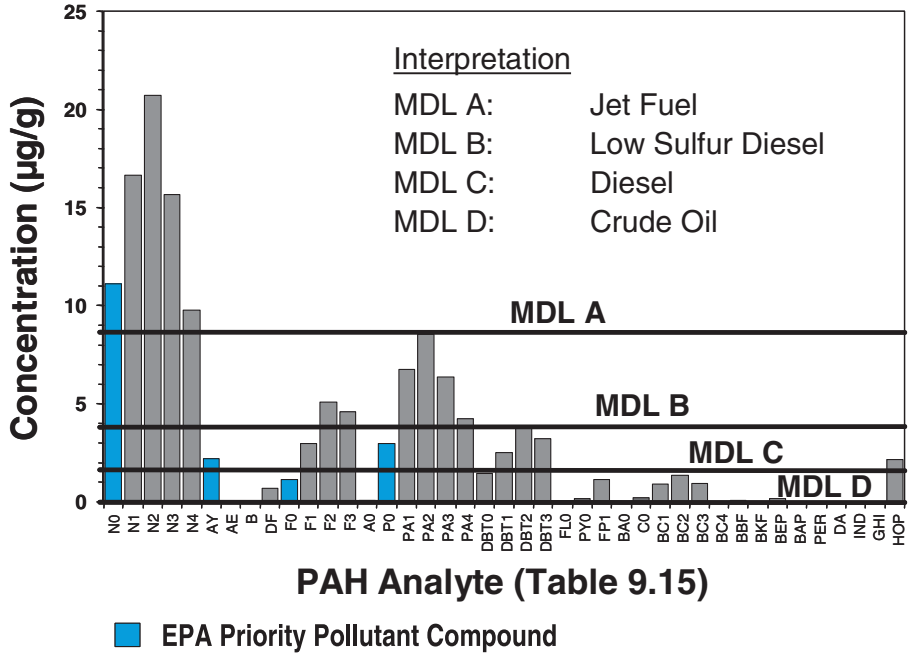


Figure 9.25
The method detection limit (MDL) affects the interpretation of the source signature (Douglas et al. *2004).*

a PAH histogram bar represents the concentration of an individual PAH compound or the collective concentration of the isomers within an alkylated PAH homolog group. Successful application of PAH compound distributions to forensic investigations is a function of the specificity and sensitivity of the method used. The forensic chemist relies heavily on the relative distributions of compounds in a sample to identify product type, source, and degree of weathering. Figure 9.25 highlights the five PAH compounds detected using the standard EPA Method 8270C target compounds compared to the PAH target analyte list for the GC/MS method discussed here. Clearly, inclusion of the alkylated PAHs and dibenzothiophenes markedly increases the potential to distinguish and describe product type, source, and degree of weathering (Douglas *et al.*, 2002). Sensitive PAH MDLs also are critical for hydrocarbon fingerprinting studies when the hydrocarbon concentrations are low because the identification of hydrocarbon type and source depends on the relative distributions of the analytes in the sample (i.e., the PAH fingerprint, Douglas *et al.*, 2004; Page *et al.*, 1995).

Nondetects due to elevated MDLs provide little forensic information. Since the PAH concentrations in a sample can vary by as much as several orders of magnitude, high MDLs capture only the signature provided by the most abundant analytes. This can result in an erroneous source identification conclusion (see Figure 9.25). Using a hypothetical example, a laboratory that generates an alkylated PAH fingerprint of crude oil using an insensitive method might

detect only petrogenic naphthalenes and misidentify the material as kerosene (see MDL A, Figure 9.25). A laboratory using a method with half the MDL of the first lab would detect petrogenic naphthalenes, fluorenes, and phenanthrenes and conclude that the material is low sulfur diesel (see MDL B, Figure 9.25). A laboratory using a method with half the MDL of the second lab would also detect alkylated dibenzothiophenes and conclude that the material is regular diesel (see MDL C, Figure 9.25). It is only the laboratory that uses a method with very low MDLs that can detect the full range of petrogenic PAHs that will conclude correctly that the material is crude oil (see MDL D, Figure 9.25). Qualitative and quantitative PAH fingerprinting benefits greatly from the use of robust measurement techniques with a known level of sensitivity. The MDL for the PAH analytes in this method are approximately 1 ng/L and 0.05 μg/kg (dry weight) for water and soil/sediment, respectively (Douglas *et al.*, 2004). This level of sensitivity is sufficient for identifying source signatures and monitoring weathering near background concentrations in most media (Boehm *et al.*, 1998; Boehm *et al.*, 2000).

The concentration and distribution of PAHs provided greater detail and specificity about the type of petroleum, tar, soot, and plant material in the field samples. The terms, petrogenic, pyrogenic, and diagenetic, are important for hydrocarbon fingerprinting because they generally indicate the presence of PAHs from different sources. For example, petroleum possesses a petrogenic PAH pattern consisting of low parent abundance relative to the alkylated PAHs (e.g., N0 < N1 < N2 > N3 > N4 in Figure 9.26a). By contrast, pyrogenic PAHs form during the partial combustion or pyrolysis of organic matter. A pyrogenic PAH pattern exhibits high parent abundance relative to the alkylated PAH (e.g., N0 < N1 < N2 < N3 < N4 in Figure 9.26p, before weathering). This downward sloping profile is caused by the pyrolytic cleavage of alkyl side chains under hot conditions—hot environments with many catalytic surfaces yield lower levels of alkylated PAHs than cooler temperatures with fewer reactive surfaces. Finally, diagenetic PAHs, like retene and perylene, form naturally due to oxidation in sediments containing specific types of decaying organic debris (e.g., Wakeham *et al.*, 1980; Venkatesen *et al.*, 1988).

A selection of reference samples discussed previously in the high resolution hydrocarbon fingerprinting section (see Figure 9.20) help graphically portray the corresponding diversity in PAH signatures (see Figure 9.26). Crude oil, for example, contains a wide range of petrogenic PAHs (see Figure 9.26a). Distillate products retain the petrogenic PAH character of the crude oil within the boiling point range of the distillate fraction. The PAH profile of diesel range distillate is predominantly petrogenic 2-ring PAHs with markedly lower relative abundances of 3- to 4-ring PAHs (see Figure 9.26b). Heavier petroleum products, like #4 Fuel Oil (see Figure 9.26c) and #6 Fuel Oil (see Figure

9.26d) contain progressively heavier fractions of crude oil with higher proportions of petrogenic 4-ring PAHs (see Figures 9.26c and 9.26d, respectively). Heavy refined petroleum distillates, like lubricating oil, are specifically distilled to operate in motors and machines with specific operating temperatures and mechanical properties. New lubricating oil contains predominantly low levels of petrogenic 3-ring PAHs (total EPA PAHs is 4 mg/kg) that becomes enriched in a complex mixture of weathered pyrogenic 2- to 6-ring PAHs (total EPA PAHs is 710 mg/kg) due to contact with combustion byproducts in the enclosed confines of a gasoline engine (see Figure 9.26e, Pruell *et al.*, 1988; Vazquez-Duhalt, 1989). By contrast, a laboratory study demonstrated that middle distillates did not accumulate pyrogenic PAHs during open flame experiments; rather the compositional changes during combustion closely matched a parallel experiment in which the fuels were simply evaporated (see Figures 9.26f to 9.26k). This finding is consistent with the fact that flame is generated in the vapor phase above the liquid fuel, and combustion byproducts (e.g., pyrogenic PAHs) rise rapidly above the flame in an unconfined combustion environment. Importantly, the signature integrity of the heavier petrogenic PAHs characteristic of each distillate is retained throughout the combustion process (0 to 99% mass reduction) and serves as a reliable signature feature of the fuel used at fire fighter training areas and arson investigations.

Additional reference samples of pavement materials and coal tar demonstrated the types of materials likely to influence the storm water runoff in areas with high densities of paved surfaces. For example, hot patch is a material commonly used to repair modern roads. It contains heavy residual petroleum consistent with the residue of distilled crude oil after the removal of light, middle, and heavy distillates. After extensive heating at the refinery and during its hot application, this material is predominantly petrogenic 2- to 4-ring PAHs mixed with a significant quantity of pyrogenic 3- to 6 ring PAHs (see Figure 9.26l). Like hot patch, modern roadway pavement material contains a mixture of heavy petrogenic and pyrogenic PAHs thinned with kerosene (notice elevated alkyl-naphthalenes) used to facilitate its application (see Figure 9.26m). By contrast, older pavement contained different types and blends of residual petroleum and tar (see Figures 9.26n and 9.26o). A crude coal tar reference sample demonstrates the enriched levels of parent 2- to 6-ring PAHs (see Figure 9.26p) that serves as the raw material from which many refined tar products are formed (e.g., creosote, anthracene oil, and tar pitch).

Comparison of field samples to suspected source and reference materials is useful, but this comparison must be considered in the context of environmental weathering. A recent study of PAH ratio stability helped identify the

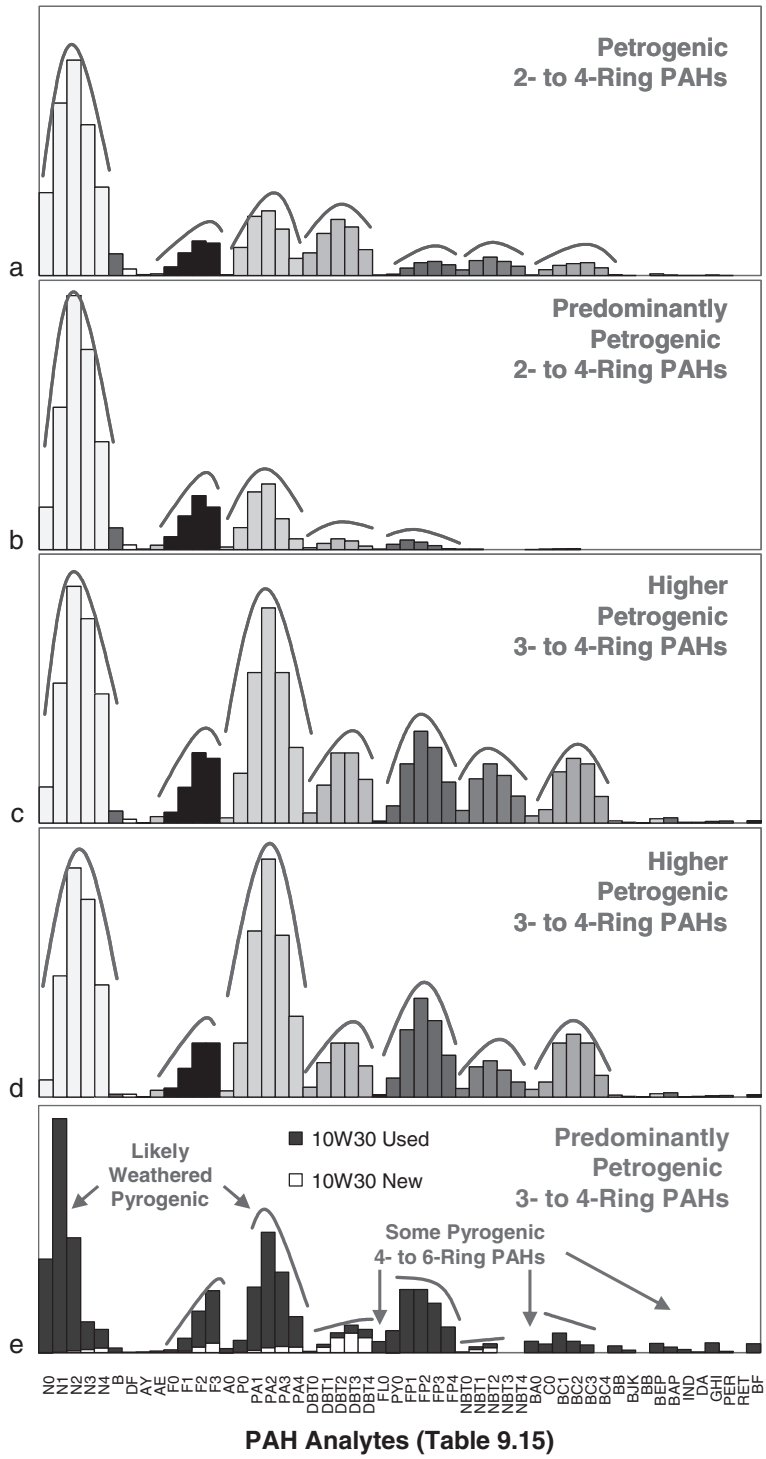


Figure 9.26

Selected PAH histograms of hydrocarbon sources.
a. *Crude oil*
Crude oil and coal used as fuels and fuel feedstocks contain a wide range of petrogenic 2- to 4-ring PAHs.
b. *#2 fuel oil*
#2 fuel oil contains petrogenic 2- and 3-ring PAHs. Low levels of petrogenic 4-ring PAHs may also be present.
c. *#4 fuel oil*
#4 fuel oil contains petrogenic 2- to 4-ring PAHs. It contains less 2-ring PAHs and more 4-ring PAHs than #2 fuel oil due to the presence of heavier petroleum fractions.
d. *#6 fuel oil*
Heavy fuel oils fall into the residual range of crude oil. They contain less 2-ring and more 4-ring PAHs than #2 and #4 Fuel Oils, although the latter distinction can be subtle.
e. *New and Used lubricating oils, are manufactured from heavy distillates that are subsequently purified and fortified with performance enhancing agents. The low levels of PAHs in new lube oil can increase by more than 10 times in used oil as a result of pyrolysis inside the engine.*
A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

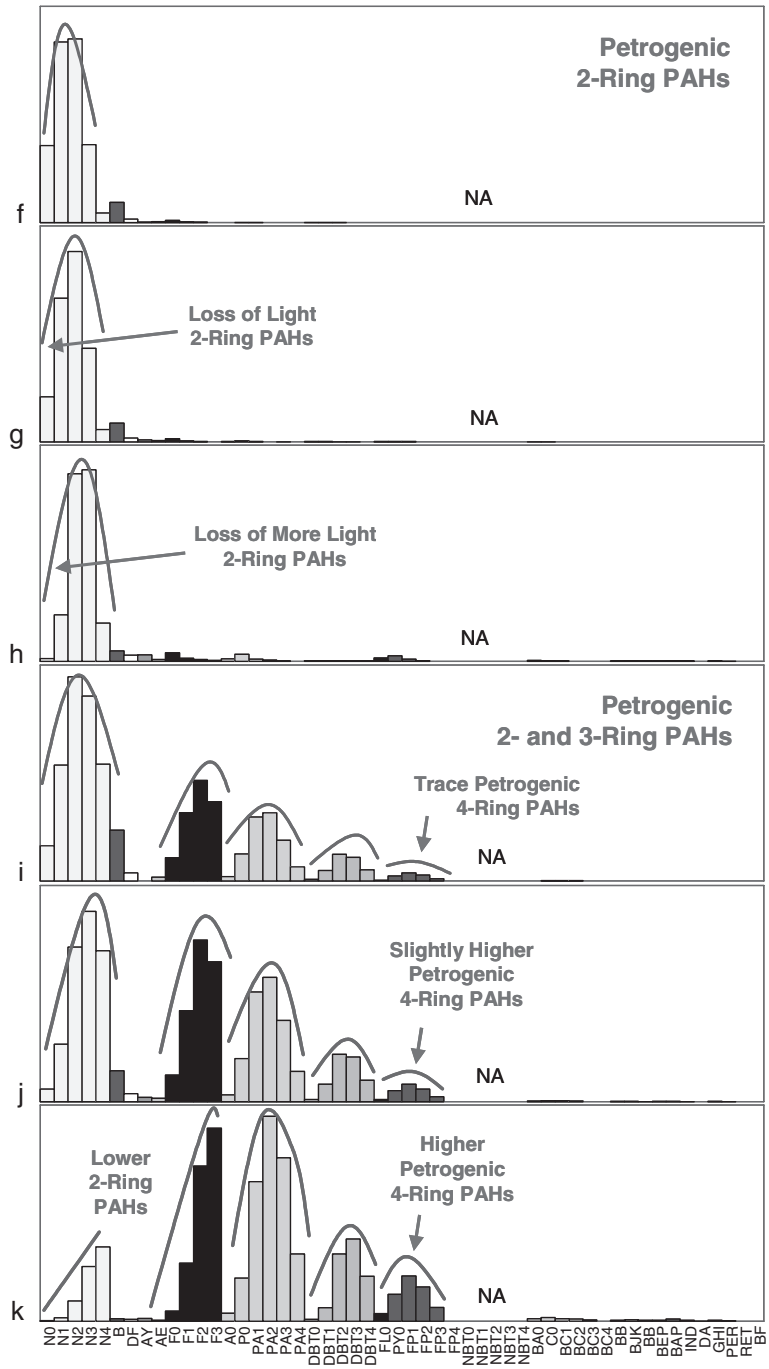


Figure 9.26

f. *Kerosene*
This light middle petroleum distillate is overwhelmingly composed of petrogenic 2-ring PAHs.
g. *Kerosene 50% combusted*
Combustion causes the preferential loss of light 2-ring PAHs.
h. *Kerosene combustion residue*
Combustion causes the continued loss of light 2-ring PAHs.
i. *Diesel*
This middle petroleum distillate contains a wider boiling point range of PAHs. The heavy end profile does not change significantly during combustion.
j. *Diesel 50% combusted*
Combustion results in the loss of lighter PAHs.
k. *Diesel combustion residue When consumed and extinguished, traces of diesel remain. Notice the absence of significant pyrogenic PAHs.*
A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

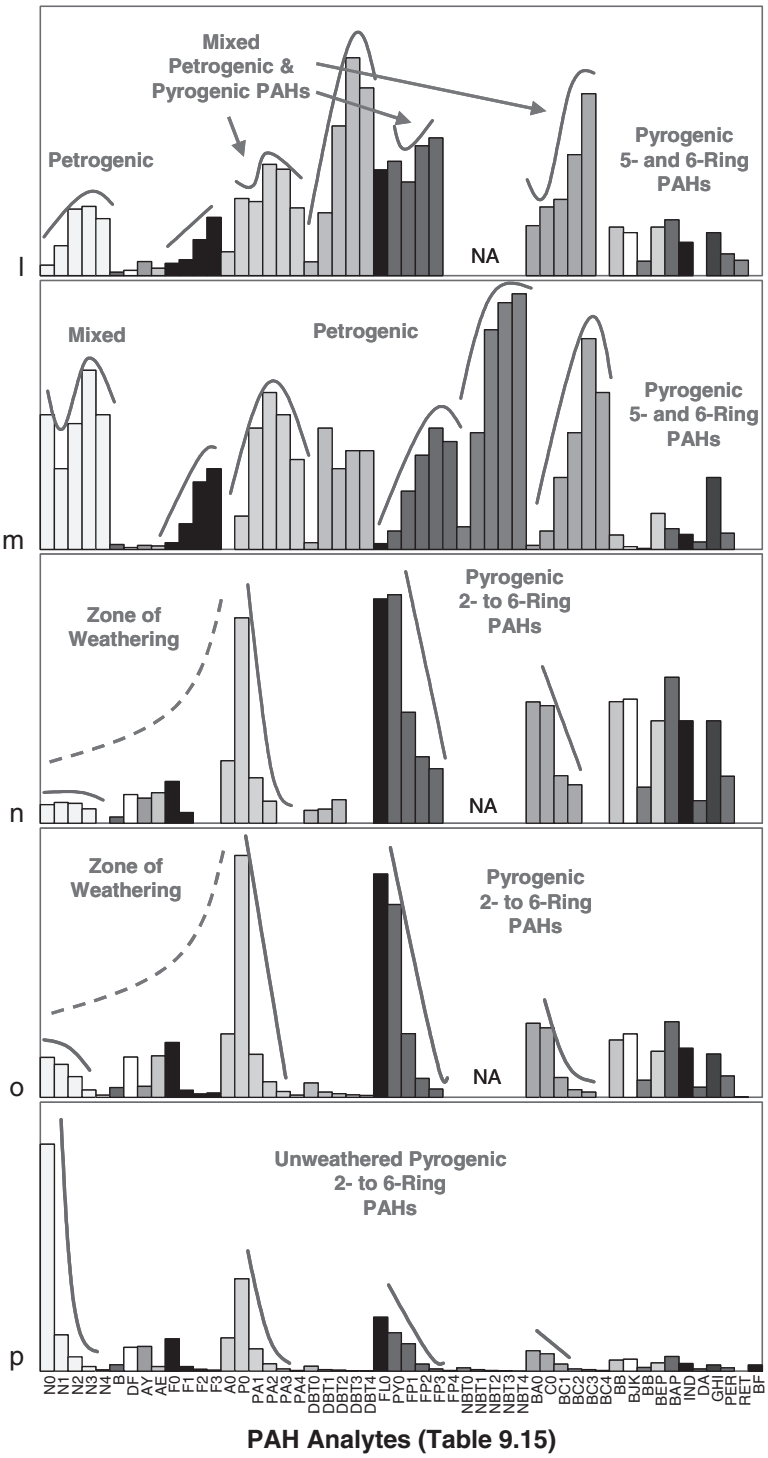


Figure 9.26—cont'd

l. *Hot patch*
This paving material is composed of heavy PAHs in the residual petroleum range. It is a complex mixture of petrogenic 2- to 4-ring PAHs mixed with pyrogenic 3- to 6-ring PAHs.

m. *Roadway pavement (c. 2005)*
Like Hot patch, this modern pavement contains mixtures of pyrogenic and petrogenic PAHs. In particular, it has higher levels of 2-ring PAHs caused by the kerosene solvent.

n. *Roadway pavement (c. 1940s)*
Mid-century pavement was frequently blended with petroleum and tar. In this sample the tar derived pyrogenic 3- to 6-ring PAHs are dominant.

o. *Roadway pavement (c. 1940s)*
Pavement from the early to mid 1900s was predominantly made with tar binders. Accordingly, the pyrogenic 2- to 6-ring PAHs from tar were well preserved in the roadway bedding.

p. *Coal tar oil*
The downward sloping profile of each homolog group is a feature shared by all pyrogenic tars. Subtle changes in several isomer ratios (e.g., FL0: PY0) are characteristic of the different processes by which they are generated. (Emsbo-Mattingly et al., *2001b)*

A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

Table 9.16

PAH Ratio Stability. The PAH ratios in the upper part of the table are stable. The ratios towards the bottom of the table are expected to change in a nonlinear or unpredictable fashion in environments susceptible to evaporation and biodegradation. See Table 9.15 for full PAH names.

PAH Ratio	Evaporation			Aerobic Biodegradation			Most Stable Ratios
	T = 0 h	T = 168 h	%Change	T = 0 wks	T = 7 wks	%Change	
Weathering Index	1.68	0.22		2.05	0.06		
FL0/PY0	1.33	1.27	5%	0.74	0.65	12%	✓
(BBF+BKF)/C0	1.22	1.20	2%	1.34	1.21	10%	✓
PER/C0	0.22	0.21	2%	0.30	0.27	8%	✓
BBF/BKF	1.05	1.02	2%	1.21	1.12	7%	✓
FL0/(FL0+FY0)	0.57	0.56	2%	0.42	0.39	7%	✓
BA0/C0	1.05	1.01	4%	0.90	0.85	5%	✓
BBF/BAP	0.84	0.85	–2%	0.58	0.56	4%	✓
BA0/(BA0+C0)	0.51	0.50	2%	0.47	0.46	3%	✓
BF/FP1	0.20	0.21	–5%	0.09	0.09	1%	✓
(BBF+BKF)/BAP	1.64	1.69	–3%	1.07	1.06	1%	✓
C0/(C0+FP1)	0.54	1.48	12%	0.49	0.49	0.20%	✓
IP/(IP+GHI)	0.52	1.54	–4%	0.47	0.47	0%	✓
IP/GHI	1.08	1.18	–9%	0.90	0.90	0%	✓
PER/BAP	0.29	1.30	–2%	0.24	0.24	–2%	✓
DBT3/PA3	0.36	0.35	2%	0.42	0.43	–3%	✓
BKF/BAP	0.80	0.83	–4%	0.48	0.50	–4%	✓
BEP/BAP	0.60	0.61	–2%	0.62	0.64	–4%	✓
DBT2/PA2	0.22	0.21	6%	0.27	0.29	–7%	✓
(BA0+C0)/(BC2+BC3)	7.68	8.10	–5%	20.15	22.96	–14%	✓
N0/(N2+N3)	4.75	0.02	100%	14.00	0.01	100%	
DBT0/(DBT2+DBT3)	1.74	0.30	83%	3.10	0.07	98%	
(P0+A0)/(PA2+PA3)	13.77	2.53	82%	13.57	0.37	97%	
(ΣDBTO-DBT4)/ (BA0+C0)	0.23	0.11	52%	0.51	0.14	72%	
BC1/FP1	0.52	0.39	24%	0.24	0.22	11%	
(FL0+PY0)/(FP2+FP3)	16.20	7.81	52%	30.26	27.50	9%	
NBT0/C0	0.13	0.18	–44%	0.21	0.23	–8%	
NBT2/BAP	0.04	0.06	–60%	0.02	0.02	–13%	
NBT3/BAP	0.02	0.03	–63%	0.02	0.02	–15%	
NBT0/BAP	0.17	0.26	–50%	0.17	0.20	–18%	
NBT2/BC2	0.16	0.27	–73%	0.27	0.33	–21%	
NBT3/BC3	0.11	0.18	–58%	0.95	1.21	–28%	
DF/F0	0.58	0.17	70%	0.17	0.51	–207%	
AE/F0	0.15	0.58	–298%	0.21	11.88	–5501%	

most stable PAHs under a range of weathering states (Uhler and Emsbo-Mattingly, 2006). In this study, the degree of weathering in tar was estimated using the sum of 2- and 3-ring PAHs divided by the sum of 4- to 6-ring PAHs (Costa and Sauer, 2005). This study demonstrated the pronounced loss of source character for most PAHs with 2-ring, 3-rings, and different ring numbers in the numerator and denominator (see Table 9.16). To a lesser extent, the diagnostic integrity of certain 4-ring PAHs was also diminished.

The most valuable PAH ratios employed compounds that did not weather significantly (e.g., 5- and 6-ring PAHs) or weathered at comparable rates (e.g., DBT2/PA2; DBT3/PA3; FL0/PY0; BF/FP1; BA0/C0; and BA0 + C0/BC2 + BC3). In practice, the more useful PAH ratios must be selected empirically based largely on the ability of the ratio to distinguish representative end-member samples over the range of PAH concentrations and weathering states observed at the site (Douglas *et al.*, 1996).

As described previously, the tiered approach to unraveling hydrocarbon signatures employs multiple lines of evidence to identify and track source signatures of fugitive hydrocarbons in the environment. These lines of evidence include (1) the molecular (boiling) ranges of hydrocarbon types, (2) relative abundances of source specific markers, and (3) weathering effects. The methods and examples of PAH fingerprinting provided in this section provide the basis for determining the most likely origins of PAH residues in environmental samples (Mauro and Emsbo-Mattingly, 1999). These aromatic source signatures often are complemented by additional lines of evidence in the form of saturated hydrocarbons and geochemical biomarkers, described as follows.

9.5.2.3 Saturates and Biomarkers by GC/MS/SIM

Saturated hydrocarbon and geochemical biomarker distributions in petroleum and coal-derived materials can be key indicators as to the product type and source. The saturated hydrocarbons are functionally defined by their common use in the identification of petrogenic product types (i.e., the distinction between distillates, residual petroleum, and coal). Saturated hydrocarbons include normal alkanes, acyclic isoprenoids, and normal alkylcyclohexanes. The geochemical biomarkers are generally present at lower concentration but serve extremely well as weathering-resistant source signatures of petroleum and coal residues (Wang *et al.*, 1999; Stout *et al.*, 2002). These biomarkers include regular and rearranged steranes, sesqui-, di-, tri-, tetra-, and pentacyclic terpanes, and mono- and tri-aromatic steranes (see Table 9.17). The saturated hydrocarbons and geochemical biomarkers are used collectively to track source signatures of petroleum and coal derived materials in the environment.

The GC/MS operating conditions for the optimal detection, separation, and measurement of these semivolatile target compounds are based on EPA Method 8270C, *Semivolatile Organic Compounds by Gas Chromatography/Mass Spectrometry (GC/MS)* (EPA, 1997). The GC operating conditions used for the generation of high resolution hydrocarbon fingerprints and alkylated PAHs work equally well for the measurement of saturates and biomarkers. Similarly, the operation of the mass spectrometer in selected ion monitoring (SIM) mode described previously can be used for the analysis of these target

Table 9.17

Biomarker target analytes measured by GC/MS/SIM.

Biomarker Analyte	Peak ID	Group ID	Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
Rearranged C15**	Q1	Sesquiterpanes	$C_{15}H_{28}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C15**	Q2	Sesquiterpanes	$C_{15}H_{28}$	123[1]	10 ng/L	2 μg/Kg
8β(H)-drimane	Q3	Sesquiterpanes	$C_{15}H_{28}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C15**	Q4	Sesquiterpanes	$C_{15}H_{28}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C16***	Q5	Sesquiterpanes	$C_{16}H_{30}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C16***	Q6	Sesquiterpanes	$C_{16}H_{30}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C16***	Q7	Sesquiterpanes	$C_{16}H_{30}$	123[1]	10 ng/L	2 μg/Kg
8β(H)-homodrimane	Q8	Sesquiterpanes	$C_{16}H_{30}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C14*	Q9	Sesquiterpanes	$C_{14}H_{26}$	123[1]	10 ng/L	2 μg/Kg
Rearranged C14*	Q10	Sesquiterpanes	$C_{14}H_{26}$	123[1]	10 ng/L	2 μg/Kg
C23 Tricyclic Terpane	T4	Triterpanes	$C_{23}H_{42}$	191[2]	10 ng/L	2 μg/Kg
C24 Tricyclic Terpane	T5	Triterpanes	$C_{24}H_{44}$	191[2]	10 ng/L	2 μg/Kg
C25 Tricyclic Terpane	T6	Triterpanes	$C_{25}H_{46}$	191[2]	10 ng/L	2 μg/Kg
C24 Tetracyclic Terpane	T6a	Triterpanes	$C_{24}H_{42}$	191[2]	10 ng/L	2 μg/Kg
C26 Tricyclic Terpane-22S	T6b	Triterpanes	$C_{26}H_{48}$	191[2]	10 ng/L	2 μg/Kg
C26 Tricyclic Terpane-22R	T6c	Triterpanes	$C_{26}H_{48}$	191[2]	10 ng/L	2 μg/Kg
C28 Tricyclic Terpane-22S	T7	Triterpanes	$C_{28}H_{52}$	191[2]	10 ng/L	2 μg/Kg
C28 Tricyclic Terpane-22R	T8	Triterpanes	$C_{28}H_{52}$	191[2]	10 ng/L	2 μg/Kg
C29 Tricyclic Terpane-22S	T9	Triterpanes	$C_{29}H_{54}$	191[2]	10 ng/L	2 μg/Kg
C29 Tricyclic Terpane-22R	T10	Triterpanes	$C_{29}H_{54}$	191[2]	10 ng/L	2 μg/Kg
18α(H)-22,29,30-Trisnorneohopane-TS	T11	Triterpanes	$C_{27}H_{46}$	191[2]	10 ng/L	2 μg/Kg
17α(H)-22,29,30-Trisnorhopane-TM	T12	Triterpanes	$C_{27}H_{46}$	191[2]	10 ng/L	2 μg/Kg
17α(H),21β(H)-28,30-Bisnorhopane & 17β(H),21β(H)-28,30-Bisnorhopane	T14a	Triterpanes	$C_{28}H_{48}$	191[2]	10 ng/L	2 μg/Kg
17α(H),21β(H)-25-Norhopane	T14b	Triterpanes	$C_{29}H_{50}$	191[2]	10 ng/L	2 μg/Kg
30-Norhopane	T15	Triterpanes	$C_{29}H_{50}$	191[2]	10 ng/L	2 μg/Kg

*tetra = methyldecahydronaphthalene.
**penta = methyldecahydronaphthalene.
***hexa = methyldecahydronaphthalene.

Table 9.17
Continued

Biomarker Analyte	Peak ID	Group ID	Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
18α(H)-30-Norneohopane-C29Ts	T16	Triterpanes	$C_{29}H_{50}$	191[2]	10 ng/L	2 μg/Kg
17α(H)-Diahopane	X	Triterpanes	$C_{30}H_{52}$	191[2]	10 ng/L	2 μg/Kg
30-Normoretane	T17	Triterpanes	$C_{29}H_{50}$	191[2]	10 ng/L	2 μg/Kg
18α(H) & 18β(H)-Oleananes	T18	Triterpanes	$C_{30}H_{52}$	191[2]	10 ng/L	2 μg/Kg
Hopane	T19	Triterpanes	$C_{30}H_{52}$	191[2]	10 ng/L	2 μg/Kg
Moretane	T20	Triterpanes	$C_{30}H_{52}$	191[2]	10 ng/L	2 μg/Kg
30-Homohopane-22S	T21	Triterpanes	$C_{31}H_{54}$	191[2]	10 ng/L	2 μg/Kg
30-Homohopane-22R	T22	Triterpanes	$C_{31}H_{54}$	191[2]	10 ng/L	2 μg/Kg
30,31-Bishomohopane-22S	T26	Triterpanes	$C_{32}H_{56}$	191[2]	10 ng/L	2 μg/Kg
30,31-Bishomohopane-22R	T27	Triterpanes	$C_{32}H_{56}$	191[2]	10 ng/L	2 μg/Kg
30,31-Trishornohopane-22S	T30	Triterpanes	$C_{33}H_{58}$	191[2]	10 ng/L	2 μg/Kg
30,31-Trishornohopane-22R	T31	Triterpanes	$C_{33}H_{58}$	191[2]	10 ng/L	2 μg/Kg
Tetrakishomohopane-22S	T32	Triterpanes	$C_{34}H_{60}$	191[2]	10 ng/L	2 μg/Kg
Tetrakishomohopane-22R	T33	Triterpanes	$C_{34}H_{60}$	191[2]	10 ng/L	2 μg/Kg
Pentakishornohopane-22S	T34	Triterpanes	$C_{35}H_{62}$	191[2]	10 ng/L	2 μg/Kg
Pentakishornohopane-22S	T35	Triterpanes	$C_{35}H_{62}$	191[2]	10 ng/L	2 μg/Kg
13β(H),17α(H)-20S-Diacholestane	S4	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
13β(H),17α(H)-20R-Diacholestane	S5	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
13β(H),17α(H)-20S-Methyldiacholestane	S8	Steranes	$C_{28}H_{50}$	217[3]	10 ng/L	2 μg/Kg
14α(H),17α(H)-20S-Cholestane	S12	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
14α(H),17α(H)-20R-Cholestane	S17	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
13β(H),17α(H)-20R-Ethyldiacholestane	S18	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg
13α(H),17β(H)-20S-Ethyldiacholestane	S19	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg
14α(H),17α(H)-20S-Methylcholestane	S20	Steranes	$C_{28}H_{50}$	217[3]	10 ng/L	2 μg/Kg
14α(H),17α(H)-20R-Methylcholestane	S24	Steranes	$C_{28}H_{50}$	217[3]	10 ng/L	2 μg/Kg
14α(H),17α(H)-20S-Ethylcholestane	S25	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg

Table 9.17

Continued

Biomarker Analyte	Peak ID	Group ID	Formula	Quant Ion (m/z)	Approximate Detection Limit	
					Water	Solid
14α(H),17α(H)-20R-Ethylcholestane	S28	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20R-Cholestane	S14	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20S-Cholestane	S15	Steranes	$C_{27}H_{48}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20R-Methylcholestane	S22	Steranes	$C_{28}H_{50}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20S-Methylcholestane	S23	Steranes	$C_{28}H_{50}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20R-Ethylcholestane	S26	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg
14β(H),17β(H)-20S-Ethylcholestane	S27	Steranes	$C_{29}H_{52}$	217[3]	10 ng/L	2 μg/Kg
C21-Monoaromatic Sterane	MAS1	Monoaromatic Steranes	$C_{21}H_{22}$	253[3]	10 ng/L	2 μg/Kg
C22-Monoaromatic Sterane	MAS2	Monoaromatic Steranes	$C_{22}H_{24}$	253[3]	10 ng/L	2 μg/Kg
C27-5β(H)-Monoaromatic Sterane (20S)	MAS3	Monoaromatic Steranes	$C_{27}H_{42}$	253[3]	10 ng/L	2 μg/Kg
C27-5β(H)-Monoaromatic Sterane (20R)	MAS4	Monoaromatic Steranes	$C_{27}H_{42}$	253[3]	10 ng/L	2 μg/Kg
C27-5α(H)-Monoaromatic Sterane (20S)	MAS5	Monoaromatic Steranes	$C_{27}H_{42}$	253[3]	10 ng/L	2 μg/Kg
C28-5β(H)-Monoaromatic Sterane (20S)	MAS6	Monoaromatic Steranes	$C_{28}H_{44}$	253[3]	10 ng/L	2 μg/Kg
C27-5α(H)-Monoaromatic Sterane (20R) & C28-5α(H)-Monoaromatic Sterane (20S)	MAS7	Monoaromatic Steranes	$C_{27}H_{42}/C_{28}H_{44}$	253[3]	10 ng/L	2 μg/Kg
C28-5α(H)-Monoaromatic Sterane (20R) C29-5β(H)-Monoaromatic Sterane (20S)	MAS8	Monoaromatic Steranes	$C_{28}H_{44}/C_{29}H_{46}$	253[3]	10 ng/L	2 μg/Kg
C29-5α(H)-Monoaromatic Sterane (20S)	MAS9	Monoaromatic Steranes	$C_{29}H_{46}$	253[3]	10 ng/L	2 μg/Kg
C28-5α(H)-Monoaromatic Sterane (20R) & C29-5β(H)-Monoaromatic Sterane (20R)	MAS10	Monoaromatic Steranes	$C_{28}H_{44}/C_{29}H_{46}$	253[3]	10 ng/L	2 μg/Kg
C29-5α(H)-Monoaromatic Sterane (20R)	MAS11	Monoaromatic Steranes	$C_{29}H_{46}$	253[3]	10 ng/L	2 μg/Kg
C20-Triaromatic Steroid Hydrocarbon	TAS1	Triaromatic Steranes	$C_{20}H_{20}$	231[3]	10 ng/L	2 μg/Kg
C21-Triaromatic Steroid Hydrocarbon	TAS2	Triaromatic Steranes	$C_{21}H_{22}$	231[3]	10 ng/L	2 μg/Kg
C26-Triaromatic Steroid Hydrocarbon (20S)	TAS3	Triaromatic Steranes	$C_{26}H_{32}$	231[3]	10 ng/L	2 μg/Kg
C26-Triaromatic Steroid Hydrocarbon (20R) & C27-Triaromatic Steroid Hydrocarbon (20S)	TAS4	Triaromatic Steranes	$C_{26}H_{32}/C_{27}H_{34}$	231[3]	10 ng/L	2 μg/Kg
C28-Triaromatic Steroid Hydrocarbon (20S)	TAS5	Triaromatic Steranes	$C_{28}H_{36}$	231[3]	10 ng/L	2 μg/Kg
C27-Triaromatic Steroid Hydrocarbon (20R)	TAS6	Triaromatic Steranes	$C_{27}H_{34}$	231[3]	10 ng/L	2 μg/Kg
C28-Triaromatic Steroid Hydrocarbon (20R)	TAS7	Triaromatic Steranes	$C_{28}H_{36}$	231[3]	10 ng/L	2 μg/Kg

Notes:
[1]Quantified using RF of trans-Decalin.
[2]Quantified using RF of Hopane.
[3]Quantified using RF of 5β(h)cholane.

analytes. Obviously, the GC/MS/SIM method used for saturates and biomarkers requires different monitored ions (primary ions listed in Table 9.17). The abbreviated mass spectra pattern and well-documented retention indices for each class of compounds can be used to identify and quantify these weathering resistant compounds (see Figure 9.19; Peters *et al.*, 2005; Stout *et al.*, 2000a; Wang *et al.*, 2006). In practice, the alkylated PAH and biomarker analysis can be carried out simultaneously for hydrocarbon impacted samples. Samples containing high levels of PAHs (e.g., pyrogenic products such as tar) should be fractionated using silica gel and the saturate fraction (often called F1) analyzed for saturates and biomarkers. Two types of calibration standards are appropriate for determination of biomarker concentrations. The first type of calibration standard includes retention time markers (e.g., 5β(H)-cholane) for the determination of expected biomarker retention times (Sauer and Boehm, 1995). The second type of calibration standard includes discrete biomarker chemicals (e.g., 17α(H),21β(H)-hopane) that can be used to generate multiconcentration calibration curves and response factors for individual and structurally similar compounds (Stout *et al.*, 2001; Douglas *et al.*, 2002).

The distinct features of saturated hydrocarbon and geochemical biomarker fingerprints are evident by comparison between a crude oil (see Figure 9.27) and crude tar (see Figure 9.28). These reference samples were discussed previously (see Figures 9.23 and 9.24) and are presented again with the full complement of normal alkanes, acyclic isoprenoids, *n*-alkylcyclohexanes, sesquiterpanes, steranes, triterpanes, monoaromatic steranes (MAS), and triaromatic steranes (TAS). The saturated hydrocarbons in the crude oil (see Figure 9.27b) span a wider molecular weight range than the coal feedstock used to make the coal tar sample (Figure 9.28b). One of the other important source-specific differences among the saturate hydrocarbons is lower relative abundance of pristane (Pr) relative to phytane (Ph) in the crude oil (see Figure 9.27b and d) when compared to the coal tar feedstock (see Figure 9.28b and d). In addition, the fingerprint of the *n*-alkylcyclohexanes is wider for the crude oil (see Figure 9.27c) and more Gaussian around the elution range of n-tridecane (n-C_{13}) for the coal tar (see Figure 9.28c). Although this feature also is observed in the *n*-alkane pattern, it is retained longer in the *n*-alkylcyclohexane pattern as the *n*-alkanes biodegrade.

It is acknowledged that weathering affects the saturated hydrocarbon patterns. When the petroleum is released into the environment the lighter hydrocarbons usually evaporate or, in some circumstances, dissolve. In addition, biodegradation preferentially reduces the normal alkanes and subsequently *n*-alkylcyclohexanes more rapidly than isoprenoids (Kaplan *et al.*, 1997; Prince, 2002). In time, even the isoprenoids can degrade, leaving the unresolved complex mixture (UCM) with few recognizable features; however, this stage of

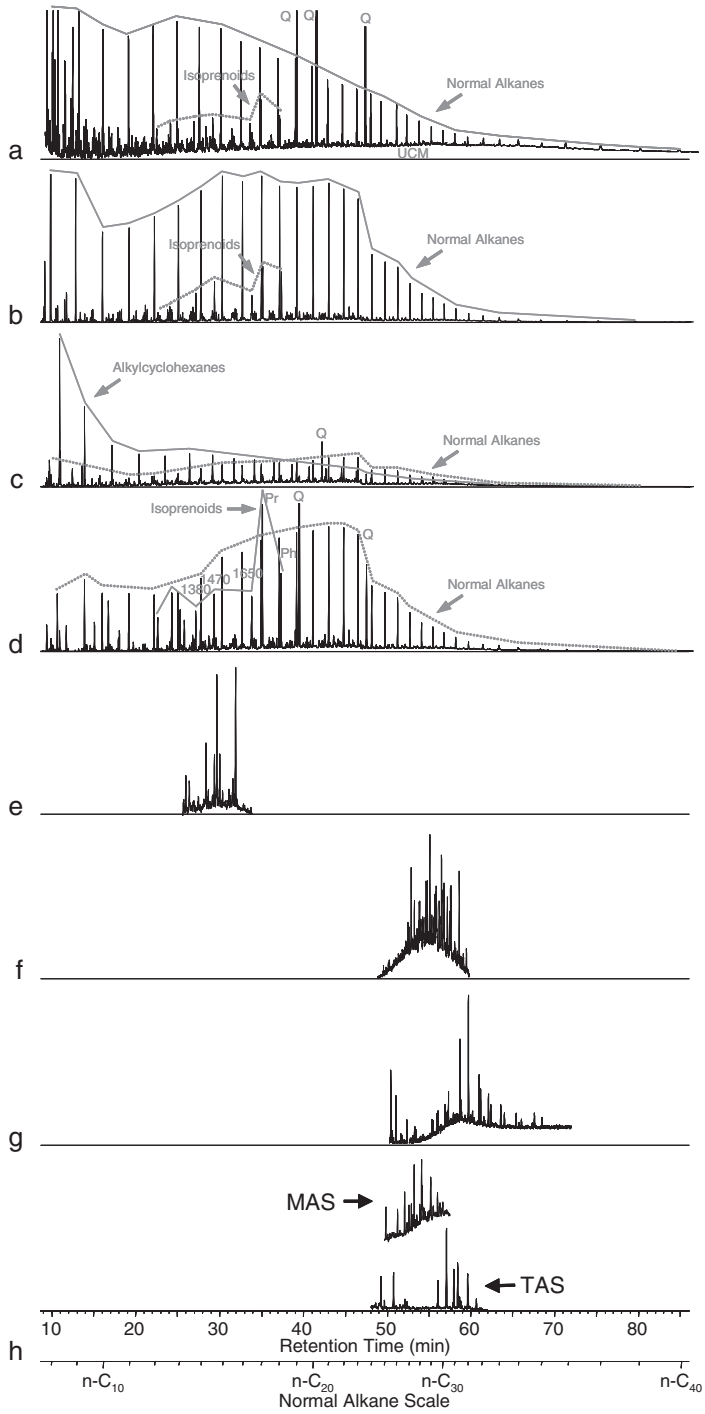


Figure 9.27

Crude oil signatures revealed by high resolution hydrocarbon fingerprints (GC/FID), saturated hydrocarbons and biomarkers (GC/MS).
***a.** Crude oil (GC/FID) Dominated by normal alkanes with lower amounts of acyclic isoprenoid hydrocarbons.*
***b.** Normal alkanes (GC/MS/SIM, m/z 85) Full range of normal alkanes and lower amounts of acyclic isoprenoid hydrocarbons.*
***c.** N-Alkylcyclohexanes (GC/MS/SIM, m/z 83) Full range of ACHs imperceptible on FID.*
***d.** Acyclic isoprenoids (GC/MS/SIM, m/z 113) Although evident with other techniques, m/z resolves the isoprenoids very clearly.*
***e.** Sesquiterpanes (GC/MS/SIM, m/z 123) Full pattern of diesel range sesquiterpanes exhibits the original crude oil signature.*
***f.** Steranes (GC/MS/SIM, m/z 217) Full pattern of heavy range steranes exhibits the original crude oil signature.*
***g.** Triterpanes (GC/MS/SIM, m/z 191) Full pattern of heavy range triterpanes exhibits the original crude oil signature.*
***h.** Mono & triaromatic steranes (GC/MS/SIM) Full pattern of heavy range MAS & TAS exhibits the original crude oil signature.*

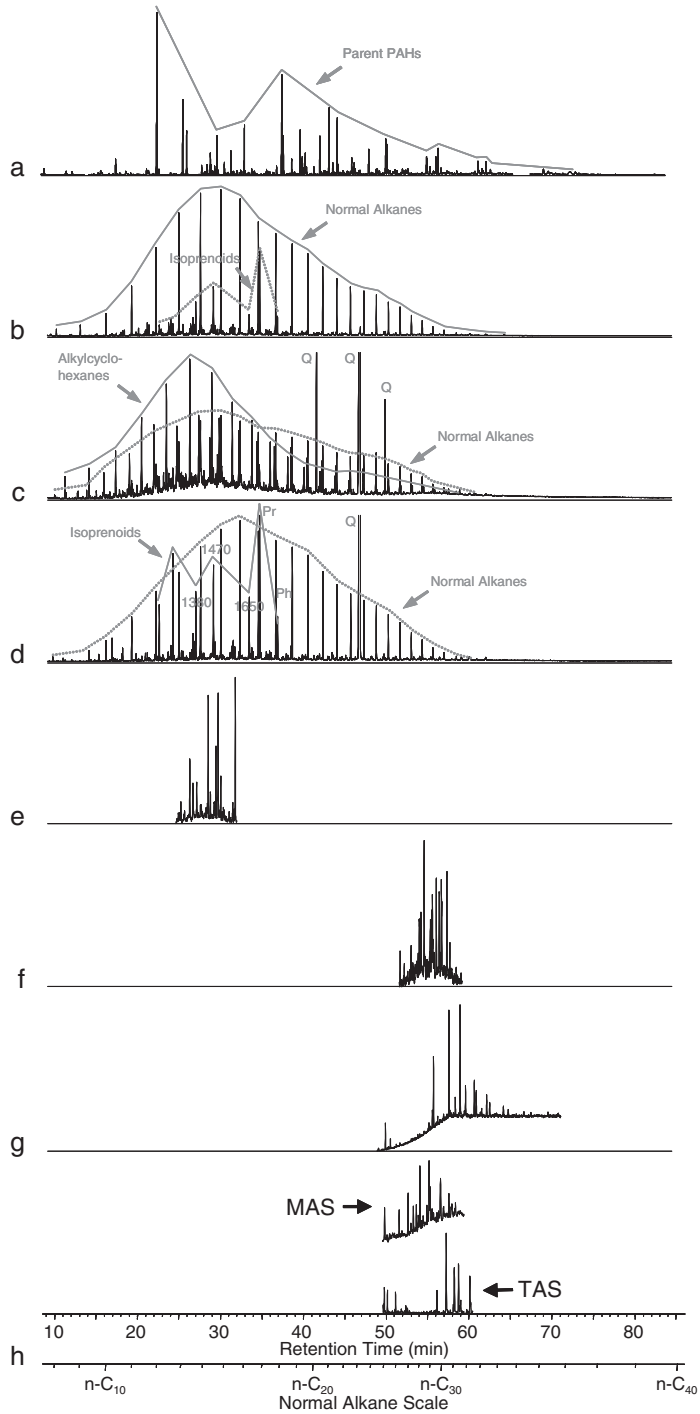


Figure 9.28

Coal tar signatures revealed by high resolution hydrocarbon fingerprints (GC/FID), saturated hydrocarbons and biomarkers (GC/MS).
a. *Crude oil (GC/FID) Dominated by normal alkanes with lower amounts of acyclic isoprenoid hydrocarbons.*
b. *Normal alkanes (GC/MS/SIM, m/z 85) Wide range of normal alkanes and lower amounts of acyclic isoprenoid hydrocarbons from residual feedstock.*
c. *N-Alkylcyclohexanes (GC/MS/SIM, m/z 83) Full range of ACHs imperceptible on FID also from residual feedstock.*
d. *Acyclic isoprenoids (GC/MS/SIM, m/z 113) Although evident with other techniques, m/z resolves the isoprenoids very clearly.*
e. *Sesquiterpanes (GC/MS/SIM, m/z 123) Full pattern of diesel range sesquiterpanes exhibits the original feedstock signature.*
f. *Steranes (GC/MS/SIM, m/z 217) Full pattern of heavy range steranes exhibits the original feedstock signature.*
g. *Triterpanes (GC/MS/SIM, m/z 191) Full pattern of heavy range triterpanes exhibits the original feedstock signature.*
h. *Mono & triaromatic steranes (GC/MS/SIM) Full pattern of heavy range MAS & TAS exhibits the original feedstock signature.*

weathering is uncommon among most field sites. Through the various stages of weathering, the saturated hydrocarbon profiles, and particularly, the late eluting contour of the UCM, can serve as one of the more easily recognizable features of fugitive petroleum. These weathering processes may not occur to any significant degree in coal residues because saturated hydrocarbons often are protected within the complex physical structure of coal. In summary, the saturated hydrocarbons help identify the type of petroleum or coal in the sample with the added capability of demonstrating the influence of weathering. When weathered, these signatures benefit from additional lines of evidence based on more recalcitrant source signatures (Prince and Douglas, 2005).

The relative abundances of specific biomarkers help geochemists identify fossil fuel reservoirs based on the mixture of initial organic matter (e.g., types of land and marine biomass), depositional environments (e.g., toxic condition and mineral substrates), and conditions of formation (e.g., temperature, pressure, and residence time) (Peters *et al.*, 2005). Having survived millennia under heated and pressurized conditions, geochemical biomarkers resist the degradation processes that occur in most natural environments and environmental timescales. These biomarkers are capable of differentiating the origins of heavy hydrocarbons (Stout *et al.*, 1999c) even when most of the more labile compounds degrade. However, they are generally present at lower concentrations than the saturated hydrocarbons and require sensitive analytical methods for accurate measurement (e.g., GC/MS/SIM). The sesquiterpane patterns of several middle distillate fuels demonstrate well the disparate geochemical signatures of the different crude oil feedstocks (notice the pattern differences between Figures 9.27e and 9.28e). The crude oil feedstock signature of heavier hydrocarbons is evident in the steranes (see Figures 9.27f and 9.28f), triterpanes (see Figures 9.27g and 9.28g), MAS and TAS (see Figures 9.27h and 9.28h) patterns. In practice, the triterpanes offer excellent discriminatory power, but the steranes also offer additional tools for distinguishing sources of heavy hydrocarbons as needed.

A detailed inspection of the geochemical biomarker patterns reveals many similar features that result from the common geochemical processes of formation; however, many differences in the relative abundances of these compounds reflect source-specific variances based on the reservoir signature and refining effects (Peters *et al.*, 1992). A selection of reference samples discussed previously in the high resolution hydrocarbon fingerprinting (see Figure 9.20) and PAH (see Figure 9.26) sections help illustrate the diversity of biomarkers signatures as represented among the triterpanes (see Figure 9.29). The most obvious differences between the crude oil and selected petroleum products are the abundance of tricyclic triterpanes relative to pentacyclic triterpanes, which is intermediate in the crude oil (see Figure 9.29a), high in

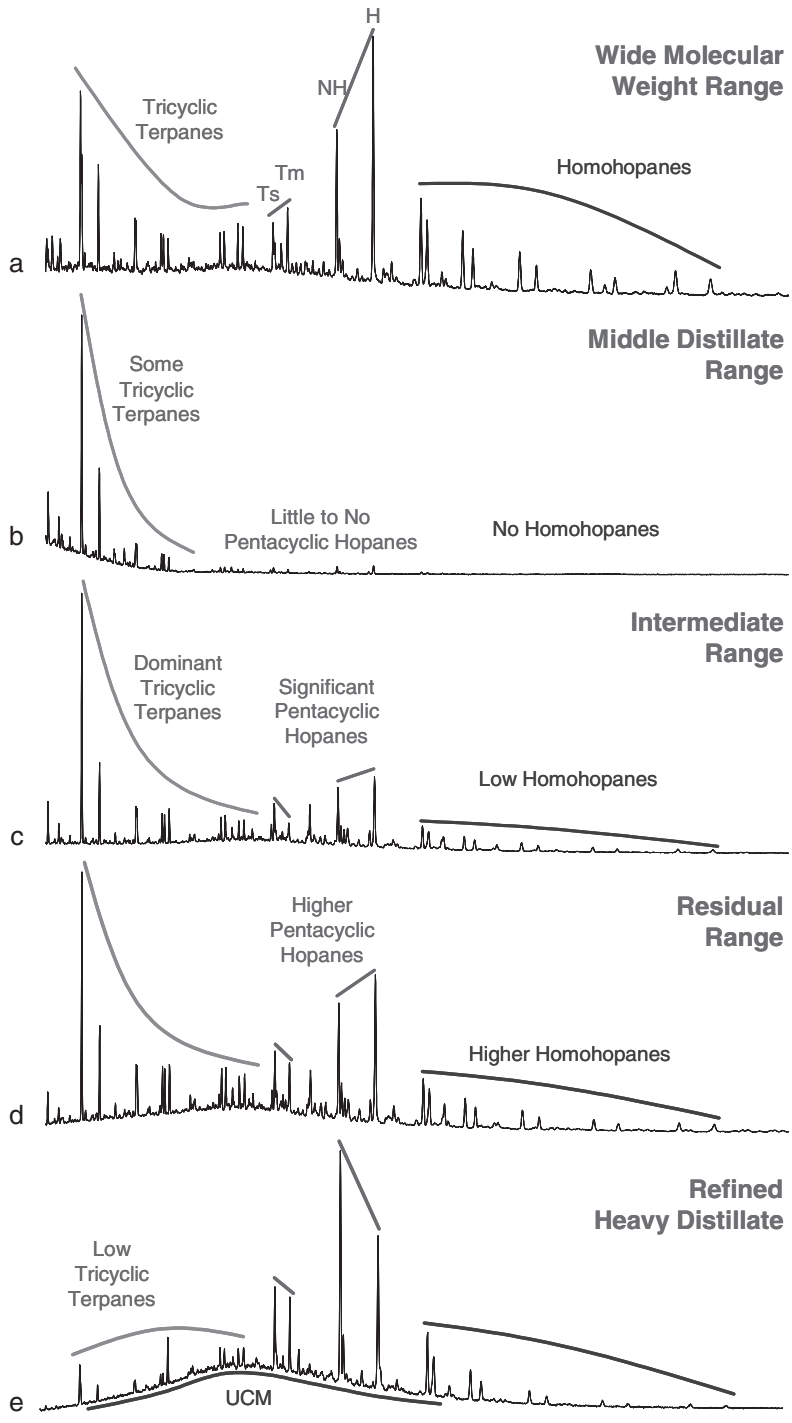


Figure 9.29
Selected triterpane fingerprints of hydrocarbon sources.
a. *Crude oil*
Heavy petroleum fraction contains a full range of triterpane biomarkers ranging from the lighter tricyclic to the heavier homohopane series. The relative abundances of biomarkers, like Ts:Tm and NH:H help identify the petroleum origin.
b. *#2 fuel oil*
Middle petroleum distillates contain only the lightest, if any, triterpane biomarkers.
c. *#4 fuel oil*
Intermediate fuel oils blend middle and heavy distillates. The levels of tricyclic triterpanes are generally more abundant than pentacyclic triterpanes.
d. *#6 fuel oil*
Heavy fuel oils generally retain the full triterpane signature of the crude oil.
e. *Lube oil*
Heavy refined distillates primarily retain a portion of the triterpane pattern commensurate with the distillation range (see UCM).

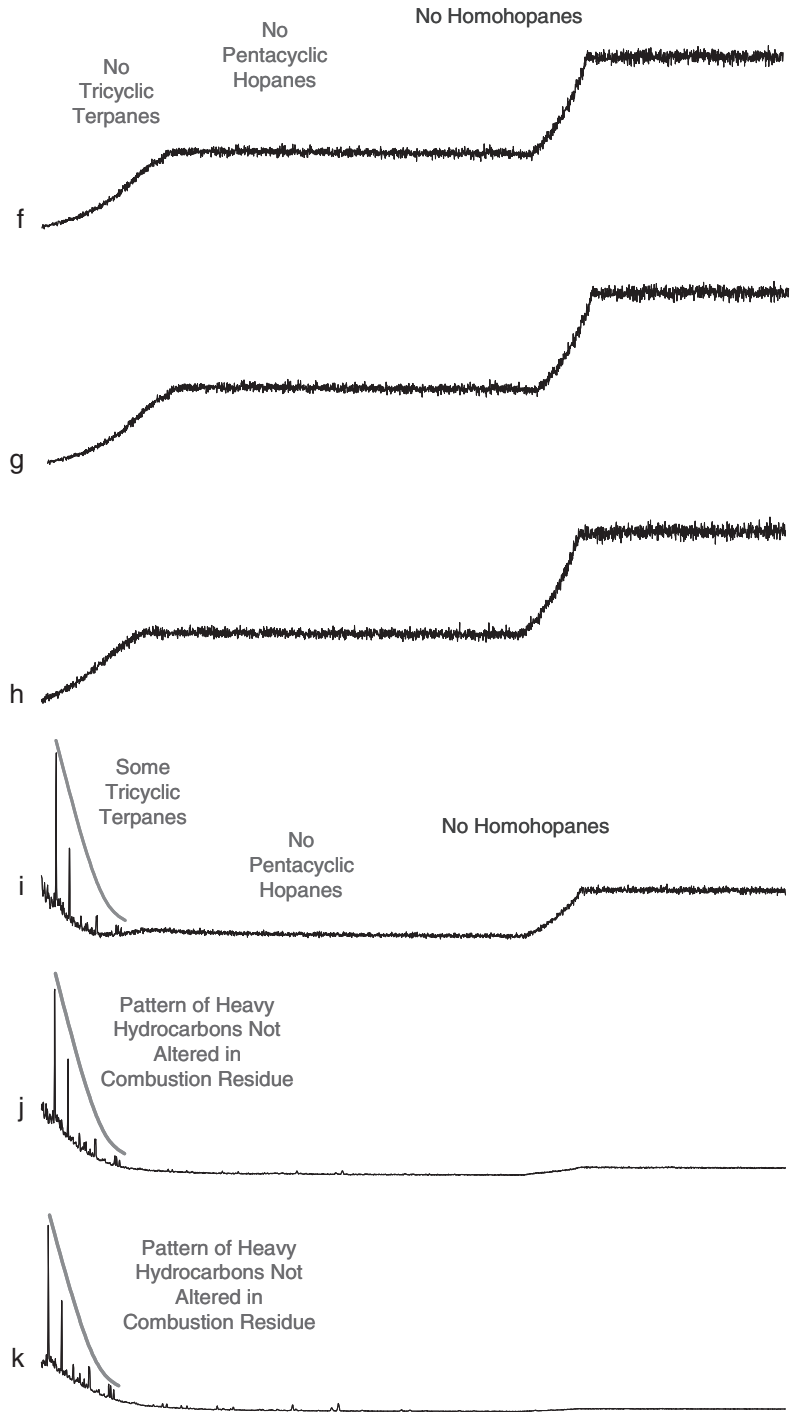


Figure 9.29

f. *Kerosene*
Light middle distillates, like kerosene, do not contain heavy petroleum hydrocarbons in the triterpane range.

g. *Kerosene*
50% combusted
Combustion causes the loss of lighter hydrocarbons and concentration of heavier hydrocarbons. Even so, no triterpanes appear.

h. *Kerosene*
Combustion residue
No triterpanes appear when the heaviest fraction of kerosene is concentrated to its maximum limit.

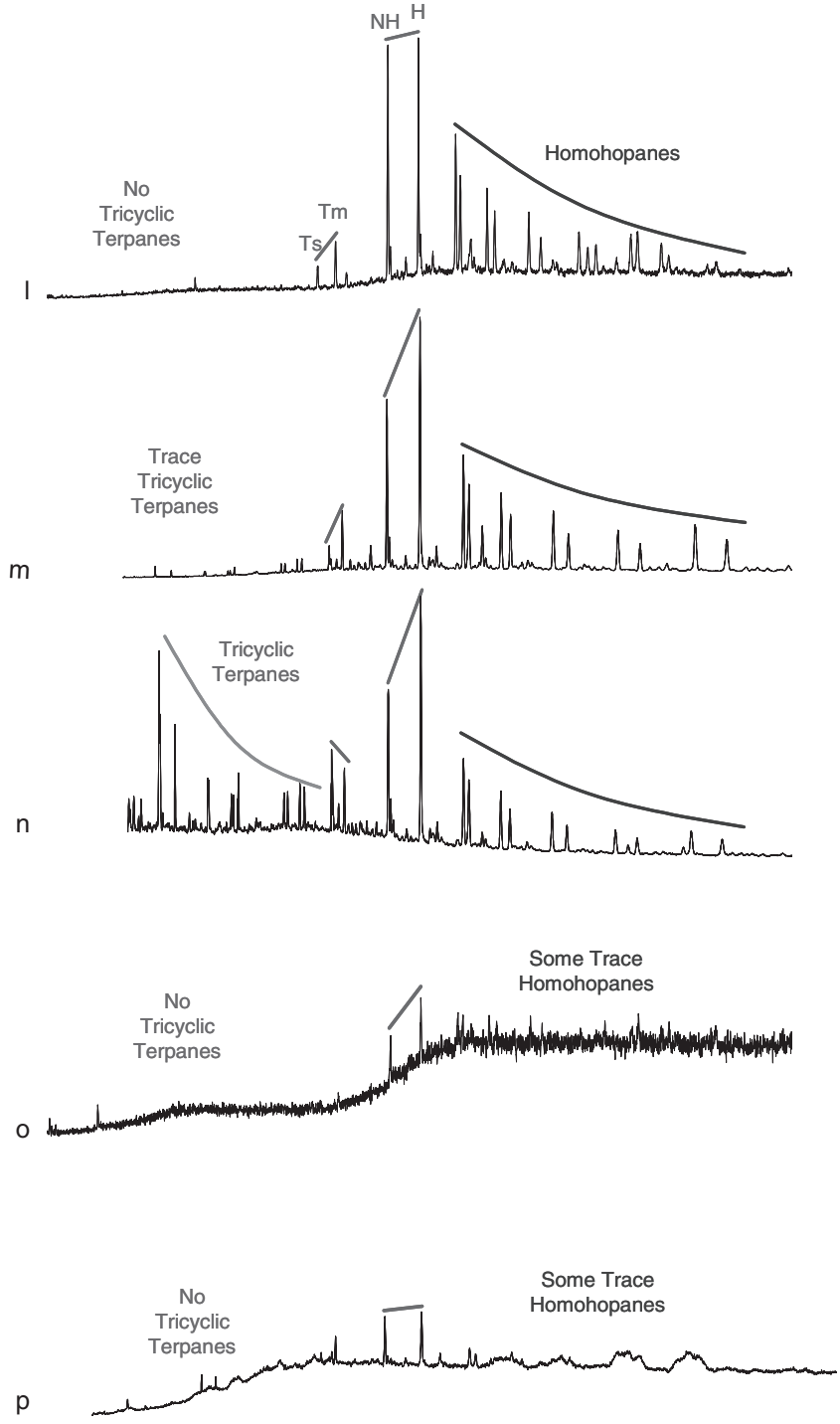
i. *Diesel*
Diesel range middle petroleum distillate contains low levels of the lightest tricyclic triterpanes.

j. *Diesel*
50% combusted
The triterpane pattern remains unaltered when half of the diesel range hydrocarbons are combusted.

k. *Diesel*
Combustion residue
The triterpane pattern remains unaltered when the heaviest (<1%) middle distillate hydrocarbons are isolated in the combustion residue.

Figure 9.29—cont'd

l. *Hot patch*
Paving material is formulated to remain solid at ambient roadway conditions. Distillation of middle and heavy petroleum distillates leaves only the heaviest triterpanes in the petroleum pitch binder.
m. *Roadway pavement (c. 2005)*
The heavy petroleum fraction in modern roadway pavement is from the same residual fraction as the hot patch sample.
n. *Roadway pavement (c. 1940s)*
The residual petroleum used to pave roads often contained lighter hydrocarbon fractions relative to modern residual petroleum products.
o. *Roadway pavement (c. 1940s)*
Variable levels of residual petroleum are found in pavement laid down in the early 1900s. None is evident in this sample. The binder in the pavement was tar.
p. *Coal tar oil*
Coal tar oil contains low levels of triterpane biomarkers. The concentrations in this coal tar oil are consistent with the roadway pavement sample illustrated above.



#2 Fuel Oil (see Figure 9.29b), and low in the heavy petroleum products (see Figure 9.29c and d). Among other indicators, a distinct feature of the lube oil sample (see Figure 9.29e) is the high ratio of C_{30}-17α(H),21β(H)norhopane (NH) to C_{30}-17α(H),21β(H)hopane (H) relative to the previously described petroleum products (see Figure 9.29a, b, c, and d). As expected, middle distillate, like kerosene and diesel, lack specific portions of the triterpane pattern as a result of refinery distillation and this pattern remains constant despite the loss of 99% of the fuel mass (see Figure 9.29f to k). By contrast, heavy petroleum residuals in roadway pavement often lack the lighter tricyclic triterpanes due to aggressive distillation techniques associated with modern refinery practices (see Figure 9.29l and m). Some pavement, especially older pavement, contains the tricyclic triterpane biomarkers that were not routinely removed at older refineries (see Figure 9.29n). Each of these residual petroleum samples exhibits unique biomarker signatures as evident in the various combinations of Ts:Tm and NH:H concentration ratios (and others). By contrast, some older pavement (see Figure 9.29o) was made with a tar-based binder (see Figure 9.29p) that contained extremely low levels of triterpanes. Collectively, the molecular weight distributions and relative abundance of diagnostic ratios among biomarkers demonstrate the complementary source identification value of fingerprinting with geochemical biomarkers—and thus, the broad application of geochemical biomarkers in most forensic investigations (Wang *et al.*, 2006).

9.5.3 ANCILLARY HYDROCARBON FINGERPRINTING TECHNIQUES

The chemical fingerprinting techniques for hydrocarbons just described have found the greatest application in environmental forensics over the past 10 years. However, other analytical methods, including some newly being applied to environmental studies, have also been applied to environmental forensics with varying degrees of success and maturity. In the following sections, some of these methods are discussed.

9.5.3.1 Conventional Simulated Distillation

Distillation analysis is a means to determine the boiling range distribution of petroleum or tar products. Simulated distillation utilized gas chromatography (GC) as a convenient means to determine the boiling range of hydrocarbons in a sample. In addition to providing the environmental forensics investigator with information about the composition of hydrocarbons under study, simulated distillation curves are often the most convenient graphical means to convey differences in petroleum or tar product composition to nontechnical decision-makers (see Figure 9.30).

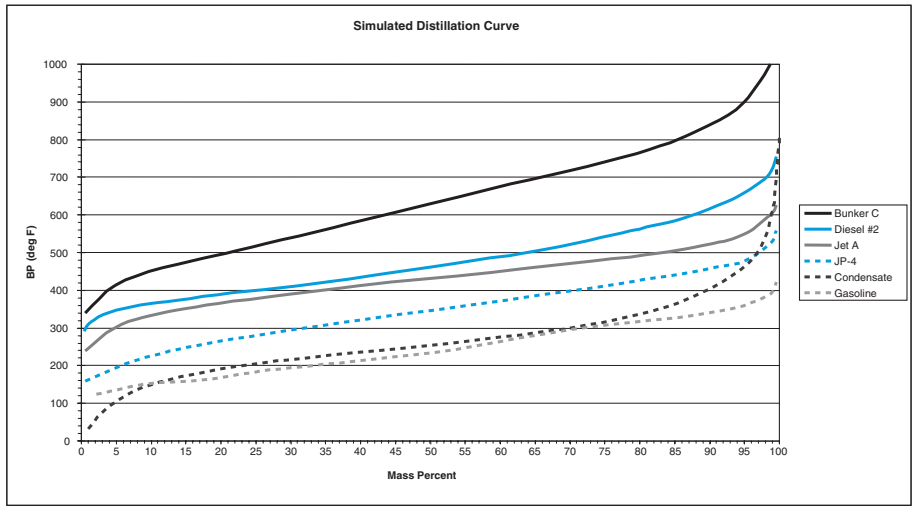


Figure 9.30
Simulated distillation curves for Gasoline, Condensate, JP-4, Jet-A, Diesel Fuel #2, and Bunker C fuel. The distinct boiling point distributions for each product can be clearly recognized. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

The standard method for determining the boiling range of hydrocarbon products by GC is ASTM Method D2887; however, other carefully calibrated GC methods can be used to reconstruct simulated distillation curves following the same guidelines presented in ASTM D2887. For example, GC analysis following EPA Method 8015B *Nonhalogenated Organics Using GC/FID* (EPA, 1997) can be used to reconstruct simulated distillation curves. In order to use EPA Method 8015B for simulated distillation analysis, the analytical chemist must calibrate the gas chromatograph with an *n*-alkane solution containing hydrocarbons that cover the distillation range of interest. From this data, a calibration plot of boiling point versus retention time is constructed for the hydrocarbons contained in the calibration mixture. Authentic samples are analyzed under the same chromatographic conditions, and the area under the chromatographic curve between each calibration point is determined. The incremental areas are summed to determine a cumulative chromatographic area; fractional amounts of hydrocarbons from each boiling increment are thus calculated. Finally, a plot of cumulative evaporated hydrocarbon versus boiling point (determined from the initial calibration curve) is constructed. Reference standards (e.g., pure gasoline, diesel fuel, tar distillate, etc.) can be analyzed in conjunction with the field samples as a measure of method quality control.

9.5.3.2 High Temperature Simulated Distillation

Gas chromatographic fingerprinting typically is conducted using conventional techniques capable of describing hydrocarbon mixtures with boiling ranges between about C_8–C_{44} (see Section 9.5.2.1). The upper boiling limit of such

methods are constrained by the thermal stability of the GC column's stationary phase (e.g., EPA Methods 8015B, 8270C, and ASTM D2887). In certain forensics investigations that deal with very high boiling petroleum or tar residues, it is advantageous to characterize hydrocarbons in samples with boiling points well in excess of conventional method's upper temperature/boiling point limit. High temperature simulated distillation techniques are available to meet this analytical need (Villalanti *et al.*, 2000; Raia *et al.*, 2004).

High temperature simulated distillation (HTSD) is an adaptation of conventional GC/FID that utilizes a nonpolar, low bleed stationary phase column (0.05–0.15 μm film thickness and 0.53 mm internal diameter), which safely can withstand temperatures up to 430˚C. The high phase ratio of the column permits the elution of hydrocarbons 260 to 316˚C below their true boiling points, and thereby permits elution and quantification of the percent mass of hydrocarbons with boiling points up to 750˚C—that is, up to C_{120}. When combined with cryogenic initial GC conditions, HTSD can expand the carbon range fingerprint from C_5–C_{120} (see Figure 9.31). Some loss of chromatographic resolution of hydrocarbons in the C_5–C_{44} range is experienced by HTSD, but this is overcome when conventional GC and HTSD fingerprinting are used in parallel (see Figure 9.32). The combination of conventional and high temperature GC provides added information in forensic investigations involving high boiling point contamination, such as heavy fuels and crude oils, petroleum asphalts, waxes, greases, and petroleum or tar residue. The HTSD method is amenable to soils, sediments, and products (including semisolids and solids).

9.5.3.3 Stable Isotopes of Bulk Hydrocarbons

Hydrocarbons derived from petroleum, coal, or recently formed biogenic material have intrinsic ratios of the stable isotopes of the carbon and hydrogen atoms that comprise each molecule (Sofer, 1984). As a bulk property, the stable isotope ratio for a given element in a hydrocarbon product or assemblage reflect the many different geochemical and biological processes to which the hydrocarbons were exposed during their formation, refining, or processing, and environmental degradation/fractionation (Whittaker *et al.*, 1995). Measurement of stable isotopes has conventionally been applied to bulk hydrocarbon fractions, for example, aliphatic or aromatic hydrocarbons separately, and this bulk level of fingerprinting has proven useful in some forensic investigations. For example, these intrinsic isotopic features of petroleum have, in the case of petroleums, been shown to be useful in distinguishing between different source(s) of petroleums in free product and in ground water (Kaplan *et al.*, 1997). Perhaps the greatest utility of bulk stable isotopes in hydrocarbon fingerprinting is in the characterization of hydrocarbon gases

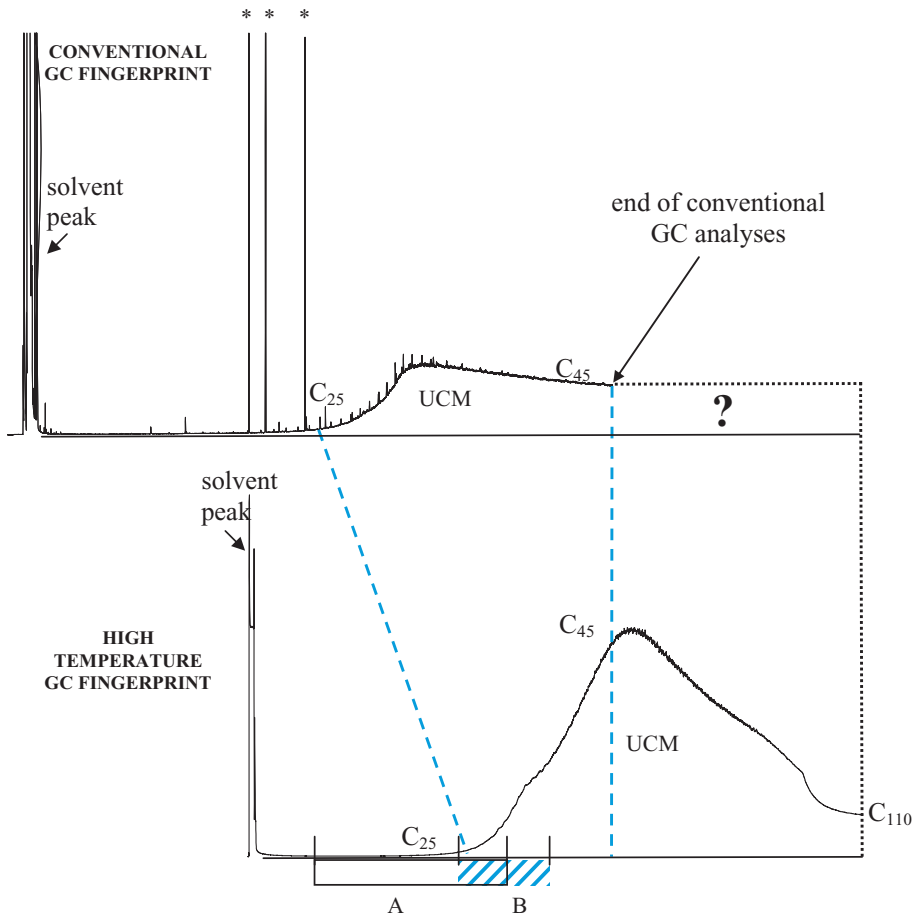


Figure 9.31

*Examples of conventional and high temperature GC fingerprints for an asphaltic material demonstrating the carbon/boiling range covered by each analysis. See text for description. * = internal standards; A = boiling range of PAH; B = boiling range of biomarkers; UCM = unresolved complex mixture.*

(e.g., methane), in which there is little molecular fingerprinting information available (Lundegard *et al.*, 1998; Lundegard, 2006).

Analytically, carbon and hydrogen isotopes are measured using an isotope ratio mass spectrometer. The isotope ratios measured for carbon and hydrogen are reported relative to international standards, in units of per mil (‰). A subsample of petroleum (pure or extracted from a sample) is combusted to produce CO_2 and H_2O. The carbon dioxide is purified, and the ratio of $^{13}C/^{12}C$ is determined. Similarly, the water produced during the combustion reaction is converted to hydrogen gas with a catalyst (Ehleringer and Rundel, 1989). The deuterium to hydrogen ratio is measured for the resulting gas. Typically, isotope ratios for carbon and hydrogen can be measured with a precision of ±0.2‰ and ±2.0‰, respectively. An example of the application of bulk stable carbon isotopes for resolution of NAPL and potential source samples is presented in Figure 9.33.

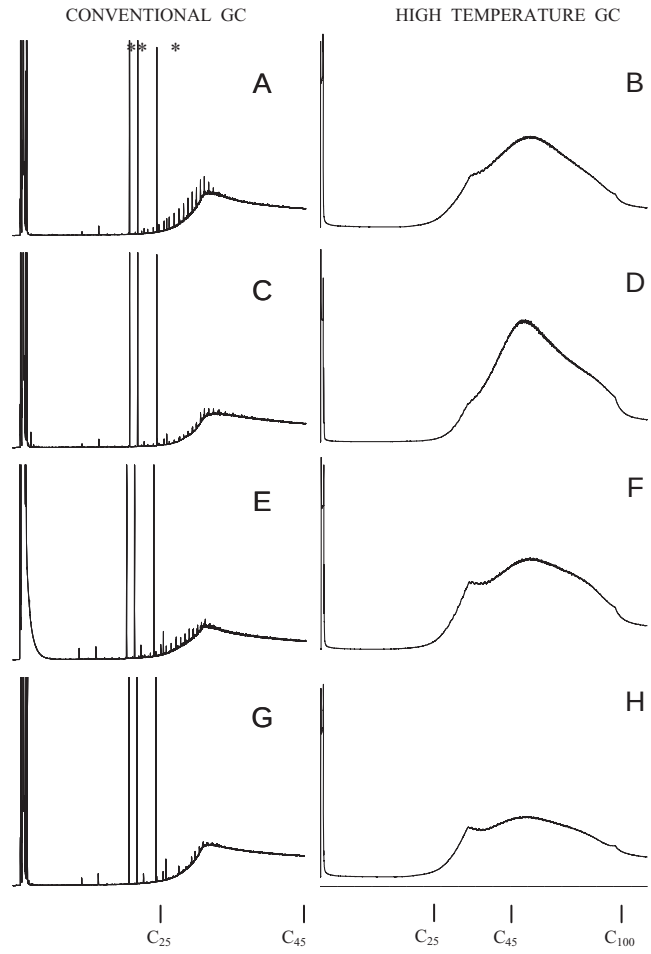


Figure 9.32

*Conventional and high temperature GC/FID chromatograms showing representatives of the asphalts and asphaltic materials found at an asphalt investigation site. (**A–B**) asphalt with prominent n-alkanes near Site A, (**C–D**) asphalt in the laboratory, (**E–F**) mineral rubber from Site B, and (**E–F**) asphaltic material near Site C with no n-alkanes. Carbon ranges depicted at bottom. * = internal standards.*

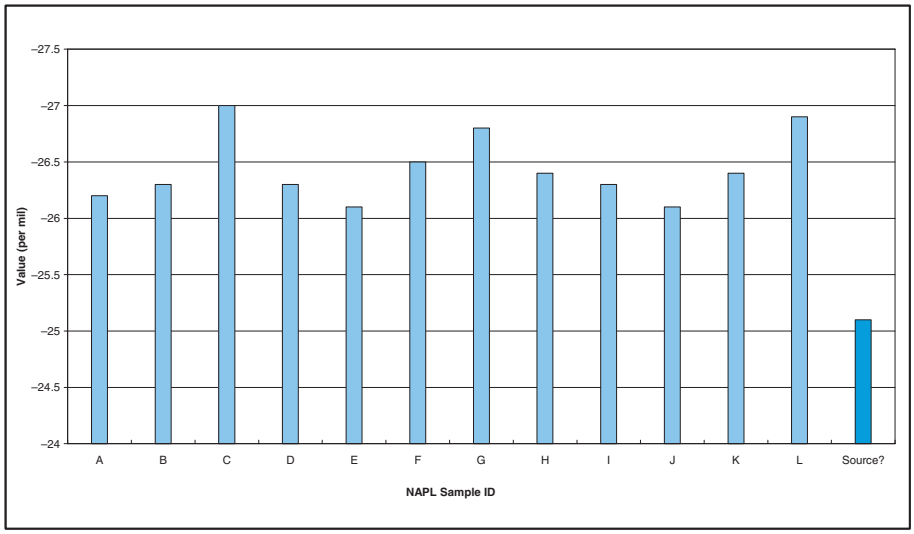


Figure 9.33

Stable isotope analyses of NAPL samples A-L and a potential source sample.

9.5.3.4 Compound-Specific Stable Isotope Measurements

Gas chromatography-isotope ratio-mass spectrometry (GCIRMS) is a specialized application of isotope ratio analysis. The applications of this method in chemical fingerprinting was recently reviewed by Slater (2003), although this review primarily focused on chlorinated ethenes. In GCIRMS, the isotopic ratios of carbon (and theoretically other isotopes) for individual hydrocarbon compounds are measured—thus, this technique often is called compound-specific isotope analysis (CSIA). This method holds particular promise for environmental forensics investigations of gasoline and other petroleum products, because isotopic differences in the chemicals that comprise blending stocks or in additives like MTBE or ethanol can be readily measured and compared among samples containing gasoline residues (Philp, 2002).

In this technique, individual hydrocarbons are first resolved from a complex sample using gas chromatography. The effluent from this separation is introduced into an isotope ratio mass spectrometer where isotope ratios are determined for individual compounds after combustion and catalytic conversion. The benefits of compound-specific isotope measurements are compelling. When coupled with sample preparation methods that isolate the aliphatic, aromatic, polar, and asphaltene constituents of an organic extract, individual hydrocarbon compounds can be isotopically measured (Mansuy *et al.*, 1997). This technique accurately identified the source of petroleum collected from bird feathers based on the carbon isotopic signature of *n*-alkanes produced during pyrolysis of the asphaltene fraction. Other workers have compared uncombusted and combusted plant and fuel materials using GCIRMS and concluded that the 2- and 3-ring PAH exhibited greater source identification potential than heavier PAH, because 5- and 6-ring PAH were more prone to isotopic alteration during pyrolysis (McRae *et al.*, 1997). O'Malley *et al.* (1996) used GCIRMS measurements of 4- and 5-ring PAH to apportion the contribution of fire, car soot, and crankcase oil in coastal sediments. This study was not compromised by pyrolytic bias, because the source material was not secondarily combusted prior to release. Mauro (2000a) and Saber *et al.* (2006) studied the PAH in MGP tars and tar-impacted soils and sediments, respectively, demonstrating the variability in isotopic compositions of individual PAHs. As demonstrated in Emsbo-Mattingly *et al.* (2003c), the primary challenge for these methods is the interfering influence of nontarget analytes that obscure the compound-specific isotopic signature in unpredictable ways as the target analytes approach ambient background concentrations.

The GCIRMS strategy has also been used for volatile hydrocarbons. A classic application of the technique involved the isotopic composition of carbon in benzene, toluene, ethylbenzene, and xylenes (Kelley *et al.*, 1997)

and MTBE (Smallwood *et al.*, 2001, 2002). One difficulty in the application of CSIA to volatile hydrocarbons or MTBE is the strong isotopic fractionation that can occur due to degradation (e.g., Meckenstock *et al.*, 1999). Groundwater samples have been used to identify one source by the depletion of ^{13}C and the other by both an enrichment in ^{13}C and the presence of a dissolved oxygenate. The mechanism of isotopic fractionation of these hydrocarbons was primarily threefold. With respect to the residual material, volatilization resulted in a slight depletion of ^{13}C, aerobic microbial degradation resulted in an enrichment of ^{13}C, and adsorption resulted in no significant change of ^{13}C (Aggarwal and Hinchee, 1991; Harrington *et al.*, 1999; Whittaker *et al.*, 1995).

The hydrogen isotope ratio of $^{2}H/^{1}H$ is interesting, because it exhibits the largest relative mass difference compared to all other elements. Like carbon, the use of GCIRMS has been used to identify the hydrogen isotopic ratio of specific compounds. Different sources of vanillin, ethanol, fatty acid methyl esters, hydrocarbons, and sterols have been identified (Jarman *et al.*, 1998a; Sessions *et al.*, 1999) and demonstrated hydrogen isotope fractionation in 80 individual lipids by algae, bacteria, and higher plants, which revealed significant compound-specific differences among distinct species. To our knowledge, this technique has not yet been used for forensic purposes except in the case of methane (Lundegard, 2006).

9.5.3.5 Two-Dimensional Gas Chromatography

Comprehensive two-dimensional GC (GC × GC) with FID detection provides another dimension to conventional chromatographic separation, greatly improving the resolution that can be attained, particularly for weathered petroleums comprised overwhelmingly of a unresolved complex mixture (Gaines *et al.*, 1999; Frysinger *et al.*, 2002; Nelson *et al.*, 2006). In this method, each analyte in the petroleum mixture is subject to two different separations achieved using two GC columns—typically a nonpolar column followed by a more polar column—connected serially by a thermal modulator, the latter of which is heated and rotated in such a way that it desorbs, spatially compresses, and injects a portion of the first GC column eluent into the second GC column. GC × GC is considered *comprehensive* because all of the analyte mass from the first column is transferred to the second column. Thus, all compounds are subjected to two different separation mechanisms. GC × GC is different than two-dimensional "heart-cut" gas chromatography where one zone of analytes eluting from the first column is isolated and subsequently separated in a different column (Bertsch, 1999).

GC × GC detectors need to have a fast response and flame ionization detectors are most widely used. When a GC × GC is coupled to a mass spectrometer,

an additional dimension of information about the sample is available. However, quadrapole mass spectrometers operating in full scan mode are too slow to properly sample a GC × GC peak unless that peak is broadened (Frysinger and Gaines, 1999). Fast time-of-flight mass spectrometers (TOFMS) that operate with spectral acquisition rates of 100–200 Hz are well-suited for GC × GC and is commercially available. The promising aspect of this technique lies in the potential to fingerprint compounds that could not previously be resolved using GC or GC/MS.

9.5.3.6 Electrospray Ionization Mass Spectrometry (ESI-MS)

In recent years, ESI-MS has been demonstrated to provide a means for the detailed characterization of hydrocarbons and nonhydrocarbons in petroleum and tar. The principles of ESI-MS have been thoroughly described (Rodgers *et al.*, 2000; Kujawinski, 2002). Basically, ESI-MS combines a soft ionization technique (ESI) with MS, which offers ultra-high resolution and mass accuracy. ESI can be performed in either positive or negative modes. Positive ESI yields exact mass information on all compounds present, resulting in extremely complex spectra containing thousands of compounds in crude oil (Qian *et al.*, 2001) and petroleum products (Rostad, 2006). Negative ESI yields less complicated MS since primarily only polar compounds are ionized and detected. Negative ESI has perhaps the greatest potential for forensic purposes since it can provide information on these underutilized (and largely ignored) polar compounds that are difficult to analyze by conventional GC or GC/MS techniques, which depend on polarity and volatility. For example, Qian *et al.* (2001) used ESI-MS to identify the exact masses of approximately 3000 nitrogen-containing aromatic compounds in crude oil. Hughley *et al.* (2001) and Rostad (2006) used negative ESI-MS to demonstrate the variability in polar constituents in various petroleum fuels. As such, ESI-MS can yield extremely specific fingerprints of the polar components in complex mixtures. Comparison of these fingerprints between samples might yield important chemical differences that could go unrecognized by any conventional GC/MS analysis. Unfortunately, the method is poorly suited for the measurement of absolute concentrations and forensic investigations in which there are no known end-member sources available for comparison. Further, the effects of weathering on these polar compounds, and thereby the effects on the polar fingerprints, is an area that requires additional research. These current drawbacks, in combination with lack of commercial availability of this technique, has slowed its entry into typical environmental forensic investigations. Consequently, despite promise, at this point it is most useful in support of the primary quantitative hydrocarbon methods described in this chapter.

9.6 POLYCHLORINATED BIPHENYL FINGERPRINTING METHODS

Polychlorinated biphenyls (PCBs) are considered toxic, persistent, and ubiquitous in the environment (Johnson *et al.*, 2006). In theory, there are 209 possible PCB congeners that contain 1 to 10 chlorine atoms attached to a biphenyl molecule (see Table 9.18).

The number of chlorines and specific molecular configuration govern the common PCB naming systems (see Figure 9.34a). The PCB congeners are often grouped by homolog corresponding to a specific level of chlorination (LOC) (see Table 9.19). For example, PCBs with four chlorines belong to the LOC4 homolog group. PCB congeners can also grouped by toxicity based on the structural similarity to 2,3,7,8-tetrachloro dibenzo-p-dioxin (Ahlborg *et al.*, 1994). The toxicity equivalence factors (TEFs) for PCB congeners are listed in Table 9.18. Herein, the proximity of adjacent chlorine atoms on the biphenyl molecule is also considered important for evaluating the behavior of PCBs in the environment (see Figure 9.34b). The aromatic rings on the biphenyl molecule are free to rotate around the central axis (see Figure 9.34c), whereas steric hindrance of ortho substituted chlorine molecules inhibits ring rotation (see Figure 9.34d). Nonortho coplanar PCBs (see Figure 9.34e) are considered the most toxic PCB congeners because they are not sterically hindered and can form dioxin like planar structures (see Figure 9.34f).

In the United States, PCB congeners (EPA, 2003) are most commonly derived from anthropogenic (man-made) substances known as Aroclors, which were widely used (see Table 9.20) and released into the environment (Houlbeck, 2001). With exceptions, the nomenclature for identifying commercial Aroclors was ARXXYY; where AR indicated Aroclor, XX represented the number of carbon molecules (i.e., 12 indicated a biphenyl structure), and YY indicated the weight percentage of chlorine in the final mixture. The domestic manufacturing of Aroclors occurred from approximately 1929 to 1979 (EPA, 1979). They were generated by reacting biphenyl with chlorine in the presence of a catalyst. The reactions produced complex mixtures of PCB congeners with varied physical and chemical properties controlled primarily by the reaction temperature and duration. Higher temperatures and more lengthy reaction times generated PCB mixtures with higher chlorine content, viscosity, and stability. At room temperature, the more mobile liquid Aroclors included AR1221, AR1232, and AR1242; the viscous liquids included AR1254; the sticky resins included AR1260 and AR1262; and the powders included AR1268 and AR1270 (EPRI, 1999 and references therein; Johnson *et al.*, 2006).

The chlorine content of the 10 most common Aroclors ranged from 20% to 70%. Middle range Aroclor liquids contained numerous PCB congeners;

Table 9.18

PCB congener analytes.

PCB Congener Analyte	CAS No.	IUPAC No.	LOC[1]	Quant Ion (m/z)	TEF[2]	Approximate Detection Limit	
						Water	Solid
2-Chlorobiphenyl	2051-60-7	PCB 1	1	188	—	1.0 ng/L	1.0 μg/Kg
3-Chlorobiphenyl	2051-61-8	PCB 2	1	188	—	1.0 ng/L	1.0 μg/Kg
4-Chlorobiphenyl	2051-62-9	PCB 3	1	188	—	1.0 ng/L	1.0 μg/Kg
2,2′-Dichlorobiphenyl	13029-08-8	PCB 4	2	222	—	1.0 ng/L	1.0 μg/Kg
2,3-Dichlorobiphenyl	16605-91-7	PCB 5	2	222	—	1.0 ng/L	1.0 μg/Kg
2,3′-Dichlorobiphenyl	25569-80-6	PCB 6	2	222	—	1.0 ng/L	1.0 μg/Kg
2,4-Dichlorobiphenyl	33284-50-3	PCB 7	2	222	—	1.0 ng/L	1.0 μg/Kg
2,4′-Dichlorobiphenyl	34883-43-7	PCB 8	2	222	—	1.0 ng/L	1.0 μg/Kg
2,5-Dichlorobiphenyl	34883-39-1	PCB 9	2	222	—	1.0 ng/L	1.0 μg/Kg
2,6-Dichlorobiphenyl	33416-45-1	PCB 10	2	222	—	1.0 ng/L	1.0 μg/Kg
3,3′-Dichlorobiphenyl	2050-67-1	PCB 11	2	222	—	1.0 ng/L	1.0 μg/Kg
3,4-Dichlorobiphenyl	2974-92-7	PCB 12	2	222	—	1.0 ng/L	1.0 μg/Kg
3,4′-Dichlorobiphenyl	2974-90-5	PCB 13	2	222	—	1.0 ng/L	1.0 μg/Kg
3,5-Dichlorobiphenyl	34883-41-5	PCB 14	2	222	—	1.0 ng/L	1.0 μg/Kg
4,4′-Dichlorobiphenyl	2050-68-2	PCB 15	2	222	—	1.0 ng/L	1.0 μg/Kg
2,2′,3-Trichlorobiphenyl	38444-78-9	PCB 16	3	256	—	1.0 ng/L	1.0 μg/Kg
2,2′,4-Trichlorobiphenyl	37680-66-3	PCB 17	3	256	—	1.0 ng/L	1.0 μg/Kg
2,2′,5-Trichlorobiphenyl	37680-65-2	PCB 18	3	256	—	1.0 ng/L	1.0 μg/Kg
2,2′,6-Trichlorobiphenyl	38444-73-4	PCB 19	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3,3′-Trichlorobiphenyl	38444-84-7	PCB 20	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3,4-Trichlorobiphenyl	55702-46-0	PCB 21	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3,4′-Trichlorobiphenyl	38444-85-8	PCB 22	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3,5-Trichlorobiphenyl	55720-44-0	PCB 23	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3,6-Trichlorobiphenyl	55702-45-9	PCB 24	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3′,4-Trichlorobiphenyl	55712-37-3	PCB 25	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3′,5-Trichlorobiphenyl	38444-81-4	PCB 26	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3′,6-Trichlorobiphenyl	38444-76-7	PCB 27	3	256	—	1.0 ng/L	1.0 μg/Kg
2,4,4′-Trichlorobiphenyl	7012-37-5	PCB 28	3	256	—	1.0 ng/L	1.0 μg/Kg
2,4,5-Trichlorobiphenyl	15862-07-4	PCB 29	3	256	—	1.0 ng/L	1.0 μg/Kg
2,4,6-Trichlorobiphenyl	35693-92-6	PCB 30	3	256	—	1.0 ng/L	1.0 μg/Kg
2,4′,5-Trichlorobiphenyl	16606-02-3	PCB 31	3	256	—	1.0 ng/L	1.0 μg/Kg
2,4′,6-Trichlorobiphenyl	38444-77-8	PCB 32	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3′,4′-Trichlorobiphenyl	38444-86-9	PCB 33	3	256	—	1.0 ng/L	1.0 μg/Kg
2,3′,5′-Trichlorobiphenyl	37680-68-5	PCB 34	3	256	—	1.0 ng/L	1.0 μg/Kg
3,3′,4-Trichlorobiphenyl	37680-69-6	PCB 35	3	256	—	1.0 ng/L	1.0 μg/Kg
3,3′,5-Trichlorobiphenyl	38444-87-0	PCB 36	3	256	—	1.0 ng/L	1.0 μg/Kg
3,4,4′-Trichlorobiphenyl	38444-90-5	PCB 37	3	256	—	1.0 ng/L	1.0 μg/Kg
3,4,5-Trichlorobiphenyl	53555-66-1	PCB 38	3	256	—	1.0 ng/L	1.0 μg/Kg
3,4′,5-Trichlorobiphenyl	38444-88-1	PCB 39	3	256	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′-Tetrachlorobiphenyl	38444-93-8	PCB 40	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4-Tetrachlorobiphenyl	52663-59-9	PCB 41	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′-Tetrachlorobiphenyl	36559-22-5	PCB 42	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5-Tetrachlorobiphenyl	70362-46-8	PCB 43	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5′-Tetrachlorobiphenyl	41464-39-5	PCB 44	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,6-Tetrachlorobiphenyl	70362-45-7	PCB 45	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,6′-Tetrachlorobiphenyl	41464-47-5	PCB 46	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′-Tetrachlorobiphenyl	2437-79-8	PCB 47	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,5-Tetrachlorobiphenyl	70362-47-9	PCB 48	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,5′-Tetrachlorobiphenyl	41464-40-8	PCB 49	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,6-Tetrachlorobiphenyl	62796-65-0	PCB 50	4	292	—	1.0 ng/L	1.0 μg/Kg

Table 9.18
Continued

PCB Congener Analyte	CAS No.	IUPAC No.	LOC[1]	Quant Ion (m/z)	TEF[2]	Approximate Detection Limit	
						Water	Solid
2,2′,4,6′-Tetrachlorobiphenyl	68194-04-7	PCB 51	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,5,5′-Tetrachlorobiphenyl	35693-99-3	PCB 52	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,5,6′-Tetrachlorobiphenyl	41464-41-9	PCB 53	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,6,6′-Tetrachlorobiphenyl	15968-05-5	PCB 54	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4-Tetrachlorobiphenyl	74338-24-2	PCB 55	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′-Tetrachlorobiphenyl	41464-43-1	PCB 56	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5-Tetrachlorobiphenyl	70424-67-8	PCB 57	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5′-Tetrachlorobiphenyl	41464-49-7	PCB 58	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,6-Tetrachlorobiphenyl	74472-33-6	PCB 59	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,4,4′-Tetrachlorobiphenyl	33025-41-1	PCB 60	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,4,5-Tetrachlorobiphenyl	33284-53-6	PCB 61	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,4,6-Tetrachlorobiphenyl	54230-22-7	PCB 62	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,4′,5-Tetrachlorobiphenyl	74472-34-7	PCB 63	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,4′,6-Tetrachlorobiphenyl	52663-58-8	PCB 64	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3,5,6-Tetrachlorobiphenyl	33284-54-7	PCB 65	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′-Tetrachlorobiphenyl	32598-10-0	PCB 66	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,5-Tetrachlorobiphenyl	73575-53-8	PCB 67	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,5′-Tetrachlorobiphenyl	73575-52-7	PCB 68	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,6-Tetrachlorobiphenyl	60233-24-1	PCB 69	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4′,5-Tetrachlorobiphenyl	32598-11-1	PCB 70	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4′,6-Tetrachlorobiphenyl	41464-46-4	PCB 71	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,5,5′-Tetrachlorobiphenyl	41464-42-0	PCB 72	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,5′,6-Tetrachlorobiphenyl	74338-23-1	PCB 73	4	292	—	1.0 ng/L	1.0 μg/Kg
2,4,4′,5-Tetrachlorobiphenyl	32690-93-0	PCB 74	4	292	—	1.0 ng/L	1.0 μg/Kg
2,4,4′,6-Tetrachlorobiphenyl	32598-12-2	PCB 75	4	292	—	1.0 ng/L	1.0 μg/Kg
2,3′,4′,5′-Tetrachlorobiphenyl	70362-48-0	PCB 76	4	292	—	1.0 ng/L	1.0 μg/Kg
3,3′,4,4′-Tetrachlorobiphenyl	32598-13-3	PCB 77	4	292	0.0005	1.0 ng/L	1.0 μg/Kg
3,3′,4,5-Tetrachlorobiphenyl	70362-49-1	PCB 78	4	292	—	1.0 ng/L	1.0 μg/Kg
3,3′,4,5′-Tetrachlorobiphenyl	41464-48-6	PCB 79	4	292	—	1.0 ng/L	1.0 μg/Kg
3,3′,5,5′-Tetrachlorobiphenyl	33284-52-5	PCB 80	4	292	—	1.0 ng/L	1.0 μg/Kg
3,4,4′,5-Tetrachlorobiphenyl	70362-50-4	PCB 81	4	292	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4-Pentachlorobiphenyl	52663-62-4	PCB 82	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5-Pentachlorobiphenyl	60145-20-2	PCB 83	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,6-Pentachlorobiphenyl	52663-60-2	PCB 84	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′-Pentachlorobiphenyl	65510-45-4	PCB 85	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5-Pentachlorobiphenyl	55312-69-1	PCB 86	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5′-Pentachlorobiphenyl	38380-02-8	PCB 87	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,6-Pentachlorobiphenyl	55215-17-3	PCB 88	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,6′-Pentachlorobiphenyl	73575-57-2	PCB 89	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5-Pentachlorobiphenyl	68194-07-0	PCB 90	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,6-Pentachlorobiphenyl	68194-05-8	PCB 91	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5,5′-Pentachlorobiphenyl	52663-61-3	PCB 92	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5,6-Pentachlorobiphenyl	73575-56-1	PCB 93	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5,6′-Pentachlorobiphenyl	73575-55-0	PCB 94	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5′,6-Pentachlorobiphenyl	38379-99-6	PCB 95	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,6,6′-Pentachlorobiphenyl	73575-54-9	PCB 96	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5′-Pentachlorobiphenyl	41464-51-1	PCB 97	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,6′-Pentachlorobiphenyl	60223-25-2	PCB 98	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′,5-Pentachlorobiphenyl	38380-01-7	PCB 99	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′,6-Pentachlorobiphenyl	39485-83-1	PCB 100	5	326	—	1.0 ng/L	1.0 μg/Kg

Table 9.18
Continued

PCB Congener Analyte	CAS No.	IUPAC No.	LOC[1]	Quant Ion (m/z)	TEF[2]	Approximate Detection Limit	
						Water	Solid
2,2′,4,5,5′-Pentachlorobiphenyl	37680-73-2	PCB 101	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,5,6′-Pentachlorobiphenyl	68194-06-9	PCB 102	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,5′,6-Pentachlorobiphenyl	60145-21-3	PCB 103	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,6,6′-Pentachlorobiphenyl	56558-16-8	PCB 104	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′-Pentachlorobiphenyl	32598-14-4	PCB 105	5	326	0.0001	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5-Pentachlorobiphenyl	70424-69-0	PCB 106	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5-Pentachlorobiphenyl	70424-68-9	PCB 107	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5′-Pentachlorobiphenyl	70362-41-3	PCB 108	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,6-Pentachlorobiphenyl	74472-35-8	PCB 109	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,6-Pentachlorobiphenyl	38380-03-9	PCB 110	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5,5′-Pentachlorobiphenyl	39635-32-0	PCB 111	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5,6-Pentachlorobiphenyl	74472-36-9	PCB 112	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5′,6-Pentachlorobiphenyl	68194-10-5	PCB 113	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,4,4′,5-Pentachlorobiphenyl	74472-37-0	PCB 114	5	326	0.0005	1.0 ng/L	1.0 μg/Kg
2,3,4,4′,6-Pentachlorobiphenyl	74472-38-1	PCB 115	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,4,5,6-Pentachlorobiphenyl	18259-05-7	PCB 116	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,4′,5,6-Pentachlorobiphenyl	68194-11-6	PCB 117	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′,5-Pentachlorobiphenyl	31508-00-6	PCB 118	5	326	0.0001	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′,6-Pentachlorobiphenyl	56558-17-9	PCB 119	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,5,5′-Pentachlorobiphenyl	68194-12-7	PCB 120	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,5′,6-Pentachlorobiphenyl	56558-18-0	PCB 121	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5′-Pentachlorobiphenyl	76842-07-4	PCB 122	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′,5′-Pentachlorobiphenyl	65510-44-3	PCB 123	5	326	0.0001	1.0 ng/L	1.0 μg/Kg
2,3′,4′,5,5′-Pentachlorobiphenyl	70424-70-3	PCB 124	5	326	—	1.0 ng/L	1.0 μg/Kg
2,3′,4′,5′,6-Pentachlorobiphenyl	74472-39-2	PCB 125	5	326	—	1.0 ng/L	1.0 μg/Kg
3,3′,4,4′,5-Pentachlorobiphenyl	57465-28-8	PCB 126	5	326	0.1	1.0 ng/L	1.0 μg/Kg
3,3′,4,5,5′-Pentachlorobiphenyl	39635-33-1	PCB 127	5	326	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′-Hexachlorobiphenyl	38380-07-3	PCB 128	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5-Hexachlorobiphenyl	55215-18-4	PCB 129	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5′-Hexachlorobiphenyl	52663-66-8	PCB 130	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,6-Hexachlorobiphenyl	61798-70-7	PCB 131	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,6′-Hexachlorobiphenyl	38380-05-1	PCB 132	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,5′-Hexachlorobiphenyl	35694-04-3	PCB 133	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,6-Hexachlorobiphenyl	52704-70-8	PCB 134	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,6′-Hexachlorobiphenyl	52744-13-5	PCB 135	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,6,6′-Hexachlorobiphenyl	38411-22-2	PCB 136	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5-Hexachlorobiphenyl	35694-06-5	PCB 137	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5′-Hexachlorobiphenyl	35065-28-2	PCB 138	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,6-Hexachlorobiphenyl	56030-56-9	PCB 139	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,6′-Hexachlorobiphenyl	59291-64-4	PCB 140	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5,5′-Hexachlorobiphenyl	52712-04-6	PCB 141	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5,6-Hexachlorobiphenyl	41411-61-4	PCB 142	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5,6′-Hexachlorobiphenyl	68194-15-0	PCB 143	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5′,6-Hexachlorobiphenyl	68194-14-9	PCB 144	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,6,6′-Hexachlorobiphenyl	74472-40-5	PCB 145	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5,5′-Hexachlorobiphenyl	51908-16-8	PCB 146	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5,6-Hexachlorobiphenyl	68194-13-8	PCB 147	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5,6′-Hexachlorobiphenyl	74472-41-6	PCB 148	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5′,6-Hexachlorobiphenyl	38380-04-0	PCB 149	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,6,6′-Hexachlorobiphenyl	68194-08-1	PCB 150	6	360	—	1.0 ng/L	1.0 μg/Kg

Table 9.18
Continued

PCB Congener Analyte	CAS No.	IUPAC No.	LOC[1]	Quant Ion (m/z)	TEF[2]	Approximate Detection Limit	
						Water	Solid
2,2′,3,5,5′,6-Hexachlorobiphenyl	52663-63-5	PCB 151	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,5,6,6′-Hexachlorobiphenyl	68194-09-2	PCB 152	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′,5,5′-Hexachlorobiphenyl	35065-27-1	PCB 153	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′,5,6′-Hexachlorobiphenyl	60145-22-4	PCB 154	6	360	—	1.0 ng/L	1.0 μg/Kg
2,2′,4,4′,6,6′-Hexachlorobiphenyl	33979-03-2	PCB 155	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5-Hexachlorobiphenyl	38380-08-4	PCB 156	6	360	0.0005	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5′-Hexachlorobiphenyl	69782-90-7	PCB 157	6	360	0.0005	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,6-Hexachlorobiphenyl	74472-42-7	PCB 158	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5,5′-Hexachlorobiphenyl	39635-35-3	PCB 159	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5,6-Hexachlorobiphenyl	41411-62-5	PCB 160	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5′,6-Hexachlorobiphenyl	74472-43-8	PCB 161	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5,5′-Hexachlorobiphenyl	39635-34-2	PCB 162	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5,6-Hexachlorobiphenyl	74472-44-9	PCB 163	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5′,6-Hexachlorobiphenyl	74472-45-0	PCB 164	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,5,5′,6-Hexachlorobiphenyl	74472-46-1	PCB 165	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3,4,4′,5,6-Hexachlorobiphenyl	41411-63-6	PCB 166	6	360	—	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′,5,5′-Hexachlorobiphenyl	52663-72-6	PCB 167	6	360	0.00001	1.0 ng/L	1.0 μg/Kg
2,3′,4,4′,5′,6-Hexachlorobiphenyl	59291-65-5	PCB 168	6	360	—	1.0 ng/L	1.0 μg/Kg
3,3′,4,4′,5,5′-Hexachlorobiphenyl	32774-16-6	PCB 169	6	360	0.01	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5-Heptachlorobiphenyl	35065-30-6	PCB 170	7	394	0.0001	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,6-Heptachlorobiphenyl	52663-71-5	PCB 171	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,5′-Heptachlorobiphenyl	52663-74-8	PCB 172	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,6-Heptachlorobiphenyl	68194-16-1	PCB 173	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,6′-Heptachlorobiphenyl	38411-25-5	PCB 174	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5′,6-Heptachlorobiphenyl	40186-70-7	PCB 175	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,6,6′-Heptachlorobiphenyl	52663-65-7	PCB 176	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5′,6′-Heptachlorobiphenyl	52663-70-4	PCB 177	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,5′,6-Heptachlorobiphenyl	52663-67-9	PCB 178	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,6,6′-Heptachlorobiphenyl	52663-64-6	PCB 179	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5,5′-Heptachlorobiphenyl	35065-29-3	PCB 180	7	394	0.00001	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5,6-Heptachlorobiphenyl	74472-47-2	PCB 181	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5,6′-Heptachlorobiphenyl	60145-23-5	PCB 182	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5′,6-Heptachlorobiphenyl	52663-69-1	PCB 183	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,6,6′-Heptachlorobiphenyl	74472-48-3	PCB 184	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5,5′,6-Heptachlorobiphenyl	52712-05-7	PCB 185	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,5,6,6′-Heptachlorobiphenyl	74472-49-4	PCB 186	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5,5′,6-Heptachlorobiphenyl	52663-68-0	PCB 187	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4′,5,6,6′-Heptachlorobiphenyl	74487-85-7	PCB 188	7	394	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5,5′-Heptachlorobiphenyl	39635-31-9	PCB 189	7	394	0.0001	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5,6-Heptachlorobiphenyl	41411-64-7	PCB 190	7	394	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5′,6-Heptachlorobiphenyl	74472-50-7	PCB 191	7	394	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,5,5′,6-Heptachlorobiphenyl	74472-51-8	PCB 192	7	394	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4′,5,5′,6-Heptachlorobiphenyl	69782-91-8	PCB 193	7	394	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,5′-Octachlorobiphenyl	35694-08-7	PCB 194	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,6-Octachlorobiphenyl	52663-78-2	PCB 195	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,6′-Octachlorobiphenyl	42740-50-1	PCB 196	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,6,6′-Octachlorobiphenyl	33091-17-7	PCB 197	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,5′,6-Octachlorobiphenyl	68194-17-2	PCB 198	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,5′,6′-Octachlorobiphenyl	52663-75-9	PCB 199	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,6,6′-Octachlorobiphenyl	52663-73-7	PCB 200	8	428	—	1.0 ng/L	1.0 μg/Kg

Table 9.18
Continued

PCB Congener Analyte	CAS No.	IUPAC No.	LOC[1]	Quant Ion (m/z)	TEF[2]	Approximate Detection Limit	
						Water	Solid
2,2′,3,3′,4,5′,6,6′-Octachlorobiphenyl	40186-71-8	PCB 201	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,5,5′,6,6′-Octachlorobiphenyl	2136-99-4	PCB 202	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5,5′,6-Octachlorobiphenyl	52663-76-0	PCB 203	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,4,4′,5,6,6′-Octachlorobiphenyl	74472-52-9	PCB 204	8	428	—	1.0 ng/L	1.0 μg/Kg
2,3,3′,4,4′,5,5′,6-Octachlorobiphenyl	74472-53-0	PCB 205	8	428	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,5′,6-Nonachlorobiphenyl	40186-72-9	PCB 206	9	464	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,6,6′-Nonachlorobiphenyl	52663-79-3	PCB 207	9	464	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,5,5′,6,6′-Nonachlorobiphenyl	52663-77-1	PCB 208	9	464	—	1.0 ng/L	1.0 μg/Kg
2,2′,3,3′,4,4′,5,5′,6,6′-Decachlorobiphenyl	2051-24-3	PCB 209	10	498	—	1.0 ng/L	1.0 μg/Kg

Notes:
[1]Level of chlorination (LOC) is synonymous with homolog group.
[2]Toxic equivalency factors (TEFs) were adopted by WHO (1994) and EPA (1996) based on Ahlborg *et al.* (1994). They represent scaling factors for relating PCB concentrations to risk relative to the most toxic polychlorinated dibenzo-p-dioxin (PCDD) known as 2,3,7,8-tetrachloro dibenzo-p-dioxin (2378-TCDD). PBC congeners with higher TEFs are considered more toxic than PCB congeners with lower or no TEFs.

Table 9.19
PCB homolog analytes.

PCB Homolog	Chemical Formula	CAS No.	LOC	No. Chlorines	No. Congeners
Monochlorobiphenyls	$C_{12}H_9Cl$	27323-18-8	1	1	3
Dichlorobiphenyls	$C_{12}H_8Cl_2$	25512-42-9	2	2	12
Trichlorobiphenyls	$C_{12}H_7Cl_3$	25323-68-6	3	3	24
Tetrachlorobiphenyls	$C_{12}H_6Cl_4$	26914-33-0	4	4	42
Pentachlorobiphenyls	$C_{12}H_5Cl_5$	25429-29-2	5	5	46
Hexachlorobiphenyls	$C_{12}H_4Cl_6$	26601-64-9	6	6	42
Heptachlorobiphenyls	$C_{12}H_3Cl_7$	28655-71-2	7	7	24
Octachlorobiphenyls	$C_{12}H_2Cl_8$	55722-26-4	8	8	12
Nonachlorobiphenyls	$C_{12}H_1Cl_9$	53742-07-7	9	9	3
Decachlorobiphenyls	$C_{12}Cl_{10}$	2051-24-3	10	10	1
Total PCB		1336-36-3			

for example, AR1242 principally contained congeners with 2, 3, 4, and 5 chlorines, and AR1254 principally contained congeners with 4, 5, 6, and 7 chlorines. AR1248 contained intermediate levels of PCB congeners with 3 through 7 chlorines. Using specialized PCB analytical methods, researchers have demonstrated the congener specific differences between each of these Aroclor products (see Figure 9.35) (Rushneck, 2004; Frame *et al.*, 1996a, 1996b). Although these higher resolution methods are most useful to the forensic investigator, a familiarity with the general patterns assist a review of more routine PCB data generated by lower resolution techniques.

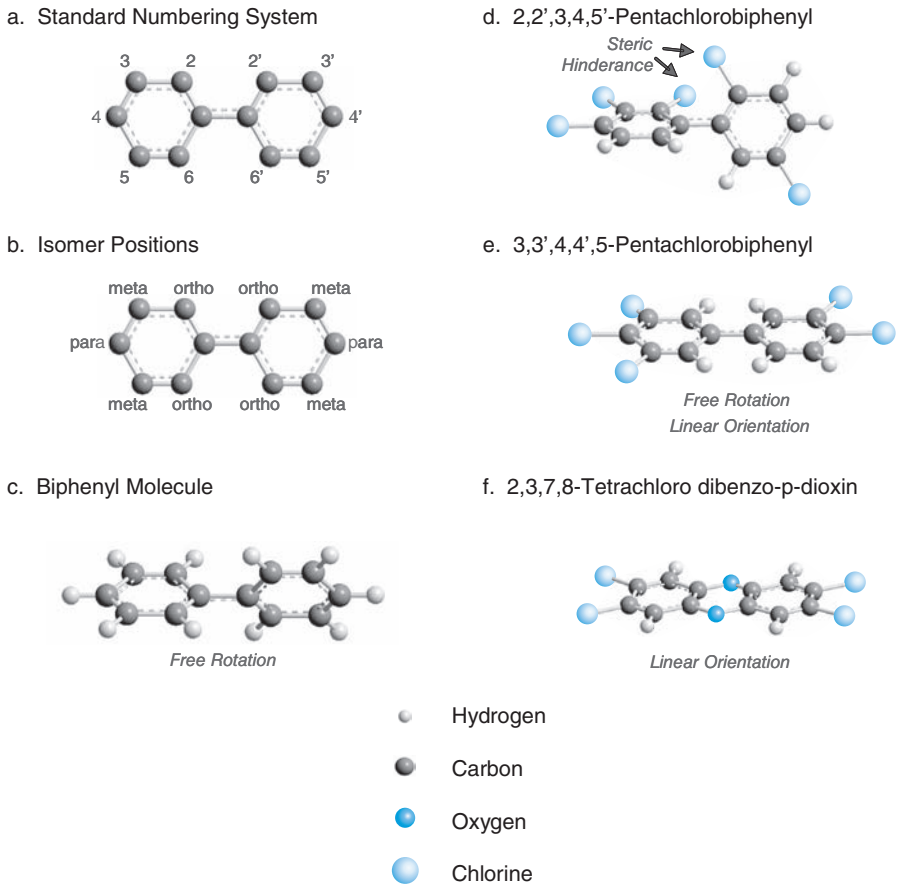


Figure 9.34
PCB Nomenclature.

9.6.1 CURRENT PRACTICES

The analytical methods used for the identification and detection of PCBs in environmental samples changed dramatically over the past several decades as GC instrumentation evolved and improved. An overview of PCB methodology can be found in Erickson (1997). Many current PCB environmental investigations rely on both historical and modern analytical data. Therefore, it is important for the forensic chemist to understand the evolution of these analytical methods, and their applications/limitations with respect to product and source identification. At present, the measurement of PCBs can be performed by several quantitative methods that permit a wide range of procedural and instrumental options. Among the many methods developed for the measurement of PCBs, the most commonly employed techniques include EPA Method 8082 (GC/ECD), EPA Method 8270C/680 (GC/MS), and EPA Method 1668A (GC/HRMS).

Table 9.20
Applications of PCB aroclors.

Application	PCB Aroclor								
	1016	1221	1232	1242	1248	1254	1260	1262	1268
Closed Systems									
Transformers				X		X	X[1]		
Capacitors	X	minor		X[1]		X			
Nominally Closed Systems									
Small capacitors	X	minor		X[1]		X			
Heat transfer systems				X					
Hydraulic equipment			X	X	X	X	X		
Vacuum pumps					X	X			
Gas-transmission turbines		X		X					
Open Systems									
Plasticizers		X	X	X	X	X	X	X	X
Carbonless carbon paper				X					
Paint additives						X	X		
Rubber plasticizer		X	X	X	X	X			X
Synthetic resin/plasticizer					X	X	X	X	X
Adhesives		X	X	X	X	X			
Wax extenders				X		X			X
Dedusting agents						X	X		
Links						X	X		
Cutting oils						X			
Pesticide extenders						X			
Sealants and caulking					X				
Lubricants		X		X	X	X			

[1]The use of AR1242 in capacitors and AR 1260 in transformers was discontinued as of 1971 in the United States.

Unlike standard EPA methods for hydrocarbons, the various PCB methods in use today include a comprehensive list of analytes and robust strategies for detecting trace concentrations in samples with complex matrices. PCB chemists have demonstrated that GC/ECD (Frame *et al.*, 1996a), GC/MS (Frame *et al.*, 1999; Emsbo-Mattingly and Durell, 2006), and GC/HRMS (EPA, 1999b; Rushneck *et al.*, 2004) methods can resolve the majority of ($n > 125$) of the PCB congeners needed to identify the original Aroclor material and changes due to volatilization or anaerobic dechlorination. Under many situations, the EPA standard methods produce data of sufficient quality to identify source signatures of PCBs.

The primary problem often faced by forensic investigators lies in the laboratory-specific execution of the procedures. PCB fingerprinting difficulties arise when comparing data generated by laboratories that use different operating conditions to analyze PCBs, for example, different capillary columns and instrument operating conditions that cause different groups of PCB congeners to coelute and be potentially misreported or not reported at all. Environmental laboratories promote efficiency and higher throughput by analyzing

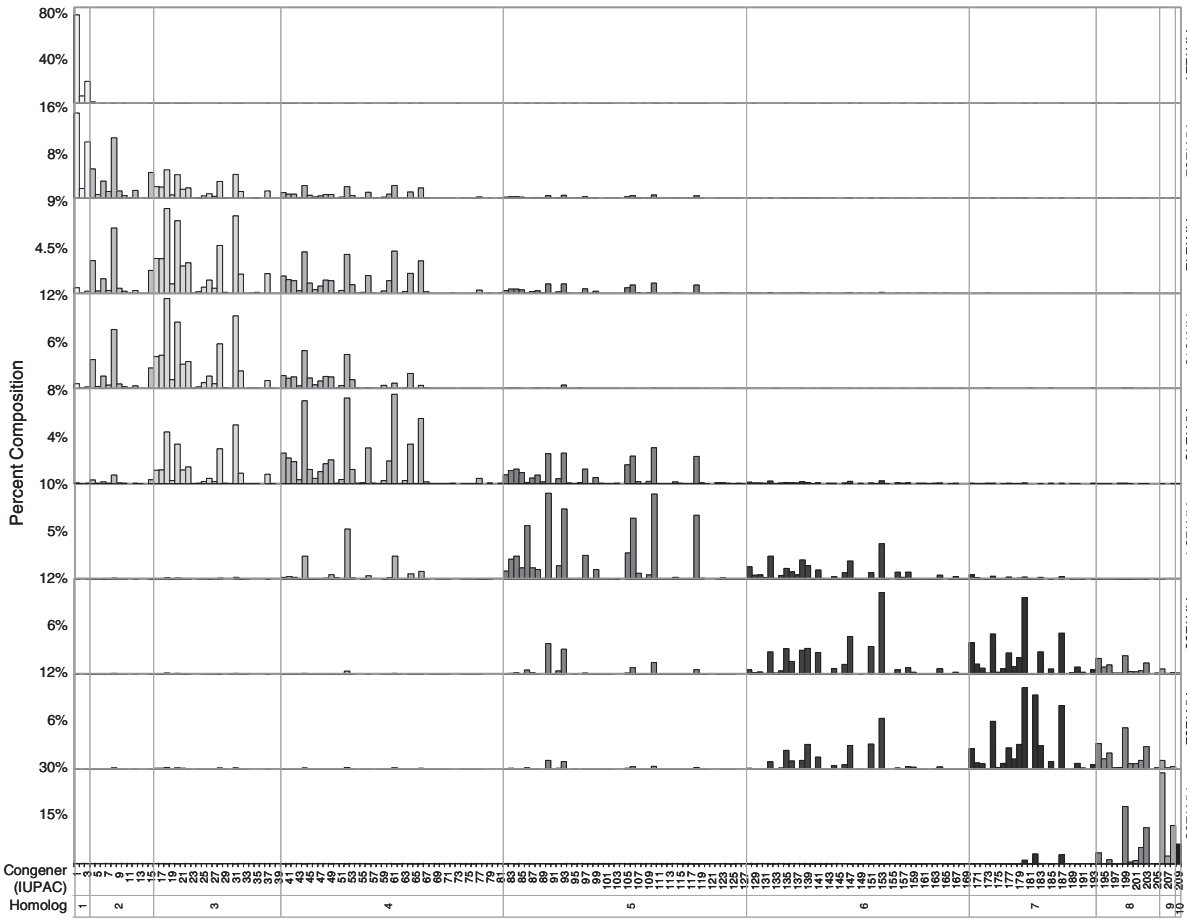


Figure 9.35

Aroclor PCB congener patterns fully resolved by GC/HRMS (Rushneck et al. *2004).*

samples as fast as possible, while still allowing the instrument analyst to qualitatively differentiate common Aroclor calibration standards (e.g., AR1221, AR1232, AR1016, AR1242, AR1248, AR1254, and AR1260). Although this approach improves the speed of analysis, it also forces many PCB congeners to coelute (e.g., merge together as one peak). PCB congener coelution is particularly problematic when a laboratory uses GC/ECD without second column or separate GC/MS confirmation. Consequently, the rapid Aroclor analysis used by most contract laboratories often sacrifices the ability to accurately observe subtle compositional differences that are essential to the forensic investigator.

A secondary problem for the forensic investigator is the use of different types of detection systems for PCB analysis that introduce varying degrees of matrix interference associated with laboratory-specific bias in the measurement of selected congeners. The PCB methods can generate different levels

of sensitivity among data populations, further confounding the ability to systematically identify and quantify the PCBs in complex environmental samples. For example, the electrolytic conductivity detector (ELCD) exhibits a first order linear increase in detector response proportional to the number of chorines on the PCB congener (a valued feature); however, this detector requires constant care and maintenance that renders it impractical for routine analyses in most laboratories. The net effect of these method compliant differences is the impaired ability to compare forensic signatures effectively within or among projects.

These limitations are more pronounced when complex mixtures of Aroclors commingle in the environment. In addition, the identification of PCB sources is most important in complex matrices (sediment and tissue) and at low detection limits associated with environmental risk. As with most forensic analyses, weathering (e.g., dechlorination, volatilization; Johnson *et al.*, 2006) of the PCB distributions will alter the original source signature, making it more difficult to identify in complex environmental samples. A constructive strategy for managing these more difficult situations is to recognize the potential need for source identification in the long term and plan for it in the short term. Examples of short-term planning strategies include the use of analytical methods with a demonstrated capability for both regulatory compliance and source identification. The following discussion provides specific guidelines that will allow the environmental scientist to optimize the forensic value of currently available PCB analytical methods.

9.6.2 PCB FORENSIC METHODS

As is the case with hydrocarbons (described previously), PCB methods that generate data with poor congener resolution and high method detection limits are of limited value for identifying the type and source of the environmental contamination. The pressure on production laboratories to lower costs and increase throughput in some cases has resulted in a lower quality analytical product that generates a highly compressed chemical signature replete with nondetected values (even though these compounds would be detected at a lower detection limit). By contrast, some laboratories optimize the resolution of the method (GC/ECD, GC/MS/SIM, and GC/HRMS) for improved sensitivity and fingerprinting quality.

The principal strategy for capturing the essential aspects of a PCB forensic signature employs many of the same procedures that are used in the sample preparation, cleanup, and analysis of hydrocarbons. Analytically, such an approach includes both a robust primary (Tier 1) chromatography analysis and, when needed, an optional confirmatory method (Tier 2). This approach

emphasizes the natural evolution of GC/ECD and GC/MS methods toward the standardized measurement of all 209 PCB congeners at gradually declining detection limits and minimal numbers of coeluting analytes. Tier 2 is recommended when the Tier 1 results require confirmation. Tier 2 is the EPA1668A method as written with an analyte list inclusive of all 209 PCB congeners. This latter, state-of the-art method attains the highest quality data available and provides very low detection limits in virtually all environmental matrices.

9.6.2.1 Sample Preparation

The sample preparation procedures described for hydrocarbons previously (see Section 9.5.1) work well for PCBs as well. Unlike the hydrocarbon method, however, the sample preparation procedure requires different PCB-related surrogate and analyte spike compounds. These QC compounds should be selected to represent the light and heavy spectrum of PCB congeners. The use of less commonly encountered PCB congeners as surrogates should be avoided when possible because all 209 PCB congeners are potentially helpful for the identification of less common PCB materials and weathering by-products. Surrogate compounds such as tetrachloro-m-xylene (TCX), 4,4′-dibromooctafluorobiphenyl (DBOB), and carbon labeled PCB congeners (MS only) are preferred surrogate compounds and have been shown to effectively monitor sample extraction efficiencies (NOAA, 1998).

9.6.2.2 Extract Cleanup

The sample extract cleanup procedure depends on the complexity of the matrix and requires consideration on a project-specific basis. Projects that require very low detection limits for the most toxic PCB congeners may require more aggressive cleanup and detection strategies than those presented here. However, it must be recognized that extensive cleanup and subsequent concentration frequently result in the progressive loss of lighter PCB congeners. This loss may not affect the determination of risk, but it can obscure the signature of AR1221, AR1232, and the by-products of anaerobic dechlorination (see Figure 9.36). In order to preserve the chemical signature of PCBs in environmental samples, we recommend using only the minimum of cleanup steps when preparing such samples for forensic analysis.

9.6.2.2.1 Alumina Cleanup (EPA Method 3611B)

The alumina cleanup step is recommended for all sediment and tissue samples to remove polar and other interfering organic compounds from the extracted sample. It can be used for water and soil if high levels of polar materials are present. Alumina is a porous and granular form of aluminum oxide that is

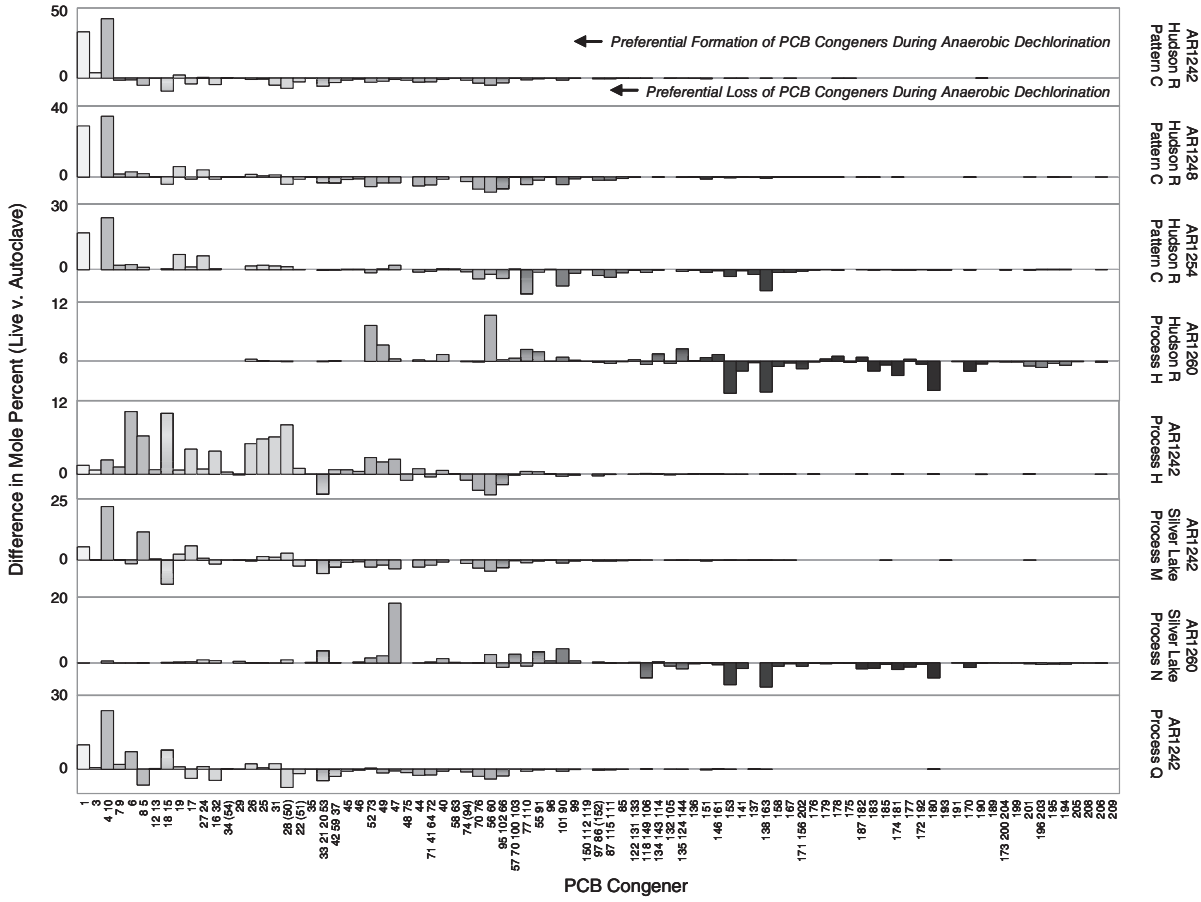


Figure 9.36

PCB congener dechlorination patterns (Johnson et al. *2006 and references therein).*

prepared at neutral pH and used in its dry-activated state. A chromatography column fitted with a polytetrafluoroethylene (PTFE) stopcock is prepared with a glass wool plug below a 10-g bed of alumina and topped with 1 cm plug of anhydrous sodium sulfate powder. The column is rinsed with DCM, loaded with the extract, and eluted with 30 mL of DCM. The extract is concentrated by N-Evap.

9.6.2.2.2 Gel-Permeation Cleanup (EPA Method 3640A)

The gel-permeation chromatography (GPC) cleanup step is recommended for tissue and heavily contaminated sediment samples (NOAA, 1998). GPC is a size exclusion technique that separates macromolecules (e.g., biogenic compounds like lipids, polymers, proteins, resins, and cellulose) from the target analytes. A SEPHADEX system is calibrated using a liquid chromatography

system equipped with an ultraviolet detector. The extract is eluted off the column after the retention time of corn oil and phthalate calibration standards. The elutriate is concentrated by KD and N-Evap.

9.6.2.2.3 Sulfur Removal with Copper Granules (EPA Method 3660B)

The copper granule cleanup method is recommended for the removal of excess sulfur in sediment sample extracts. The copper granules are prepared by acid rinse, water wash, solvent extraction, and N-Evap drying. The sample extract is added to a vial containing approximately 2 g of copper granules and shaken or vortexed for 15 minutes. The extract is transferred to a clean vial by pipette.

9.6.2.2.4 Sulfuric Acid (EPA Method 3665A)

The sulfuric acid cleanup removes complex hydrocarbons and reactive pesticides that cause baseline rise and excessively complex chromatograms, respectively. The sample extract is exchanged to hexane and shaken vigorously or vortexed with a 1 : 1 mixture of sulfuric acid and reagent water for one minute. Transfer the hexane extract on top to a clean vial. Add 1 mL of clean hexane and repeat the mixing procedure to assure a quantitative transfer. Concentrate the extract by N-Evap. The procedure is repeated if the extract retains a brown color after it is allowed to settle.

9.6.2.3 Internal Standards (EPA Method 8000B)

The use of the internal standard method (EPA 1966) for quantifying PCB congeners is strongly recommended for the GC/MS methods. This approach generally produces a more accurate quantitative analysis because it incorporates changes in extract volume after the sample preparation is complete (USACE, 2005). Like the surrogate compounds, the internal standards are most often brominated, fluorinated, or stable isotopically labeled analogs of specific target compounds, or are closely related compounds whose presence in environmental samples is highly unlikely. The preinjection volume generally is adjusted to approximately 1.0 mL (depending on the estimated concentration of PCBs in the sample) and spiked with at least one internal standard.

9.6.2.4 PCB Analysis by GC/ECD (EPA Method 8082)

The use of GC/ECD is more useful for forensic studies when PCB congener resolution is maximized using an extended GC run time and dual capillary columns. The ECD is an advantageous detector because it is very sensitive for chlorine-containing compounds and insensitive to background hydrocarbons. The instrument conditions that have demonstrated excellent analyte separation are presented in Table 9.21.

Table 9.21
Example of GC/ECD conditions for the analysis of PCB congeners. Equivalent conditions may be used if data quality guidelines are achieved.

GC Parameter	Program Settings
Column—High resolution equivalent) capillary.	SPB-octyl, 0.25 mm × 30 m × 0.25 μm (or DB-1, 0.25 mm × 30 m × 0.25 μm (or equivalent) DB-5, 0.25 mm × 30 m × 0.25 μm (or equivalent)
Initial column temperature:	75°C
Initial hold time:	2 minutes
Program rate 1:	15°C/minute
Final temp.:	150°C
Program rate 2:	2.5°C/minute
Final temp.:	290°C
Final hold time:	1 minute
Injector temperature:	270°C
Detector temperature:	290°C
Column flow rate:	~1 mL/min. (helium)
Injector:	Splitless
Sample volume:	1.0 μL

High resolution capillary GC columns that provide reliable congener separations include SPB-octyl (30 m long × 0.25 mm ID × 0.25 μm film, or equivalent), DB-1 (30 m long × 0.25 mm ID × 0.25 μm film, or equivalent), and DB-5 (30 m long × 0.25 mm ID × 0.25 μm film, or equivalent). The dual column combinations that work well are SPB-octyl and DB-1 or SPB-octyl and DB-5. Others are possible, but these column pairs are used frequently.

Using this method, the following minimum resolution is required:

- SPB-octyl column: PCB34 from PCB23 and PCB187 from PCB182
- DB-1 column: PCB156 from PCB157

Each of these columns will produce other coelutions; however, the coeluting congeners on one column usually are resolved on the other. Thus, the use of multiple columns together yields a high resolution PCB fingerprint with more than 125 discrete analytes. Longer versions of these columns have also been used successfully to promote PCB congener separation. However, full resolution of all 209 PCB congeners will not likely be routinely achieved with GC/ECD techniques due to its inability to resolve coeluting congeners based on the mass spectra. Nevertheless, the full resolution of the 209 congeners is often not required for source identification projects if an adequate number of congeners are resolved.

The calibration standards (209 congeners plus surrogates and IS compounds) are often too numerous to blend into a single stock standard for the generation of a multiconcentration series of initial calibration standards. Consequently, the multiconcentration calibration standards can be made

from a subset of the full PCB calibration that minimally includes the surrogates, internal standards, toxic PCBs (congeners with TEFs in Table 9.18), and representative congeners at each level of chlorination that elute at the beginning and end of the homolog range. A second calibration standard or series of standards should include all remaining congeners, such that together the subseries of standards include 209 PCB congeners.

Ongoing precision and accuracy for the determinative method can be evaluated by running daily continuing calibration standards at the beginning and end of the sample sequence. It is also recommended that each initial calibration be followed by the analysis of at least one Aroclor standard—preferably the Aroclor(s) present in the forensic study area. This reference standard will help benchmark instrument performance in data acquired over time and by multiple laboratories.

The concentrations of PCB congeners can be summed by level of chlorination to calculate the PCB homolog concentrations. Comparison of the PCB congener pattern in the Aroclor reference standards and samples can be used to determine the likely identity of Aroclor(s) in the sample. Accordingly, it is often useful to use the individual PCB congener concentrations to calculate the homolog, Aroclor, total, and TEF-weighted PCB concentrations.

9.6.2.5 PCB Analysis by GC/MS (EPA Method 8270C/680)

In an effort to improve the reliability of PCB congener identification, many laboratories have adopted the use of low resolution GC/MS to separate coeluting compounds. The main advantages of this method are that positive identification can be obtained from the mass spectra and coeluting isomers can be identified. The MS detector is less sensitive than the ECD detector when operated in full scan mode; however, when operated in the selected ion mode (SIM) it rivals the sensitivity of the ECD detector. The addition of high volume injection can further improve congener sensitivy that is on par with the HRMS method. This method has been successfully applied to environmental site assessment studies with PCB concentrations ranging from background to highly contaminated (National Academy of Sciences, 2000).

The use of the instrument conditions described for the GC/ECD method can be applied and optimized for the GC/MS method. The difference is that the laboratory may not need to analyze the extract in dual column mode, because the MS will generate mass spectra that confirm the analyte identity. Using these method specifications, analytical laboratories can routinely resolve more than 180 PCB congeners in water, soil, sediment, NAPL, wipe, air filter, and tissue samples, and greatly improve the value of the data to the forensic chemist.

The EPA reference for the low resolution MS technique can include adaptations of EPA Method 8270C and Method 680. The essential difference is that Method 8270C is a performance-based instrument method that includes the measurement of PCB concentrations in extracts from an assortment of matrices and cleanup steps. By contrast, Method 680 includes sample preparation and cleanup procedures for water and solid matrices. In addition, the analytical section of Method 680 is tailored to the analysis of PCB congeners and correction of analytical interferences caused by coeluting PCB congeners of different levels of chlorination. Unfortunately, Method 680 was not promulgated as a final method, so regulators do not always accept these results. Therefore, some laboratories prefer to reference the low-resolution MS method as EPA Method 8270C or as EPA Method 8270C/680.

9.6.2.6 PCB Analysis by GC/HRMS (EPA Method 1668A)

The analysis of PCBs by GC/HRMS using EPA 1668A is one of the best methods for identifying and tracking PCB source signatures in soil, sediment, water, tissue, and multiphase samples. This method provides the procedures for the preparation of common environmental sample matrices, removal of interferences, and detection of PCB congeners at trace concentrations. However, the high cost and lengthy period of analysis of this method often requires the use of more cost efficient methods (EPA Methods 8082 and 8270C) to limit the number of samples that require GC/HRMS confirmation.

Many aspects of EPA Method 1668A overlap with the procedures described for the EPA Method 8270C procedure. Samples are spiked with carbon isotope labeled analogs of toxic and homolog-specific PCB congeners. A one liter aliquot of aqueous sample is extracted using solid-phase extraction, separatory funnel extraction, or continuous liquid/liquid extraction. A 10 gram aliquot of solid samples is extracted in a Soxhlet/Dean-Stark (SDS) extractor.

The tissue procedure is more involved. A 10 g aliquot of homogenized tissue sample is fortified with isotopically labeled internal standards, mixed with anhydrous sodium sulfate, allowed to dry for 12 to 24 hours, and extracted for 18 to 24 hours using methylene chloride : hexane (1 : 1) in a Soxhlet extractor. The extract is evaporated to dryness, and the lipid content is determined. The extract is spiked with a labeled cleanup standard and back-extracted with sulfuric acid and/or base.

Various combinations of extract cleanups may be performed based on the presence of suspected interferences. Tissue samples are passed through an solid phase chromotography column that removes lipids by passage through alternating beds of silica gel, potassium silicate, and anhydrous sodium sulfate.

High molecular weight compounds are removed from soil, sediment, and tissue extracts using gel permeation chromatography. Extracts with excess polar and nonpolar interferences are purified by silica gel or Florisil chromatography. Activated carbon and high-performance liquid chromatography (HPLC) can be used for further isolation of specific congener groups (e.g., the more toxic PCB congeners associated with TEFs in Table 9.18). The final extract is concentrated to 100 μL and spiked with carbon isotope labeled PCB congener internal standards.

Determinative analysis is accomplished by GC/HRMS. Target analytes are identified by retention time and mass spectra. Two quantitation methods are used. Isotope dilution quantitation is used for the toxic PCBs and selected congeners that represent the elution range of each homolog. The remaining PCBs are measured with an internal standard technique, like that used for EPA Method 8270C.

A recent case study illustrated the benefit of using higher resolution PCB congener methods. The laboratory analyzed more than one thousand sediment samples using a less costly GC/ECD technique (EPA 8082). At a 10% frequency, sediment samples were sent to a specialized laboratory that used GC/MS (EPA 8270C/680) and GC/HRMS (EPA 1668A) interchangeably based on instrument availability. Under most circumstances, the use of multiple methods is discouraged. However, in this case, the accomplished PCB laboratory demonstrated that it could generate equivalent data at the project required detection limit of 1 mg/kg dry weight of total Aroclors using both methods. In addition to the applicable QA/QC parameters described previously, the laboratory demonstrated method equivalency by analyzing comparable initial calibrations with 209 PCB congeners, Aroclor standards, and representative sample extracts using both EPA 8270C/680 and EPA 1668A.

The results generated by both laboratories were combined and compared to the Aroclor standards using principal components analysis (PCA) (see Figure 9.37). This statistical tool (Pirouette, Version 3.02, Infometrix, Seattle, WA) is a factor analysis method that generates new independent variables (i.e., factors) that are linear combinations of the original input variables (e.g., PCB congener concentrations). This method reduces the dimensionality of the data to a few important principal components (axes) that best describe variations in the data. The first axis (1st PC) demonstrates the most prominent trend and successive axes (2nd PC, 3rd PC, etc.) demonstrate additional trends in decreasing order of statistical importance. In this investigation, one-half of the reporting limit was substituted for nondetected analytes. A sensitivity analysis demonstrated that other substitution values for nondetects (0, 0.1, 0.5, and 1 times the reporting limit) yielded the same results. Prior to PCA, the chemical concentration data were log-transformed and z-score

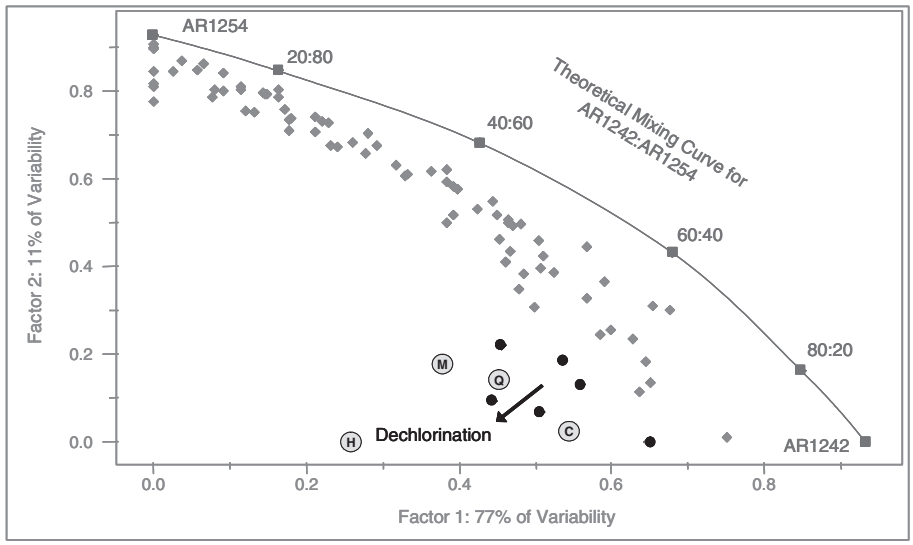


Figure 9.37

PCB congener patterns in sediment samples compared to a theoretical mixing curve (line) and AR1242 dechlorination data (letters representing AR1242 dechlorination patterns C, H, M, and Q in Johnson et al. *2006 and references therein). These data demonstrated the presence of Aroclor mixtures of AR1254 and AR1242 in many field samples. Some sediment samples contained mixtures of dechlorination by-products that most closely matched dechlorination patterns C and Q. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.*

normalized to reduce the effect of widely varying concentrations among samples and analytes. The primary objective of the PCA conducted for this study was to aid in the classification of field samples based on their chemical similarities or differences, without any preclassification as to their nature/source(s). The results of a PCA are presented using a two-dimensional factor score plots that represents 77% and 11% of the variability in Factors 1 and 2, respectively.

PCA initially helped explore the compositional features of PCB congeners patterns by creating a PCA model based on the PCA loadings of sediment samples collected throughout the study area. Thereafter, the Aroclor standard data were added to the field sample model and the presence of mixtures of AR1242 and AR1254 were identified. A mixing gradient was calculated between AR1242 and AR1254 to estimate the percentage of each Aroclor in each sample. Many of the samples fell off the mixing gradient, so the PCB congener results were recombined and reduced to match the analytes described in Johnson *et al.* (2006) (see Figure 9.36). The PCA model of the field samples in the combined PCB congener dataset were again compared to the AR1242 : AR1254 mixing gradient and published dechlorination patterns (see Figure 9.37). The PCA score plot demonstrates that some of the sediment samples that fell off the mixing gradient contained residues of anaerobic dechlorination with greatest compositional affinity with dechlorination patterns C and Q. In summary, high resolution PCB data helps identify the PCB origin, commingled Aroclors, and anaerobic dechlorination patterns. Although these compositional features are evaluated and verified on an individual basis between sample and Aroclor results, statistical tools, like PCA,

help explore multianalyte signatures (e.g., PCB congeners, PAHs and others) and illustrate these comparisons simultaneously.

REFERENCES

Aggarwal, P.K. and Hinchee, R.E. (1991) Monitoring in situ biodegradation of hydrocarbons by using stable carbon isotopes. *Environ. Sci. Technol.* 25, 1178–1180.

Ahlborg, U.G., Becking, G.C., Birnbaum, L.S., Brouwer, A., Derks, H.J.G.M., Feeley, M. *et al.* (1994) Toxic equivalency factors for dioxin-like PCBs. *Chemosphere* 28(6), 1049–1067.

Atlas, R.M., Boehm, P.D., and Calder, J.A. (1981) Chemical and biological weathering of oil from the Amoco Cadiz spillage within the littoral zone. Estuarine, *Coastal and Shelf Science* 12, 589–608.

Ahlborg, U.G., Becking, G.C., Birnbaum, L.S., Brouwer, A., Derks, H.J.G.M., Eruly, M., Goler, G., Hanberg, A., Larson, J.C., Liem, A.K.D., Safe, S.H., Schlatter, C., Weern, F., Younes, M., and Yrjanheikkl, E. (1994) Toxic equivalency factors for dioxin-like PCBs: Report on a WHO-ECEH and IPCS consultation, December 1993. *Chemosphere*, 28(6), 1049–1067.

American Society of Testing and Materials (1997) Standard test method for determination of MTBE, ETBE, TAME, DIPE, tertiary-amyl alcohol and C_1–C_4 alcohols in gasoline by gas chromatography. ASTM 05.02, 107–114.

Beall, P.W., Stout, S.A., Douglas, G.S., and Uhler, A. (2002) On the role of process forensics in the characterization of fugitive gasoline. *Env. Claims J.* 14(4), 487–505.

Bertsch, W. (1999) Two-dimensional gas chromatography. Concepts, instrumentation, and applications—Part 1: Fundamentals, conventional two-dimensional gas chromatography, selected applications. *J. High Resol. Chromatogr.* 22, 647–665.

Behar, F. and Pelet, R. (1984) Characterization of asphaltenes by pyrolysis and chromatography. *J. Anal. Appl. Pyrol.* 7, 121–135.

Boehm, P.D. and Farrington, J.W. (1984) Aspects of the polycyclic aromatic hydrocarbon geochemistry of recent sediments in the Georges Bank region. *Environ. Sci. Technol.* 18(11), 840–845.

Boehm, P.D., Page, D.S., Gilfillan, E.S., Bence, A.E., Burns, W.A., and Mankiewicz, P.J. (1998) Study of the fates and effects of the *Exxon Valdez* oil spill on benthic sediments in two bays in Prince William sound, Alaska. 1. Study design, chemistry, and source fingerprinting. *Environ. Sci. Technol.* 32, 567–576.

Boehm, P.D., Douglas, G.S., Brown, J.S., Page, D.S., Bence, A.E., Burns, W.A., and Mankiewicz, P.J. (2000) Comment on "Natural hydrocarbon background in benthic sediments of Prince William Sound, Alaska: Oil *vs* coal". *Environ. Sci. Technol.* 34, 2064–2065.

Bradley, L.J.N., Magee, B.H., and Allen, S.L. (1994) Background levels of polycyclic aromatic hydrocarbons (PAH) and selected metals in New England urban soils. *J. Soil. Contam.* 3, 349–361.

Bruce, L.G. (1993) Refined gasoline in the subsurface. *Am. Assoc. Petrol. Geol. Bull*. 77, 212–224.

Buddle, W.L. and Eichelberger, J.W. (1979) *Organic analysis using gas chromatography/mass spectrometry—A technique & procedures manual.* Michigan, Ann Arbor Sciences.

Chen, J., Fu, J., Sheng, G., Liu, D., and Zhang, J. (1996) Diamondoid hydrocarbon ratios: Novel maturity indices for highly mature crude oils. *Org. Geochem.* 25(3/4), 179–190.

Colombo, J.C., Pelletier, E., Brochu, C., and Khalil, M. (1989) Determination of hydrocarbon sources using n-alkane and polyaromatic hydrocarbon distribution indexes: Case Study—Rio de La Plata Estuary, Argentina. *Environ. Sci. Technol*. 23, 888–894.

Corseuil, H.X., Hunt, C.S., Santos, R.C., and Alverez, P.J. (1998) The influence of the gasoline oxygenate ethanol on aerobic and anerobic BTX biodegradation. *Wat. Res.* 32, 2065–2072.

Costa, H.J. and Sauer, T.C. (2005) Forensic approaches and considerations in identifying PAH background. *Environ. Foren.* 6, 9–16.

Douglas, G.S., Prince, R.C., Butler, E.L., and Steinhauer, W.G. (1994) The use of internal chemical indicators in petroleum and refined products to evaluate the extent of biodegradation. In Situ and On-Site Bioreclamation Conf.

Douglas, G.S. and Uhler, A.D. (1993) Optimizing EPA methods for petroleum-contaminated site assessments. *Environ. Testing Anal.* May/June 1993, 46–53.

Douglas, G.S., Naughton, G., Rhodes, I., and Kerr, J.M. (2001) Performance of TPHCWG analytical method for crude oils in soils. In *Risk-Based Decision-Making for Assessing Petroleum Impacts at Exploration and Production Sites*, S.J. McMillen, R.I. Magaw, and R.L. Carovillano, eds. Department of Energy and the Petroleum Environmental Research Forum. Government Printing Office. ISBN 0-9717288-0-1.

Douglas, G.S., Hardenstine, J., Owens, E.H., and Prince, R.C. (2002) The OSSA II pipeline oil spill: The character and weathering of the spilled oil. *Spill Science & Technol. Bull.* 7(3–4), 135–148.

Douglas, G.S., Burns, W.A., Bence, A.E., Page, D.S., and Boehm, P.B. (2004) Optimizing detection limits for the analysis of petroleum hydrocarbons in complex samples. *Environ. Sci. Technol.* 38(14), 3958–3964.

Douglas, G.S., Bence, A.E., Prince, R.C., McMillen, S.J., and Butler, E.L. (1996) Environmental stability of selected petroleum hydrocarbon source and weathering ratios. *Environ. Sci. Technol.* 30, 2332–2339.

Dutch, M., Long, E., Kammim, W., and Redman, S. (1998) Puget Sound ambient monitoring program marine sediment monitoring component—final quality assurance project & implementation plan. Measures of bioeffects associated with toxicants in Puget Sound: Survey of sediment contamination, toxicity, and benthic macroinfaunal community structure. Washington State Department of Ecology. Environmental Investigations and Laboratory Services Program. Marine Sediment Monitoring Unit.

Eglinton, G., Scott, P.M., Besky, T., Burlingame, A.L., and Calvin, M. (1964) Hydrocarbons of biological origin from a one-billion-year-old sediment. *Science* 145, 263–264.

Ehleringer, J.R. and Rundel, P.W. (1989) Stable isotopes: History, units, and instrumentation. In *Stable Isotopes in Ecological Research*, J.R. Ehleringer and K.A. Nagy, eds. Springer Verlag, NY, 1–15.

Emsbo-Mattingly, S.D., Stout, S.A., Uhler, A.D., and McCarthy, K.J. (2003a) Identifying manufactured gas plant tar in the environment using chemical fingerprinting. *Environ. Claims J*. 15(4), 477–490.

Emsbo-Mattingly, S.D. and Durell, G. (2003b) Identifying PCB mixtures and compositional changes in sediments with low resolution mass spectrometry and environmental forensic interpretation tools. *In Situ and On-Site Bioremediation, 7th Int'l. Symposium.* Orlando, FL.

Emsbo-Mattingly, S.D., Boehm, P.D., and Coleman, A. (2003c) Identifying PAHs from manufactured gas plant sites. *EPRI Technical Report 1005289.* Palo Alto, CA.

Emsbo-Mattingly, S.D. and Mauro, D.M. (1999) Occurrence of refined petroleum products at former manufactured gas plant (MGP) sites. International Conference on Contaminated Soil and Water. October 18–21, 1999 UMass–Amherst.

Emsbo-Mattingly, S.D., McCarthy, K.J., Uhler, A.D., Stout, S.A., Boehm, P.D., and Douglas, G.S. (2001a) Identifying and differentiating high and low temperature

tars at contaminated sites. *Journal of Contaminated Soil, Sediment & Water,* June/July Issue, 59–60.

Emsbo-Mattingly, S.D., Uhler, A.D., Stout S.A., McCarthy, K.J., Douglas, G.S., Brown, J.S., and Boehm, P.D. (2001b) Polycyclic aromatic hydrocarbon (PAH) chemistry of MGP tar and source identification in sediment. In *Sediments Guidance Compendium*, A. Coleman, ed. *Electric Power Research Institute, Technical Report 1005216*, 1–41.

EPA. (1979) Press Release. EPA Bans PCB Manufacture; Phases Out Uses. April 19, 1979.

EPA. (1983) Methods for chemical analysis of water and wastes. EPA-600-4-79-020 Environmental Monitoring and Support Laboratory, Cincinnati, OH.

EPA. (1985) Method 680, determination of pesticides and PCBS in water and soil/sediment by gas chromatography/mass spectrometry. Physical and Chemical Methods Branch, Environmental Monitoring and Support Laboratory (EMSL), Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, Ohio.

EPA. (1993a) Environmental monitoring and assessment program EMAP-Estuaries Virginian Province 1993 Quality assurance project plan. Office of Research and Development, Washington, DC.

EPA. (1993b) Guidance for assessing chemical contaminant data used for fish advisories. Volume 1. Fish sampling and anlaysis. EPA-823-R-93-002. Office of Water, Washington, DC.

EPA. (1995) Quality assurance for superfund environmental data collection activities—Publication 9200.2-16FS, February 1993, and EPA Order 5360.1, draft 1995 quality assurance order. Environmental Protection Agency, Washington, DC.

EPA. (1996) PCBs: Cancer Dose-Response Assessment and Application to Environmental Mixtures. EPA/600/P-96/001F.

EPA. (1997) Test methods for evaluating solid waste (SW-846). Update III. U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

EPA. (1998) Support document for the PCB disposal amendments, Final rule: Request for data in support of a TSCA use authorization for non-liquid PCB applications (http://www.epa.gov/fedrgstr/EPA-TOX/1998/June/Day-29/support.pdf).

EPA. (1999a) Guidelines establishing test procedures for the analysis of oil and grease and non-polar material under the Clean Water Act and Resource

Conservation and Recovery Act; Final Rule. 40 CFR Parts 136 and 260 [FRL-6341-9] RIN 2040-AC63.

EPA. (1999b) Method 1668, Revision A: Chlorinated biphenyl congeners in water, soil, sediment, and tissue by HRGC/HRMS. Office of Water EPA-821-R-00-002.

EPA. (1999c) EPA's program for cleaner vehicles and cleaner gasoline (EPA420-F-99-051, December 1999). Environmental Protection Agency, Washington, DC.

EPA. (1999d) Compendium method TO-15. Determination of volatile organic compounds (VOCs) in air collected in specially-prepared canisters and analyzed by gas chromatography/mass spectrometry (GC/MS). In *Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air,* 2e. Center for Environmental Research Information, Office of Research and Development, U.S. Environmental Protection Agency, Cincinnati, OH.

EPA. (2001) The relationship between SW-846, PBMs, and innovative analytical technologies. EPA 542-R-01-015. Office of Solid Waste and Emergency Response, Washington, DC.

EPA. (2002a) Guidance for developing quality systems for environmental programs. EPA QA/G-1. EPA/240/R-02/008. Office of Environmenatl Information, Washington, DC.

EPA. (2002b) OSWER Draft Guidance for evaluating the vapor intrusion to indoor air pathway from groundwater and soils (Subsurface Vapor Intrusion Guidance) November 2002 EPA530-D-02-004. Washington, DC.

EPA. (2002c) EPA Methods 3511-Organic compounds in water by microscale solvent extraction (MSE). Test methods for evaluating solid waste (SW-846). U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

EPA. (2002d) EPA Methods 3570-Organic compounds in solids by microscale solvent extraction (MSE). Test methods for evaluating solid waste (SW-846). U.S. Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington, DC.

EPA. (2003) PCB ID—Table of PCB Congeners & Other Species (http://www.epa.gov/toxteam/pcbid/table.htm).

EPA. (2005) Waste Management System; Testing and Monitoring Activities; Final Rule: Methods Innovation Rule and SW-846 Final Update IIIB.

EPA. (2006) SW-846 Test Methods—Frequently Asked Questions (http://www.epa.gov/sw-846/faqs_org.htm).

EPRI. (1995) Microscale Solvent Extraction (MSE) Methods for Analysis of Soils and Waters, Volumes I & II. EPRI Report, TR-105317-V1 and -V2, March 1995.

EPRI. (1999) The PCB Information Manual. Electric Power Research Institute TR-114091-V1. Palo Alto, CA.

Erickson, M.D. (1997) Analytical Chemistry of PCBs, 2e. Lewis Publishers, New York.

Federal Register. (1994) National Oil and Hazardous Substances Pollution Contingency Plan—Final Rule 40 CFR Parts 9 and 300.

Fort, R.C. and Schleyer, P.V.R. (1964) Adamantane: Consequences of the diamondoid structure. *Chemical Reviews* 277–300.

Fort, R.C. (1976) *Adamantane: The Chemistry of Diamond Molecules.* Marcel Dekker, New York.

Frame, G.M., Wagner, R.E., Carnahan, J.C., Brown, J.F., May, R.J., Smullen, L.A., and Bedard, D.L. (1996a) Comprehensive, quantitative, congener-specific analyses of eight aroclors and complete PCB congener assignments on DB-1 capillary GC columns. *Chemosphere* 33(4), 603–623.

Frame, G.M., Cochran, J.W., and Bowadt, S.S. (1996b) Complete PCB congener distributions for 17 Aroclor mixtures determined by 3 HRGC systems optimized for comprehensive, quantitative, congener specific analysis. *J. High Resol. Chromatogr.* 19, 657–668.

Frame, G.M. (1999) Improved procedure for single DB-XLB column GC-MS-SIM quantitation of PCB congener distributions and characterization of two different preparations sold as "Aroclor 1254." *J. High. Resol. Chromatogr.* 22(10), 533–540.

Frysinger, G.S. and Gaines, R.B. (1999) Comprehensive two-dimensional gas chromatography with mass spectrometry (GC × GC-MS) applied to the analysis of petroleum. *J. High Resol. Chromatogr.* 22, 251–255.

Frysinger, G.S., Gaines, R.B., and Reddy, C.M. (2002) GC × GC—A new analytical tool for environmental forensics. *Environ. For.* 3(1), 27–34.

Gaines, R.B., Frysinger, G.S., Hendrick-Smith, M.S., and Stuart, J.D. (1999) Oil spill identification by comprehensive two-dimensional gas chromatography. *Environ. Sci. Technol.* 33, 2106–2112.

Gibbs, L.M. (1990) Gasoline Additives—When and Why. SAE Technical Paper Series Int'l. Fuels and Lubricants Mtg. & Exposition, Tulsa, OK, 21.

Gibbs, L.M. (1993) How gasoline has changed SAE International Fuels and Lubricants Meeting and Exposition. SAE—The Engineering Society for Advancing Mobility Land Sea Air and Space International, Warrendale, PA, S01007–S010025.

Guetens, E.G., Jr., DeJovine, J.M., and Yogis, G.J. (1982) TBA aids methanol/fuel mix. *Hydrocarbon Processing* May 1982, 113–117.

Grice, K., Alexander, R., and Kagi, R.I. (2000) Diamondoid hydrocarbon ratios as indicators of biodegradation in Australian crude oils. *Org. Geochem.* 31, 67–73.

Harrington, R.R., Poulson, S.R., Drever, J.I., Colberg, P.J.S., and Kelly, E.F. (1999) Carbon isotope systematics of monoaromatic hydrocarbons: vaporization and adsorption experiments. *Org. Geochem*. 30, 765–775.

Henry, C.P. (1988) Additives for middle distillate and kerosine fuels. In *Distillate Fuel: Contamination, Storage and Handling. ASTM STP 1005*, H.L. Chesneau and M.M. Dorris, eds. American Society for Testing and Materials, Philadelphia, PA, 105–113.

Hites, R.A., Laflamme, R.E., and Windsor, J.G., Jr. (1980) Polycyclic aromatic hydrocarbons in an anoxic sediment core from the Pettaquamscutt River (Rhode Island, U.S.A.). *Geochim. Cosmochim. Acta* 44, 873–878.

Houlbeck, I. (2001) Polychloriated biphenyls (PCB) contaminated sites worldwide. In *PCBs: Recent advances in environmental toxicology and heatlh effects*, L.W. Robertson and L.G. Hansen, eds. Univerity of Kentucy Press.

Hughey, C.A., Hendrickson, C.L., Rodgers, R.P., and Marshall, A.G. (2001) Elemental composition analysis of processed and unprocessed diesel fuel by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy and Fuels* 15, 1186–1193.

Issac, H.J., (2001) *A Century of Separation Science.* Marcel Dekker, New York.

Jarman, W.M., Hilkert, A., Bacon, C.E., Collister, J.W., Ballschmiter, K., and Risebrough, R.W. (1998) Compound-specific carbon isotopic analysis of Aroclors, clophens, kaneclors, and phenoclors. *Environ. Sci. Technol.* 32(6), 833–836.

Jennings, W. (1980) Evolution and application of the fused silica column. *J. of High Res. Chrom. & Chrom. Commun*. 3, 601–608.

Jennings, W. (1987) *Analytical Gas Chromatography.* Academic Press, New York.

Johnson, G.W., Quensen, J.F., Ciarenzelli, J.R., and Hamilton, M.C. (2006) Polychlorinated bipenyls. In *Environmental Foreniscs: Contaminant Specific Guide*, R.D. Morrision and B.L. Murphy, eds. Elsevier Academic Press, New York.

Johnson, M.D. and Morrison, R.D. (1996) Petroleum fingerprinting: Dating a gasoline release. *Environmental Protection*, Sept. 1996, 37–39.

Kaplan, I.R., Galperin, Y., Alimi, H., Lee, R.P., and Lu, S.-T. (1996) Patterns of chemical changes during environmental alteration of hydrocarbon fuels. *Ground Water Monitoring and Remediation,* Fall 1996, 113–124.

Kaplan, I.R., Galperin, Y., Lu, S.-T., and Lee, R.-P. (1997) Forensic environmental geochemistry: Differentiation of fuel-types, their sources and release time. *Org. Geochem.* 27, 289–317.

Kelly T.J. and Holdren, M.W. (1995) Applicability of canisters for sample storage in the determination of hazardous air pollutants, *Atmos. Environ.*, 29, 2595–2608.

Kelley, C.A., Hammer, B.T., and Coffin, R.B. (1997) Concentrations and stable isotope values of BTEX in gasoline-contaminated groundwater. *Environ. Sci. Technol*. 31, 2469–2472.

Kennicutt, M.C. II. (1988) The effect of biodegradation on crude oil bulk and molecular composition. *Oil Chem. Pollut.* 4, 89–112.

Koester, C. (1999) Evaluation of analytical methods for the detection of ethanol in ground and surface water. In *Potential Ground and Surface Water Impacts*, D. Rice and G. Gannon, eds. Lawrence Livermore National Laboratory Report UCRL-AR-135949, Livermore, CA.

Koester, C. (2001) Evaluation of storage and analysis protocols for environmental water samples containing ethanol. Lawrence Livermore National Laboratory, Livermore, CA.

Kujawinski, E.B. (2002) Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI FT-ICR MS): Characterization of complex environmental mixtures. *Environ. Forensics*, 3, 207–216.

Lautamo, R., Schirmer, R., and Jennings, W. (1998) A brief history of column developments in gas chromatography. *Amer. Laboratory* 30(25), 28–32.

Lee, M.L., Prado, G.P., Howard, J.B., and Hites, R.A. (1977) Source identification of urban airborne PAHs by GC/MS and high resolution MS. *Biomed. Mass Spectrom.* 4, 182–186.

Lee, M.L., Vassilaros, D.L., White, C.M., and Novotny, M. (1979) Retention indices for programmed-temperature capillary column gas chromatography of polycyclic aromatic hydrocarbons. *Anal. Chem.* 51(6), 768–773.

Lundegard, P.D. (2006) Methane. In *Environmental Forensics, Contaminant Specific Guide*, R.D. Morrison and B.L. Murphy, eds. Academic Press, Boston, 97–110.

Lundegard, P.D., Haddad, R., and Brearley, M. (1998) Methane associated with a large gasoline spill: Forensic determination of origin and source. *Environ. Geosci.* 5, 69–78.

Macchia, T., Mancinelli, R., Gentili, S., Lugaresi, E.C., Raponi, A., and Taggi, F. (1995) Ethanol in biological fluids-headspace GC measurement. *J. Anal. Toxicol.* 19, 241–246.

Maney, J.P. and Wait, A.D. (1991) The importance of measurement integrity. *Environ. Lab* 3(5), 20–25.

Mansuy, L., Philp, R.P., and Allen, J. (1997) Source identification of oil spills based on the isotopic composition of individual components in weathered oil samples. *Environ. Sci. Technol.* 31, 3417–3425.

Mauro, D. (2000a) Chemical source attribution at former MGP sites. *EPRI Technical Report No. 1000728,* EPRI, Palo Alto, CA.

Mauro, D. (2000b) Simultaneous determination of volatile and semivolatile organic compounds in soil. *Remediation,* 10(4), Autumn 2000.

Mauro, D. and Emsbo-Mattingly, S.D. (1999) Simultaneous measurement of volatile and semivolatile compounds: Introducing Methods 3511 and 3570. In *15th Annual International Conference on Contaminated Soils and Water.* Amherst, MA October 18–21, 1999.

McCarthy, K.J. and Uhler, R.M. (1994) Optimizing gas chromatography conditions for improved hydrocarbon analysis. In *Proceedings Ninth Annual Conference on Contaminated Soils*. Amherst, MA.

McCarthy, K.J., Uhler, A.D., and Stout, S.A. (1998) Weathering affects petroleum identification. *Soil & Groundwater Cleanup,* Aug/Sept. 1998.

McCarthy, K.J., Mattingly, S.D., Uhler, A.D., Rezendes, A., Stout, S.A., and Douglas, G.S. (2006) Forensic characterization of subsurface and indoor air for evidence of vapor intrusion at contaminated sites. 16th Annual AEHS Meeting & West Coast Conference on Soils, Sediments and Water, San Diego, CA.

McCarver-May, D.G. and Durisin, L. (1997) An accurate, automated, simultaneous gas chromatographic headspace measurement of whole blood ethanol and acetaldehyde for human *in vivo* studies. *J. Anal. Toxicol*. 21, 134–141.

McLafferty, F.W. and Turecek, F. (1993) *Interpretation of Mass Spectrometry.* University Science Books, Sausalito, California.

McRae, C., Love, G.D., and Snape, C.E. (1997) Application of gas chromatography-isotope ratio mass spectrometry to source polycyclic aromatic hydrocarbon

emissions. *American Chemical Society, Division of Environmental Chemistry—Preprints of Extended Abstracts* 37, 315–315.

Meckenstock, R.U., Morasch, B.W.R., Schink, B., Annweller, E., Michaelis, W., and Richnow, H.H. (1999) 13C/12C isotope fractionation of aromatic hydrocarbons during microbial degradation. *Environ. Microbiol.* 1, 409–414.

Moran, M.J., Clawges, R.M., and Zogorski, J.S. (2000) Identifying the usage patterns of methyl tert-butyl ether (MTBE) and other oxygenates in gasoline using gasoline surveys. American Chemical Society, Division of Environmental Chemistry. March 26–30, 2000. Exploring the Environmental Issues of Mobile, Recalcitrant Compounds in Gasoline. 40(1), 209–212.

Morrison, R.D. (2000) *Environmental Forensics*. CRC Press, Boca Raton, 351.

Mulroy, P.T. and Ou, L.-T. (1998) Degradation of tetraethyllead during the degradation of leaded gasoline hydrocarbons in soil. *Environ. Toxicol. Chem.* 17, 777–782.

Murphy, B. and Morrison, R. (2002) *Environmental Forensics: A Contaminant Specific Approach*. Elsevier Publishing Co., San Francisco, CA, 466–531.

National Academy of Sciences (2000) A risk-management strategy for PCB contaminated sediments committee on remediation of PCB-contaminated sediments. National Academy Press Washington, DC.

Nelson, R.K., Kile, B.M., Plata, D.L., Sylva, S.P., Xu, L., Reddy, C.M., *et al.* (2006) Tracking the weathering of an oil spill with comprehensive two-dimensional gas chromatography. *Environ. Forensics* 7, 33–44.

NIST. (2006) NIST/EPA/NIH Mass Spectral Library with Search Program: (Data Version: NIST 05, Software Version 2.0d). National Institute of Standards and Technology, Gaithersburg, MD.

NOAA. (1993) Trace organic analytical procedures. In *Sampling and Analytical Methods of the National Status and Trends Program Mussel Watch Project: 1984–1992*. NOAA Technical Memorandum NOS/ORCA 71. National Oceanic and Atmospheric Administration, Silver Spring, MD.

NOAA. (1998) Trace organic analytical procedures. In *Sampling and Analytical Methods of the National Status and Trends Program Mussel Watch Project: 1993–1996 Update*. NOAA Technical Memorandum NOS/ORCA/CMBAD 130. National Oceanic and Atmospheric Administration, Silver Spring, MD.

NOAA. (2002) NOAA Damage Assessment Center Rapid Assessment Program Sampling Guidelines.

Nordtest (1991) Nordtest Method. Oil spill identification, NORDTEST, Esbo, Finland, 24.

Novotny, M., Strand, J.W., Smith, S.L., Wiesler, D., and Schwende, F.J. (1980) Compositional studies of coal tar by capillary gas chromatography/mass spectrometry. *Fuel* 60, 213–220.

O'Malley, V.P., Abrajano, T.A., Jr., and Hellou, J. (1996) Stable carbon isotopic apportionment of individual polycyclic aromatic hydrocarbons in St. John's Harbour, Newfoundland. *Environ. Sci. Technol*. 20, 634–639.

Overton, E.B., McFall, J.A., Mascarella, S.W., Steele, C.F., Antoine, S.A., Politzer, I.R., and Laseter, J.L. (1981) Identification of petroleum residue sources after a fire and oil spill. *Proceedings of the 1981 Oil Spill Conference.* American Petroleum Institute Publication #720, 541–546.

Page, D.S., Boehm, P.D., Douglas, G.S., and Bence, A.E. (1995) Identification of hydrocarbon sources in benthic sediments of Prince William Sound and the Gulf of Alaska following the *Exxon Valdez* oil spill. In: *Exxon Oil Spill: Fate and Effects in Alaskan Waters*, ASTM Special Technical Publication #219, Peter G. Wells, James N. Butler, and Jane S. Hughes, eds., American Society for Testing and Materials, Philadelphia.

Peters, K.E., Scheuerman, G.L., Lee, C.Y., Moldowan, J.M., Reynolds, R.N., and Pena, M.M. (1992) Effects of refinery processes on biological markers. *Energy Fuels* 6, 560–577.

Peters, K.E. and Moldowan, J.M. (1993) *The Biomarker Guide*. Prentice Hall Inc., Englewood Cliffs, NJ, 363.

Peters, K.E., Moldowan, J.M., McCaffrey, M.A., and Fago, F.J. (1996) Selective biodegradation of extended hopanes to 25-norhopanes in petroleum reservoirs. Insights from molecular mechanics. *Org. Geochem.* 24, 765–783.

Peters, K.E., Walters, C.C., and Moldowan, J.M. (2005) *The Biomarker Guide*, Volumes 1 and 2, Cambridge University Press.

Philp, R.P. (2002) Application of stable isotopes and radioisotopes in environmental forensics. In *Introduction to Environmental Forensics*, B.L. Murphy and R.D. Morrison, eds. Academic Press, San Diego, CA, 99–136.

Prince, R.C. (2002) Biodegradation of petroleum and other hydrocarbons. In *Encyclopedia of Environmental Microbiology*, G. Bitton, ed. John Wiley, New York, 2402–2416.

Prince, R.C. and Douglas, G.S. (2005) Quantification of hydrocarbon biodegradation using internal markers. In *Manual of Soil Analysis—Monitoring and Assessing Soil Bioremediation*, R. Margesin and F. Schinner, eds. Springer-Verlag, Berlin, 179–188.

Plourde, K.L., Hendrick, M.S., Greimann, D.E., and Reilly, T.R. (1995) Nets: A novel approach to thin sheen oil sampling. *Proceedings of the Second International Oil Spill Research and Development Forum*.

Pruell, R.J. and Quinn, J.Q. (1988) Accumulation of polycyclic aromatic hydrocarbons in crankcase oil. *Environ. Pollut.* 49(2), 89–97.

Qian, K., Rodgers, R.P., Hendrickson, C.L., Emmett, M.R., and Marshall, A.G. (2001) Reading chemical fine print: Resolution and identification of 3000 nitrogen-containing aromatic compounds from a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry of heavy petroleum crude oil. *Energy and Fuels* 15, 492–498.

Quinn, J.G. and Latimer, J.S. (2003) Protocols for the collection of water, sediment, and tissue samples for chemical analysis of petroleum hydrocarbons. Prepared for the RIOST OIL-Spill Science Workshop, November 24, 2003.

Raia, J.C., Blakley, C.R., Fuex, A.N., Villalanti, D.C., and Fahrenthold, P.D. (2004) Evaluation of environmental samples containing heavy hydrocarbon components in environmental forensic investigations. *Environ. Forensics* 5, 21–32.

Rodgers, R., Blumer, E.N., Freitas, M.A., and Marshall, A.G. (2000) Complete compositional monitoring of the weathering of transportation fuels based on elemental compositions from Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* 34(9), 1671–1678.

Rostad, C.E. (2006) Differentiation of commercial fuels based upon polar components using negative electrospray ionization/mass spectrometry. *Environ. Forensics* 7, 5–14.

Rushneck, D.R., Beliveau, A., Fowler, B., Hamilton, C., Hoover, D., Kaye, K. *et al.* (2004) Concentrations of dioxin-like PCB congeners in unweathered Aroclors by HRGC/HRMS using EPA Method 1668A, *Chemosphere* 54, 79–87.

Saber, D., Mauro, D., and Sirivedhin, T. (2006) Environmental forensics investigation in sediments near a former manufactured gas plant. *Environ. Forensics* 7, 65–75.

Sauer, T.C. and Boehm, P.D. (1991) The use of defensible analytical chemical measurements for oil spill natural resources damage assessment. *Proc. International Oil Spill Conference*, 363.

Sauer, T.C. and Boehm, P.D. (1995) Hydrocarbon chemistry analytical methods for oil spill assessment. *Marine Spill Response Corporation Technical Report Series 95-032.*

Sauer, T.C. and Uhler, A.D. (1994–1995) Pollutant source identification and allocation: Advances in hydrocarbon fingerprinting. *Remediation, Winter Issue*, 25–50.

Sessions, A.L., Burgoyne, T.W., Schimmelmann, A., and Hayes, J.M. (1999) Fractionation of hydrogen isotopes in lipid biosynthesis. *Org. Geochem.* 30, 1193–1200.

Smallwood, B.J., Philp, R.P., and Allen, J.D. (2002) Stable carbon isotope composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. *Organic Geochemistry* 33, 149–159.

Smallwood, B.J. *et al.* (2001) The use of stable isotopes to differentiate specific source markers for MTBE. *Environmental Forensics.* 2(3), 215–221.

Sofer, Z. (1984) Stable carbon isotope compositions of crude oils: Application to source depositional environments and petroleum alteration. *Am. Assoc. Petrol. Geol. Bull.* 68, 31–49.

Squillace, P.J., Pope, D.A., and Price, C.V. (1995) Occurrence of the gasoline additive MTBE in shallow ground water in urban and agricultural areas. FS-114-95, United States Geological Survey, National Water Quality Assessment Program, Rapid City, SD, 4.

Stout, S.A., Davidson, J.M., McCarthy, K.J., and Uhler, A.D. (1999a) Gasoline additives—Usage of lead and MTBE. *Soil & Groundwater Cleanup,* Feb/Mar 1999.

Stout, S.A. and Douglas, G.S. (2004) Diamondoid hydrocarbons—Application in the chemical fingerprinting of natural gas condensate and gasoline. *Environmental Forensics Journal*, 5, 225–235.

Stout, S.A., McCarthy, K.J., Seavey, J.A., and Uhler, A.D. (1999b) Application of low boiling biomarkers in assessing liability for fugitive middle distillate petroleum products. *Proc. 9th Ann. West Coast Conf. Contaminated Soils and Water*, March, 1999 (Abstract).

Stout, S.A., Naples, W.P., Uhler, A.D., McCarthy, K.J., Roberts, L.G., and Uhler, R.M. (2000a) Use of quantitative biomarker analysis in the differentiation and characterization of spilled oil. SPE 61460 Prepared for Presentation at the SPE International Conference on Health, Safety, and the Environment in Oil and Gas Exploration and Production. 26–28 June 2000 Stavanger, Norway. Society of Petroleum Engineers, Inc., Richardson, TX, 15.

Stout, S.A., Uhler, A.D., and McCarthy, K.J. (1999c) Biomarkers—Underutilized components in the forensic toolkit. *Soil & Groundwater Cleanup* June/July 1999, 58–59.

Stout, S.A. and Uhler, A.D. (2003) Distinguishing background hydrocarbons from contamination using chemical fingerprinting. *Environ. Claims J.* 15(2), 241–259.

Stout, S.A., Uhler, A.D., and McCarthy, K.J. (2004) Characterizing the source of fugitive middle distillate fuels—A case study involving railroad diesel fuel, Mandan, North Dakota. *Environ. Claims J.,* 16(2), 157–172.

Stout, S.A., Uhler, A.D., McCarthy, K.J., and Emsbo-Mattingly, S.D. (2002) Chemical fingerprinting of hydrocarbons. In *Introduction to Environmental Forensics*, B. Murphy and R. Morrison, eds. Academic Press, New York, 135–260.

Stout, S.A., Uhler, A.D., and McCarthy, K.J. (2005) Middle distillate fuel fingerprinting using drimane-based bicyclic sesquiterpanes. *Environ. Forensics* 6(3), 241–252.

Stout, S.A., Douglas, G.S., and Uhler, A.D. (2006) Automotive gasoline. In *Environmental Forensics: A Contaminant Specific Approach*. B. Murphy and R. Morrison, eds. Elsevier Publishing Co., San Francisco, CA, 466–531.

Tangerman, A. (1997) Highly sensitive gas chromatographic analysis of ethanol in whole blood, serum, urine, and fecal supernants by the direct injection method. *Clin. Chem.* 43, 1003–1009.

Tswett, M. (1906) *Ber. Dtsch. Bot. Ges.* 24, 316.

Uhler, A.D., Stout, S.A., and McCarthy, K.J. (1998–1999) Increase success of assessments at petroleum sites in 5 steps. *Soil & Groundwater Cleanup*, Dec/Jan 1998–1999, 13–19.

Uhler, R.M., Healey, E.M., McCarthy, K.J., Uhler, A.D., and Stout, S.A. (2003) Molecular fingerprinting of gasoline by a modified EPA 8260 gas chromatography/mass spectrometry method. *Int. J. Environ. Anal. Chem.* 83(1), 1–20.

Uhler, A.D., and Emsbo-Mattingly, S.D. (2006) Environmental stability of PAH source indices in pyrogenic tars. *Bulletin of Environ. Contam. and Tox.* 76, 689–696.

USACE. (2005) Guidance for evaluating performance-based chemical data. EM 200-1-10 Department Of The Army U.S. Army Corps of Engineers, Washington, DC. 20314-1000.

Vazquez-Duhalt, R. (1989) Environmental impact of used motor oil. Sci. *Total Environ.* 79, 1–23.

Venkatesan, M.I. (1988) Occurrence and possible sources of perylene in marine sediments—A review. *Mar. Chem.* 25, 1–27.

Villalanti, D.C., Raia, J.C., and Maynard, J.B. (2000) High-temperature simulated distillation applications in petroleum characterization. In *Encyclopedia of*

Analytical Chemistry, R.A. Meyers, ed. John Wiley & Sons Ltd, Chichester, UK, 6726–6741.

Wait, D.A. and Douglas, G.S. (1995) QA for non-standard analytical methods. *Environ. Lab.* 7(6), 30–34.

Watts, M.T. and MacDonald, O.L. (1990) The effect of sodium chloride concentration, water content, and protein on the gas chromatographic headspace analysis of ethanol in plasma. *Amer. J. Clin. Pathol.* 93, 357–362.

Wakeham, S.G., Schaffner, C., and Giger, W. (1980) Polycyclic aromatic hydrocarbons in recent lake sediments—I. Compounds having anthropogenic origins. *Geochim. Cosmochim. Acta* 44, 403–413.

Wakim, J.M., Schwendener, H., and Shimosato, J. (1990) Gasoline octane improvers. CEH Marketing Research Report. SRI International. Menlo Park, CA.

Wang, Z., Fingas, M., and Li, K. (1994) Fractionation of a light crude oil and identification and quantitation of aliphatic, aromatic, and biomarker compounds by GC-FID and GC-MS, part I. *J. Chromatogr. Sci.* 32, 361–382.

Wang, Z., Fingas, M., and Sergy, G. (1994) Study of 22-year-old *Arrow* oil samples using biomarker compounds by GC/MS. *Environ. Sci. Technol*. 28, 1733–1746.

Wang, Z. and Fingas, M. (1999a) Identification of the source(s) of unknown spilled oils. Paper #162. In 1999 International Oil Spill Conference, 9.

Wang, Z., Fingas, M., and Page, D.S. (1999b) Oil spill identification. *J. Chrom. A* 842, 369–411.

Wang, Z., Stout, S.A., and Fingas, M. (2006) Forensic fingerprinting of biomarkers for oil spill identification and source identification (A Review). *Environ. Forensics*, 7(2), TBD.

Waples, D.W. and Machihara, T. (1990) Application of sterane and triterpane biomarkers in petroleum exploration. *Bulletin of Canadian Petroleum Geology* 38, 357–380.

Whittaker, M., Pollard, S.J.T., and Fallick, T.E. (1995) Characterisation of refractory wastes at heavy oil-contaminated sites: A review of conventional and novel analytical methods. *Environ. Tech.* 16, 1009–1033.

Williams, J.A., Bjoroy, M., Dolcater, D.L., and Winters, J.C. (1986) Biodegradation in South Texas Eocene oils—Effects on aromatics and biomarkers. *Org. Geochem*. 10, 451–461.

Wingert, W.S. (1992) GC-MS analysis of diamondoid hydrocarbons in Smackover petroleums. *Fuel* 71, 37–43.

Youngblood, W.W. and Blumer, M. (1975) Polycyclic aromatic hydrocarbons in the environment: Homologous series in soils and recent marine sediments. *Geochim. Cosmochim. Acta* 39, 1303–1314.

CHAPTER 10

APPLICATION OF STABLE ISOTOPES AND RADIOISOTOPES IN ENVIRONMENTAL FORENSICS

R. Paul Philp and Emilie Jardé

10.1 Introduction 455
10.2 Radioisotope Age Dating of Contaminants in Sediments and Groundwater 458
10.2.1 Sediment Dating 458
10.2.2 Groundwater Dating 459
10.3 Use of Stable Isotopes to Identify Contaminant Sources 460
10.3.1 Origin of Isotopic Carbon Differences 461
10.3.2 Reporting Isotope Measurements 462
10.3.3 Bulk Isotope Values 463
10.3.4 Carbon Isotope GCIRMS for Hydrocarbons 467
10.3.5 Sources and Sink of Atmospheric Gases 476
10.3.6 Use of GCIRMS for MTBE 477
10.4 Other Isotopes 479
10.5 Characterization of Chlorinated Compounds 483
10.6 Biodegradation 485
10.7 Combining Isotope Methods with Other Methods 491
References 497

10.1 INTRODUCTION

In recent years there have been many published studies concerned with the origin and fate of pollutants in aquatic environments such as rivers, lakes, groundwater, and coastal waters (Macko *et al.*, 1981; Farran *et al.*, 1987; Galt *et al.*, 1991; Eganhouse *et al.*, 1993; Oros *et al.*, 1996; Dempster *et al.*, 1997; Smirnov *et al.*, 1998; Wang *et al.*, 2004; Ricking *et al.*, 2005). Sources of contaminants are many and varied and include natural seepage of crude oils; ship traffic; supertankers spilling crude oils; leaking storage tanks, pipelines, polychlorinated biphenyls (PCBs), pesticides, dioxins, and many other types of chemical spills from a wide variety of sources (Aislabie *et al.*, 2004; Alzaga *et al.*, 2004; Ohe *et al.*, 2004; Chapman and Riddle, 2005; Franco *et al.*, 2005; Hong *et al.*, 2005; Lee *et al.*, 2005; Chen *et al.*, 2006; Samara *et al.*, 2006).

Identification, quantification, and monitoring the fate of these pollutants in the environment are important in providing a better response to such spills and for the purposes of determining responsibility for the spill and liability for the clean-up.

Common approaches for characterization of such spills, comprised of organic compounds, and identification of potential sources, generally rely upon analyses by gas chromatography (GC) and gas chromatography-mass spectrometry (GCMS). Correlations are made on the basis of the molecular distribution of saturate or aromatic hydrocarbons and more specifically, in the case of crude oil spills, biomarker fingerprints (Wang *et al.*, 1994; Brenner *et al.*, 2002; Peters and Fowler, 2002; Stout *et al.*, 2004; Wang *et al.*, 2005; Wang and Christensen, 2005). In certain cases GC and GCMS data may be ambiguous, and inconclusive, as result of weathering processes such as evaporation, photooxidation, water-washing, and biodegradation changing the distribution of components in the spilled material vs. the original sample (Milner *et al.*, 1977; Lafargue and Barker, 1988; Palmer, 1993). Depending on the nature of the compounds involved, evaporation can occur in the first few hours after a spill and remove the more volatile components. If the accident occurs in an aquatic environment, water-washing will remove the more water-soluble components, such as hydrocarbons below C_{15}, and some of the C_{15+} aromatic compounds that are more water-soluble than paraffins (Lafargue and Barker, 1988; Palmer, 1993; Fuentes *et al.*, 1996). At the same time, biodegradation will also start to affect the distribution of individual compounds, although the rate will depend upon the nature of the spill and environment into which it spilled (Thornton *et al.*, 2001; Brenner *et al.*, 2002; Gray *et al.*, 2002; Mancini *et al.*, 2002; Meckenstock, 2002; Schüth *et al.*, 2003; Griebler *et al.*, 2004; Schirmer and Butler, 2004; Kuder *et al.*, 2005; Haws *et al.*, 2006).

Correlation of spilled material with its suspected source, whether it is hydrocarbon-based or of some other chemical origin, requires discriminative parameters that are relatively insensitive to weathering processes. GC and GCMS are not always the ultimate answer to such a problem. One alternative approach is the use of stable isotopic compositions, typically carbon or hydrogen isotopic values (Hartman and Hammond, 1981; Macko *et al.*, 1981; Macko and Parker, 1983; Farran *et al.*, 1987; Wassenaar and Hobson, 2000; Fang *et al.*, 2002; Pond *et al.*, 2002; Wang *et al.*, 2004; Glaser *et al.*, 2005; Gürgey *et al.*, 2005; Horii *et al.*, 2005; Wen *et al.*, 2005) and in selected cases chlorine, nitrogen, oxygen or sulfur isotope compositions (Warmerdarm *et al.*, 1995; Sturchio *et al.*, 1998; Beneteau *et al.*, 1999; Reddy *et al.*, 2000; Coffin *et al.*, 2001; Jendrzejewski *et al.*, 2001; Drenzek *et al.*, 2002; Numata *et al.*, 2002; Shouakar-Stash *et al.*, 2003). The most common parameter would be carbon isotopes $^{13}C/^{12}C$, but in relatively recent years there has been an increasing number of

papers reporting the use $^{35}Cl/^{34}Cl$; D (deuterium)/H; $^{18}O/^{16}O$; in a variety of geochemical applications, including spill/source correlations and rate of biodegradation of individual contaminants (Philp *et al.*, 2002; Pond *et al.*, 2002; Griebler *et al.*, 2004; Kuder *et al.*, 2005; Zwank *et al.*, 2005). Sources of a chemical, or timing of a contaminant release, can be traced using isotopic signatures since these signatures have been shown to vary from different manufacturers as a result of variations in source materials, conditions, and pathways used to synthesize specific compounds. The approach has been used successfully with variations in the $\delta^{13}C$ values of BTEX (Dempster *et al.*, 1997); $\delta^{13}C$ in MTBE (Smallwood *et al.*, 2001); $\delta^{13}C$ in PCBs (Drenzek *et al.*, 2002; Yanik *et al.*, 2003); $\delta^{13}C$, $\delta^{2}H$, and $\delta^{37}Cl$ in chlorinated solvents (Shouakar-Stash *et al.*, 2003); and $\delta^{13}C$ and $\delta^{15}N$ in trinitrotoluene (Coffin *et al.*, 2001). It should be noted that although this approach has been very successful it is not going to be successful in every case. For example, no differences in $\delta^{13}C$ were found in PAHs from two different creosote-contaminated sites (Hammer *et al.*, 1998).

Isotopic compositions may be determined as either a bulk isotopic value for the total spill, or through the use of combined gas chromatography-isotope ratio mass spectrometry (GCIRMS; Hayes *et al.*, 1990; O'Malley *et al.*, 1994; Boreham *et al.*, 1995; Dowling *et al.*, 1995; O'Malley *et al.*, 1996) to determine the isotopic composition of individual compounds in a mixture. Carbon and hydrogen isotopic values for the individual compounds are now routinely determined by GCIRMS and attempts are under way to develop methods for determination of chlorine isotopes of individual compounds by GCIRMS. However, for chlorine most of the published methods currently require isolation of the individual compounds prior to isotopic determination (Warmerdam *et al.*, 1995; Holt *et al.*, 1997; Sturchio *et al.*, 1998; Holt *et al.*, 2001; Jendrzejewski *et al.*, 2001; Drenzek *et al.*, 2002; Shouakar-Stash *et al.*, 2003).

Uses of isotopes in environmental forensics can be divided into two main areas, those associated with stable isotopes and those associated with radioisotopes. The division is actually more specific than that since radioisotope applications generally are associated with age dating applications, whereas stable isotopes commonly are associated with correlations or source determination. Radioisotopes commonly used for dating of sediments and groundwater in the context of environmental forensics include cesium-137 ^{137}Cs; lead-210 ^{210}Pb; and tritium ^{3}H and will be described in more detail later. Although there are other radioisotopes that may have applications in various geochemical exploration topics, it should be remembered that this chapter is concerned with environmental applications. Stable isotopes have found a much greater range of applications in environmental studies than have radioisotopes and such a bias is naturally reflected in this chapter.

The chapter is divided into six sections covering these general areas. The first will deal with radioisotope age dating; the second with the use of stable isotopes in the identification of contaminant sources (especially C and H); the third with uses of other stable isotopes in environmental forensics; the fourth with specific use of isotopes for chlorinated compounds; the fifth with biodegradation processes; and finally the sixth will examine the combined use of isotopes and data from other techniques in environmental forensics. It is the intent of this chapter to provide you with a comprehensive insight into the use of isotopes in environmental forensics rather than exhaustive review of all the published literature.

10.2 RADIOISOTOPE AGE DATING OF CONTAMINANTS IN SEDIMENTS AND GROUNDWATER

The need, or desire, to age date a contaminant in a sediment or groundwater like most environmental matters is basically driven by money and the need to assign blame or determine the responsible party. In many cases a contaminated site will have a long list of previous owners, and the ability to provide an indication as to when any contaminant spill occurred will enable some of the previous owners to be absolved of clean-up costs and other liability. Hence the main uses of radioisotopes in environmental forensics are for dating sediments and groundwater. However, it should be noted that unlike other techniques described later, this particular application of radioisotopes is still in its infancy and at the present time is probably more commonly used for research purposes.

10.2.1 SEDIMENT DATING

Sediment cores can be dated to determine historical inputs to aquatic systems, however, it needs to be emphasized there are limitations to the methods. Disturbance of the sediment cores in any way by storms or animals can preclude accurate dating of the cores or require adjustments in interpretation. The two main isotopes used for sediment dating are ^{137}Cs and ^{210}Pb. ^{137}Cs is derived from atmospheric nuclear weapons testing and when removed from the atmosphere is incorporated into the sediments. Once deposited it is bound into the sediment by sorbing to fine clay particles and organic material and is essentially nonexchangeable (Tamura and Jacobs, 1960; Lieser and Steinkopff, 1989). The first observation of ^{137}Cs in sediments dates back to about 1954 and reaches a well-documented maximum in 1963, followed by minor increases in 1971 and 1974 (Ritchie and McHenry, 1990). The global deposition of ^{137}Cs

can be estimated based on the better documented deposition of strontium-90 (^{90}Sr) (Hermanson, 1990). However, for many purposes having the absolute deposition rate is unnecessary since dating can be accomplished by noting the 1963 ^{137}Cs peak and measuring its depth in the sediment (Pennington and Cambray, 1973). ^{137}Cs sediment dating has been used for a number of purposes, including reconstructing a history of arsenic inputs along a tidal waterway in Tacoma, Washington (Davis *et al.*, 1997); tracking historical trends for PCBs and pesticides in reservoir sediments from central and southeastern United States (van Metre *et al.*, 1997); and dating the occurrence of dioxins and dibenzofurans in remote lakes (Czuczwa and Hites, 1984; Juttner *et al.*, 1997).

^{210}Pb has a half life of 22.6 years and is useful for dating sediments deposited during the past 100 years, particularly when incorporated with ^{137}Cs dating. Uranium-238 (^{238}U) decays to radon-222 (^{222}Rn) and then through a series of isotopes to ^{210}Pb. Radon is a gas that seeps up into the atmosphere from ^{238}U containing rocks. Once it enters the atmosphere it soon decays to ^{210}Pb, which falls out and can be deposited in sediments where it decays. The decrease in ^{210}Pb with sediment depth indicates the age of a sediment level and of any undisturbed sediments at that level. If the sediments contain ^{238}U, a correction for ^{210}Pb formed in the sediments may be necessary. ^{210}Pb measurements have been used to date Lake Ontario sediments and historical input of PCBs, pesticides and other hydrophobic organic compounds (Wong *et al.*, 1995) and metal accumulations in Lake Zurich (von Gunten *et al.*, 1997). Gallon *et al.* (2006) analyzed sediment cores from Canadian Shield headwater lakes and dated them using Pb isotopes. The depth distributions of stable Pb isotope ratios show the presence of several isotopically distinct Pb types since the preindustrial period, allowing the geographical area impacted by the smelter Pb emissions to be traced.

10.2.2 GROUNDWATER DATING

In addition to sediment dating, isotopes have been used for dating groundwater for use in determination of hydrogeological parameters. Such parameters are used for reverse plume modeling to determine source location and release date, modeling of leachate and infiltration, to link groundwater and surface contamination. Determination of the age of the groundwater can also indicate where active local recharge is occurring. Age determinations of the groundwater at two different locations along a streamline can be used to give an estimate of groundwater velocity.

Tritium is a radioactive isotope of hydrogen (^{3}H) with a half-life of 12.3 years, which makes it particularly useful for dating modern groundwaters (Clark and Fritz, 1997). The presence of significant amounts of tritium in the

groundwater is evidence for active recharge. Tritium is formed in the atmosphere by high energy cosmic radiation bombarding the earth, the decay of radioactive materials, in nuclear reactors, and from the atmospheric testing of thermonuclear devices. Tritium is deposited on the earth's surface with precipitation and makes its way to the groundwater. The amount of tritium in the atmosphere increased after 1953 due to thermonuclear testing and reached a peak in 1963, after which it started to decline. In areas with small net infiltration and large overburden, the 1963 peak may still occur in the unsaturated or vadose zone and may be reflected in groundwater samples (Dincer *et al.*, 1974). If the 1963 peak is preserved at a known distance from a recharge zone, flow velocity can be estimated, as demonstrated at a waste disposal site in Glouster, Ontario (Michel *et al.*, 1984).

Groundwater can also be dated by measuring the decay product of tritium, ^{3}He. The problem is complicated by the fact that the concentration of ^{3}He normally in the atmosphere that was dissolved in the groundwater during recharge must be estimated and subtracted. In addition a prerequisite for using this approach is a rapid vertical flow of the groundwater such that losses by gas diffusion from the top of the water table are limited. Comparison between the use of this method and the tritium 1963 bomb peak method resulted in dates that differed by only 15% (Schlosser *et al.*, 1989).

Finally two decay products of ^{238}U also offer the possibility of yielding age-related information on ground water samples, in addition to the information available for sediments. Radium 226 has a half life of 1620 years, whereas radon-222 has a half life of 3.8 days. With accurate information on dissolution rates, there is a good possibility of establishing fairly accurate dates over the past 1000–5000 year period (Hillaire-Marcel and Ghaleb, 1997). This may be of significant interest to hydrogeologists and of direct relevance to obtaining information on recharge rates.

Although not related to radioisotopes, chlorofluorocarbons (CFCs) provide another age marker in sediments and groundwater; this approach will be reviewed in Chapter 15.

10.3 USE OF STABLE ISOTOPES TO IDENTIFY CONTAMINANT SOURCES

For many years, as described in other chapters of this book, techniques such as gas chromatography (GC) and gas chromatography-mass spectrometry (GCMS) have been used extensively for the purpose of identifying contaminant sources. Identification in this sense can mean unambiguous identification of each individual compound in the sample or it might mean the ability to correlate a spilled product with its suspected source thus identifying the

source of the spill. The isotopic composition of a specific compound, or the spilled material, will play little, if any, role in the task of actually identifying individual components in the contaminant. The power in the use of stable isotopes is their ability to permit correlations between source samples and spilled samples even if the samples are weathered. This surge in the use of isotopes for correlation purposes has been due in large part to the development of combined gas chromatograph-isotope ratio mass spectrometer (GCIRMS). Before discussing this development, let us consider some of the basic concepts of stable isotope geochemistry.

10.3.1 ORIGIN OF ISOTOPIC CARBON DIFFERENCES

Carbon exists as a mixture of two stable isotopes, ^{12}C and ^{13}C, with the approximate natural abundance of $^{12}C/^{13}C$ ratio being 99:1. Reference notation is explained in Section 10.3.2. The carbon isotopic composition of living organic matter in part depends on the species but also is determined by a number of environmental properties. Atmospheric carbon dioxide is assimilated by living plants during photosynthesis. The nature of the plants and whether they assimilate CO_2 via a C_3 or C_4 photosynthetic cycle determines the extent of preferential assimilation of the lighter ^{12}C isotope. This results in a 5 to 25% depletion in ^{13}C. The amount of fractionation depends on the pathway followed with the so-called C_4 plants typically being associated with warmer and more arid climates and in general having isotopic values in the –10‰ to –18% range. C_3 plants on the other hand are more typically associated with cooler and wetter climates, and generally have lighter isotopic values in the –22% to –30% range. The C_3 pathway operates in about 85% of plant species and dominates in most terrestrial ecosystems. The natural vegetation in temperate and high latitude regions, as well as tropical forests, is almost exclusively C_3. Major crops such as wheat, rye, barley, legumes, cotton, tobacco, tubers (including sugar beets), and fallow grasses are C_3 plants. The C_4 pathway evolved as atmospheric CO_2 concentrations began to drop in the early Tertiary. C_4 plants represent less than 5% of floral species but dominate in hot open ecosystems such as tropical and temperate grasslands and include common crops such as sugar cane, corn, and sorghum. As a point of interest it should be noted that the isotopic composition of food products marked as 100% natural provides a useful tool for customs and excise departments to monitor and check the origin of these and other imports (Hillaire-Marcel, 1986). Desert plants such as cacti operate under yet another mechanism of photosynthesis referred to as the CAM photosynthesis cycle, and can operate under both the C_3 and C_4 cycles. Photosynthesis will have an effect on the composition of the residual CO_2 in the atmosphere and lead to an enrichment of the heavier ^{13}C isotope. The composition of atmospheric CO_2 also varies

depending upon whether it is over land or over ocean and upon surrounding temperatures.

Fossil fuels, coal, crude, and natural gas are formed as a result of a series of very complex and long-term reactions where the living organic matter dies and is deposited in the sedimentary environment. A significant amount of this material will be degraded and recycled as CO_2 and water. However, a small proportion is buried and ultimately converted to fossil fuels of one type or other depending on the nature of the depositional environment. During this formation and generation stage, organic matter from many different sources is deposited in a sedimentary sink and becomes a heterogeneous mixture of organic material. At the same time the unique isotopic signatures associated with specific sources of organic matter are lost as a result of this mixing process. Hydrocarbons generated in this process result in the formation of crude oils or other fossil fuels, which possess isotopic signatures; however, these signatures can no longer be related to specific sources of organic material. From an environmental forensic point of view the fact that it is not possible to relate the isotopic numbers to a specific source of organic material is not critical since the more important application is the ability to use these isotopic values for correlation of the spilled product with its suspected source(s). As will be seen later, isotope values are even more valuable in this application when the spilled product is a single component or a mixture of relatively light hydrocarbons that may not contain any of the components commonly used for correlation purposes following analysis by GCMS. Correlations on the basis of isotopic compositions utilize either the bulk isotopic compositions or the isotopic composition of individual compounds as determined by combined GCIRMS (discussed in more detail later). In complex mixtures, the bulk isotopic composition of the mixture actually will reflect the isotopic compositions of all the components in the mixture and their relative proportions.

So in summary, the carbon isotopic composition of any compound or material is a reflection of the relative proportions of the two stable isotopes and a measure of how depleted the compound or material is in the heavier isotope relative to a standard. The next question, therefore, is how is this depletion measured?

10.3.2 REPORTING ISOTOPE MEASUREMENTS

There are two common methods for the determination of the carbon isotopic composition of a sample. It is possible to measure either the bulk carbon isotope value or the isotopic composition of each individual compound in the mixture. The principle behind the determination of either of these values is

very similar. Each method requires complete combustion of the sample, regardless of its origin, and conversion to CO_2 and water. For bulk determinations, the combustion is undertaken in a sealed tube in the presence of CuO. The CO_2 then is transferred and analyzed in a stable isotope ratio mass spectrometer and the isotopic composition measured relative to that of a standard material (Vienna Pee Dee Belemnite or VPDB) whose isotopic composition has been assigned a value of 0. Virtually all environmental samples will contain less ^{13}C than the standard, and thus being depleted in ^{13}C will have negative $\delta^{13}C$ values.

Stable carbon-isotope ratios ($R = {}^{13}C/{}^{12}C$) are expressed relative to a standard and typically in delta notation, where $\delta^{13}C = (R_{sample}/R_{standard} - 1) \times 1000$ (units are % or per mil or parts per thousand). Although this refers specifically to carbon isotopes, the ratios for the other elements such as O, H, Cl, S, or N are expressed in the same way relative to their specific standard. Figure 10.1 shows the typical isotopic compositions for a range of organic material of different geological ages. Note both the differences between various types of organic material and the relatively narrow spread of values for the bulk isotopic composition of crude oils.

10.3.3 BULK ISOTOPE VALUES

Although the bulk isotope numbers represent weighted averages of all components in a mixture they have still been used successfully in many environmental applications. For example in the case of crude oils, correlations can be made using the bulk isotopic compositions of the saturate and aromatic fractions rather than the whole oil itself. In order to do this the oils are fractionated by a simple column chromatography approach using alumina or silica as the solid support and solvents of varying polarity to separate the required fractions. (It should be noted that in this application, the saturate and aromatic fractions typically are composed of the C_{15+} fraction since the lighter components are lost during fractionation.) It is a very simple application, since the isotopic values for the saturate and aromatic fractions to be correlated are plotted against each other (see Figure 10.2). Samples that are related will plot very close to each other; those that are not related will plot in totally different areas. The use of bulk isotopes in this manner is a preliminary screening tool and simply because two samples plot close to each other does not necessarily mean that they are unequivocally related to each other. Once a preliminary relationship between samples has been established, more definitive correlations should be undertaken using techniques such as GC and GCMS. Correlation of degraded and nondegraded samples is a little more complicated, depending on the degree of degradation. The isotopic

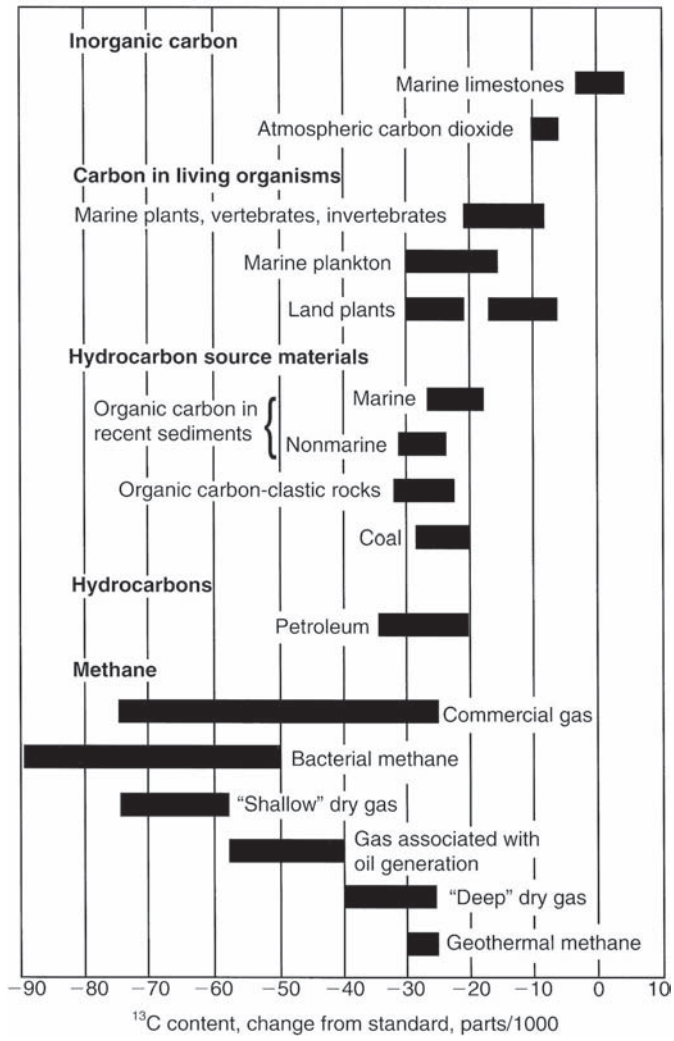


Figure 10.1
Variations in the isotopic composition of different types of organic matter. Note both the differences between various types of organic material and the relatively narrow spread of values for crude oils. (Reprinted from Fuex (1997), with permission from Elsevier Science.)

composition of a sample can change as a result of biodegradation, particularly in the case of crude oils due to the removal of the predominant n-alkanes. In general the saturate fraction of the biodegraded crude oil sample will become isotopically heavier as a result of preferential removal of the isotopically lighter compounds.

It should be noted that although utilization of the bulk isotopic values works well for crude oils and heavier refined products, it becomes a little more difficult for the lighter refined products. Separation of the lighter refined products into saturate and aromatic fractions by column chromatography in the manner just described will lead to loss of significant proportions of the individual fractions due to evaporation. These losses may have an effect on the bulk isotopic composition of the individual fractions, making it difficult

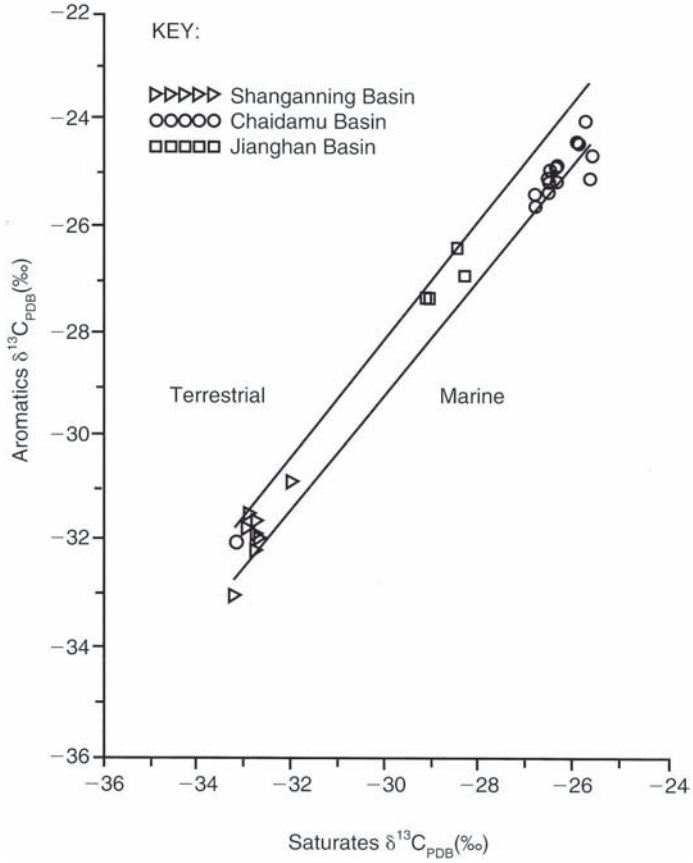


Figure 10.2
A plot of the $\delta^{13}C$ values for the saturates vs. aromatics provides a means of correlating samples of similar origin as shown in this figure using China crude oils from different locations.

to make accurate correlations. Once again, though, the bulk isotopic composition of the nonfractionated samples should be thought of as a preliminary screening tool to determine similarities and differences between selected samples.

A classic example of the use of bulk carbon isotopes in a forensic type study would be the work of Kvenvolden *et al.* (1995). Tar ball residues from the beaches of Prince William Sound were collected several years after the Exxon Valdez accident and characterized on the basis of their bulk carbon isotope ratios. A summary of the bulk isotope numbers for a number of these residues is shown in Table 10.1. The important point to notice here is the presence of two distinct families of tar ball residues. One family corresponds closely to the Exxon Valdez oil on the basis of its isotopic composition, and the second, and isotopically heavier, group can be related to a California crude oil on the basis of its biomarker fingerprints. How did the California crude end up in Prince William Sound? In retrospect it is very simple—prior to the discovery of crude oil in Alaska it had to be imported from California.

Table 10.1

Examples of bulk isotope numbers for a number of residues collected from Prince William Sound several years after the Exxon Valdez (EV) accident indicated the presence of two distinct families of tar ball residues. GCMS was used to confirm that one family was from Exxon Valdez crude whereas the other was from California crude imported to Alaska many years earlier (modified from Kvenvolden et al.*, 1995).*

Location	$\delta^{13}C$ (‰)	Description
Bligh Reef (EV Oil)	−29.2	North Slope Crude EV tanker
EV Residues		
Naked I.	−29.4	Oiled cobbles, McPherson Bay
Eleanor I.	−29.1	Solid oil on rock, Northwest Bay
Evans I.	−29.3	Solid oil mat, Shelter Bay
Non-EV Residues		
Old Valdez	−23.6	Asphalt mats, old asphalt plants
Naked I.	−23.9	Tar on metashale, Bass Harbor
Evans I.	−23.7	Tar on rock, Shelter Bay
Other		
Latouche I.	−24.8	Tar in mine dump
Knight I.	−25.3	Solid oil on cobble, Herring Bay
Pavements		
Whittier	−24.8	Street pavement > 20 years old
Valdez	−27.2	New airport runway
Seward	−28.1	New airport runway
Cold Bay	−22.7	End of old airport runway
Amchitka	−23.4	World War II C airport runway

Much of this oil was temporarily stored in tanks on the edge of Prince William Sound. In 1964 a massive earthquake struck the region of Valdez, the tanks ruptured, oil spilled into the Sound and in view of the magnitude of the disaster, the oil was left in the Sound to degrade naturally with very little cleanup. Crude oils contain many compounds that are relatively resistant to biodegradation and hence 40 years later they can still be detected in these tar residues. However, the point of this example is to demonstrate that it was the bulk carbon isotopes that originally permitted the distinction to be made between the two groups of oils. Furthermore, the same approach was extended to tar residues on buildings, airport runways, and other locations showing that most of the bitumens used there prior to the discovery of the Alaskan reserves could be correlated with an origin from California crudes on the basis of their isotopic compositions.

Although most emphasis has been placed on the use of bulk carbon isotopes for correlation purposes, there are a number of forensic applications that have used other stable isotopes. For example Kaplan *et al.* (1997) showed that a plot of deuterium against carbon isotope values for groundwater extracts and suspected gasoline sources was a viable correlation tool (see Figure 10.3). Determination of δD values is a more time consuming process than determination of δC values and, to date, has not been used so extensively in environmental forensic studies. However, the capability to determine δD values for individual compounds could change this situation dramatically (Sessions

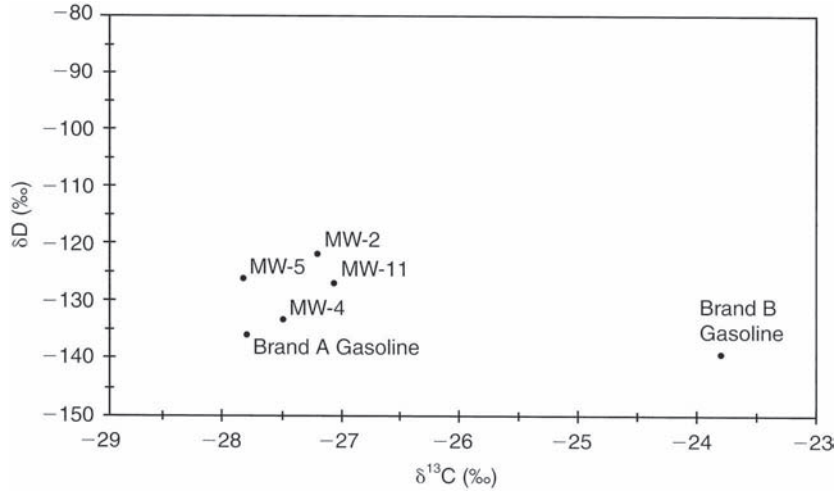


Figure 10.3
Kaplan et al. *(1997) showed, using this figure, that a plot of* $\delta^{13}C$ *vs.* δD *values for groundwater extracts and suspected gasoline sources was a viable correlation tool and could distinguish gasoline samples from different sources. MW, monitoring well. (Reprinted from Kaplan* et al. *(1997), with permission from Elsevier Science.)*

et al., 1999; Li *et al.*, 2001; Philp *et al.*, 2002; Pond *et al.*, 2002; Wang *et al.*, 2004). In addition, problems such as the possibility of H exchange reactions with the groundwater have to be minimized or eliminated to prevent the possibility of erroneous values being obtained for the H/D ratio.

10.3.4 CARBON ISOTOPE GCIRMS FOR HYDROCARBONS

Bulk isotopic values have been readily available for many years but one of the most significant analytical advances in geochemistry in the past few years has undoubtedly been the development of combined gas chromatography-isotope ratio mass spectrometry (GCIRMS). This technique permits continuous-flow acquisition of $\delta^{13}C$ values for individual components in complex mixtures in real time and without the need to physically isolate each individual compound (Freeman *et al.*, 1990; Hayes *et al.*, 1990; Mansuy *et al.*, 1997; Abrajano and Sherwood Lollar, 1999). Chemical fingerprinting of the n-alkane fraction in crude oils and refined products in combination with isotopic characterization of carbon in the individual homologues has been used successfully to allocate sources of sediment contamination (Rogers *et al.*, 1999; Smallwood *et al.*, 2002) and bird-feather oiling (Mansuy *et al.*, 1997; Mazeas and Budzinski, 2002a). Hough *et al.* (2006) used the $\delta^{13}C$ values of individual n-alkanes to identify the source of hydrocarbons in soils contaminated with residual heavy or weathered petroleum wastes. A comprehensive review of the subject was presented by Meier-Augenstein (1999) and Schmidt *et al.* (2004), and specific details of the actual technique will not be presented again in this paper.

The most recent development in the area of the combined GCIRMS is the capability to determine the H/D ratio of individual compounds as well as C

isotopes of individual compounds. Indeed, Pond *et al.* (2002) proposed the use of hydrogen isotopic composition of longer chain n-alkanes (n-C_{19} to n-C_{27}) for source discrimination since the hydrogen isotopic signature of crude oil components vary more significantly than the carbon values and do not show discernable changes during weathering or degradation of crude oils. There have been a number of applications in the petroleum geochemistry literature describing the use of both the C and H isotope values of crude oils and saturate hydrocarbon fractions for undertaking oil/oil and oil/source rock correlation studies (Peters and Fowler, 2002; Philp *et al.*, 2002; Milkov, 2005). However, the number of applications in the environmental field, particularly for H/D determinations, has been fairly limited and most of them have been simply designed to show applicability of the technique (Hilkert *et al.*, 1999; Sessions *et al.*, 1999; Wiesenberg *et al.*, 2004). However, the combination of the carbon and hydrogen isotope measurements of individual compounds, determined by GCIRMS, will provide a powerful tool in the future to discriminate compounds from different sources as it has been in the petroleum exploration field. A publication by Sherwood Lollar *et al.* (2001) demonstrated that biodegradation of toluene produced a significant change in the H/D isotopic composition of the toluene. This has led to a number of studies where changes in both carbon and hydrogen isotope compositions of groundwater contaminants such as MTBE and PCE have been used as a tool to indicate the onset of natural attenuation (Sherwood Lollar *et al.*, 2001; Slater *et al.*, 2001; Gray *et al.*, 2002; Kolhatkar *et al.*, 2002; Griebler *et al.*, 2004; Kuder *et al.*, 2005). Utilization of changes in the isotopic compositions in this manner has attracted the attention of the regulators and this approach is now being utilized at many contaminated sites to monitor progress of natural attenuation in conjunction with concentration changes.

The basic concept for the determination of the isotopic composition of individual compounds is the same as for the bulk isotopic values in that the components are completely combusted to CO_2 and water and the isotopic composition of the resulting CO_2 determined. However, the big difference is that these values are determined in real time as the individual compounds elute from the GC column. The separated compounds pass through a combustion tube where they are combusted, the CO_2 and water pass through a separator to remove the water, and the CO_2 continues into the mass spectrometer (see Figure 10.4). With the GCIRMS system, the combustion has to be complete in the time it takes for the sample to pass through the reactor, which is typically 12 to 15 secs. However, the preliminary output from the analysis is a m/z 44 chromatogram, which appears to be very similar to a conventional GC trace but actually depicts the CO_2 produced from the various components in the mixture (see Figure 10.5). It should be noted that determination of the

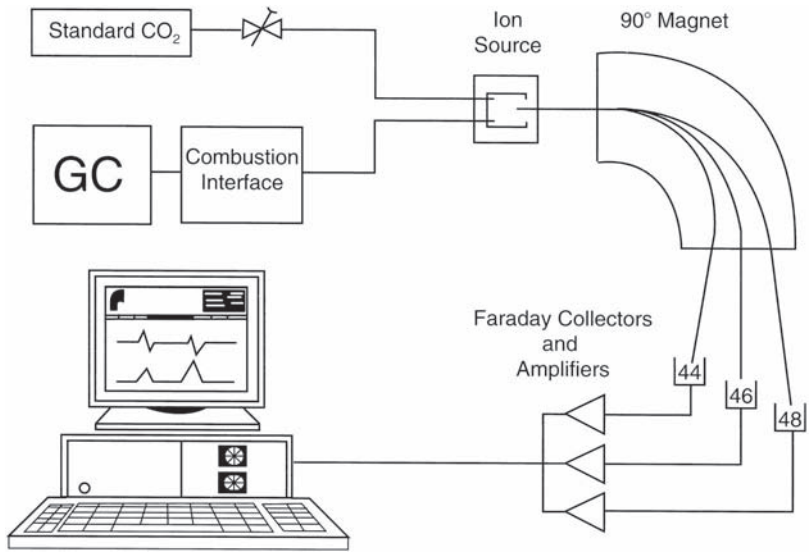


Figure 10.4

In a GCIRMS system, separated compounds elute from the GC column and pass through a combustion tube; the CO_2 and water generated in this manner pass through a separator to remove the water, and the CO_2 continues into the mass spectrometer, where the $\delta^{13}C$ values are determined relative to the standard Pee Dee Belemnite.

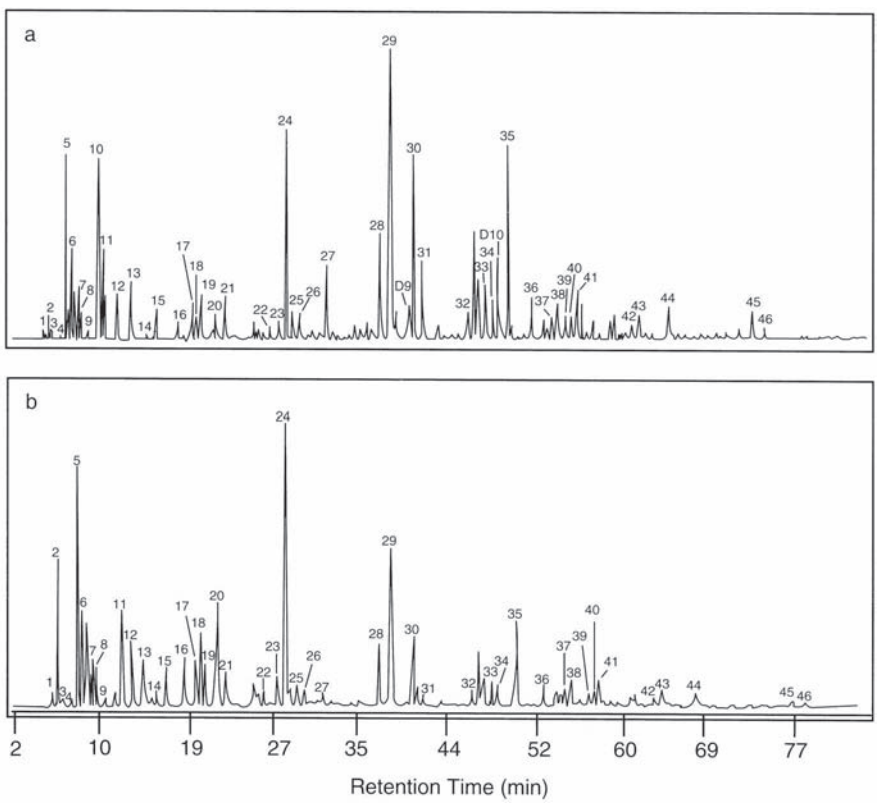


Figure 10.5

Chromatograms for two different gasoline samples. The initial output from the GCIRMS analysis is a m/z 44 chromatogram that appears to be very similar to a conventional GC trace but actually depicts the CO_2 produced from the various components in the mixture. Data for each individual component are then converted into $\delta^{13}C$ values relative to the PDB standard.

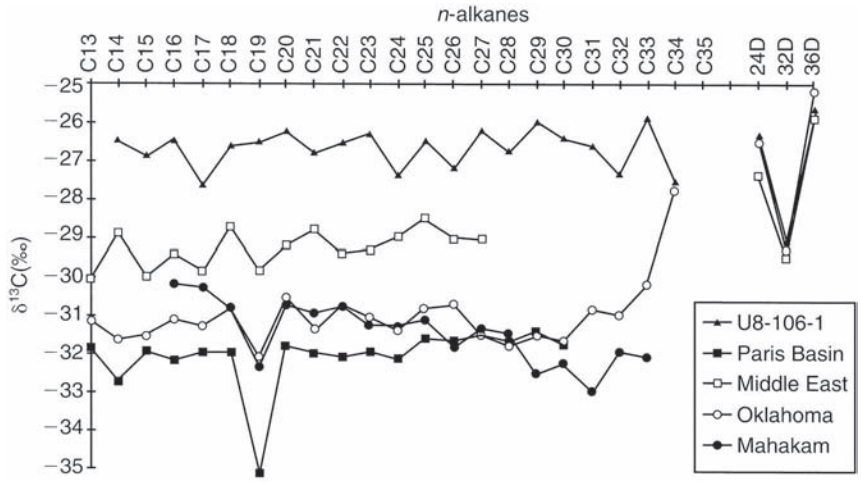


Figure 10.6
Characterization of a number of oils from different geographical regions by GCIRMS showed that this technique was a viable approach for differentiating these oils. Again this information should be used in conjunction with the GC and GCMS data. (Reprinted with permission from Mansuy et al. *(1997), American Chemical Society.)*

H/D ratios uses an interface held at a much higher temperature, 1400˚C, than used for carbon determination, and basically each compound is pyrolyzed rather than oxidized and the resulting D and H enters the mass spectrometer where the ratio of mass 2 to 1 is measured.

Data for each individual component are then converted into δ^{13}C values relative to the PDB standard. It should be noted that although it sounds simple, there are a number of problems attached to the technique, particularly that of coelution. Any coelution between two peaks can have a significant effect on the isotopic composition of the other component. This requires manual manipulation and processing of the data in order to minimize and correct for this interference. Despite all these problems, the data obtained from this approach can be very useful for correlation purposes, as will be demonstrated later. To illustrate the nature of the results obtained in this way, Figure 10.6 shows the results from the analyses of several oils from various geographic origins. The isotopic compositions for each of the major components have been determined and are plotted in this diagram. Differences between these samples are immediately apparent and reflect their different origins. In a second example, results from the analyses of four gasoline samples are shown in Figure 10.7.

These results show that the four gasolines can readily be divided into two groups based on the isotopic compositions of individual components in the gasolines. GCIRMS is particularly important in the case of refined products such as gasoline, which although derived from different feedstocks may be very similar chromatographically and difficult to discriminate using conventional techniques such as GC and GCMS. In early environmental applications of GCIRMS, Kelley *et al.* (1995, 1997) measured the concentrations and

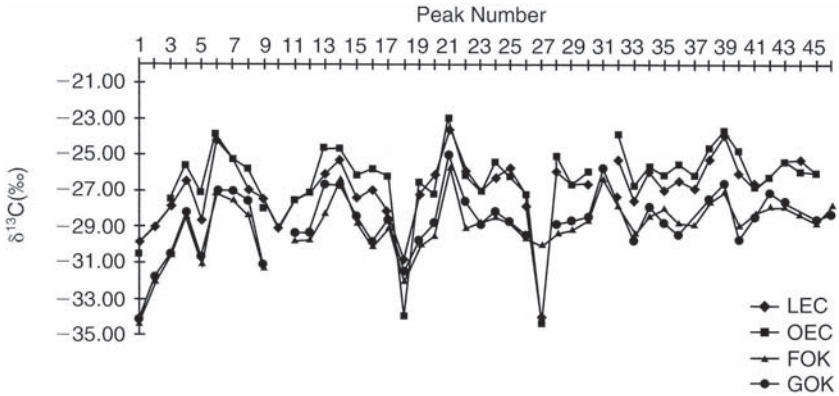


Figure 10.7
A similar example to Figure 10.6 but instead of crude oils, GCIRMS has been used to differentiate gasolines from Oklahoma (FOK, GOK) and the East Coast (LEC, OEC). (Reprinted with permission from Mansuy et al. *(1997), American Chemical Society.)*

isotopic compositions of BTEX (benzene, toluene, ethylbenzene, xylene) compounds collected from a gasoline contaminated site in Southern California. From the results it was evident that there were two isotopically distinct families of BTEX compounds that permitted the conclusion to be drawn that the leaded and unleaded gasolines at this site were derived from different sources.

Oils that have been artificially weathered in the laboratory have been analyzed by GC and GCIRMS as shown in Figure 10.8. With the exception of the most extensively biodegraded sample, all these samples can be correlated on the basis of the isotopic composition of the individual n-alkanes (see Figure 10.9). This approach works for samples that are weathered as a result of evaporation, water-washing, or biodegradation. For the most part, use of GCIRMS requires well-resolved components to eliminate the possibility of coelution, but extensively biodegraded samples can be correlated by isolating and pyrolysing the asphaltenes and analyzing the pyrolysis products of the asphaltenes by GCIRMS (Mansuy *et al.*, 1997).

Birds are often victims of oil spills and wash ashore covered in oil. In certain situations, the source of the oil is unknown, but a number of suspected sources may be available. To determine possible origins for the spilled oil on the birds, oil-covered bird feathers can be extracted, analyzed by GCIRMS, and the data compared to the suspected sources. In the example used for Figure 10.10a and 10.10b, the oil and bird feather extracts had characteristics of a light fuel oil marked by the absence of biomarkers. The oil collected from the bird feathers was weathered and light hydrocarbons had been lost by evaporation and water-washing. The suspect oil was analyzed by GCIRMS without any fractionation step prior to the analysis. Comparison of the isotopic values for the individual components in the two samples suggested the suspect oil as the possible source for the oil on the bird feathers (see Figure 10.10c). The

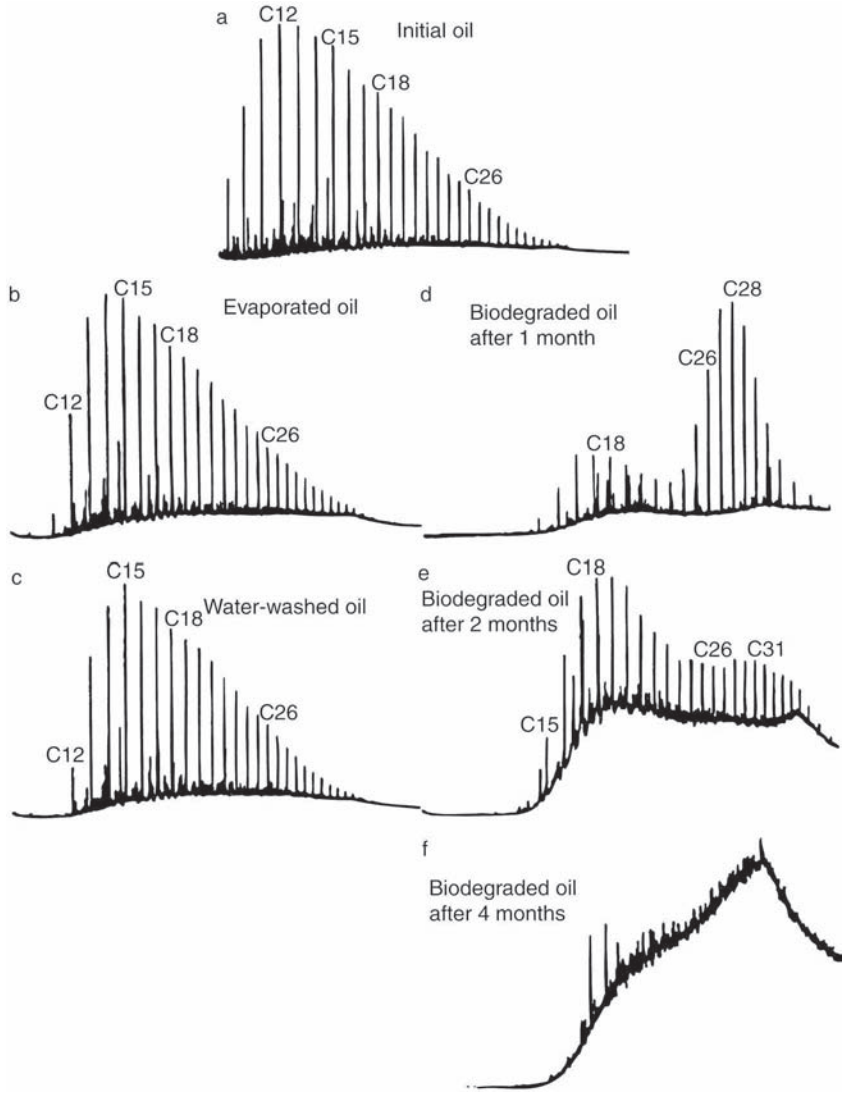


Figure 10.8
Artificial weathering of crude oils in the laboratory were analysed by GC and GCIRMS. The resulting GC traces are shown in this figure and it can be seen that even after a fairly short period of time biodegradation was quite extensive. (Reprinted with permission from Mansuy et al. *(1997), American Chemical Society.)*

standard deviation calculated between the two samples (0.36%) is close enough to the range of the reproducibility (0.21–0.26%) to consider that these two light fuel oils are derived from the same source. These results provided a high level of confidence, suggesting similar origins for the oils found on the bird feathers and the suspected sources. The correlations could be supported by the conventional GC and GCMS fingerprinting but the absence of the conventional biomarkers make this a more difficult task.

Determination of the origin of polycyclic aromatic hydrocarbons (PAHs) provides another challenging environmental problem. PAHs in the environment may be derived from sources such as coal tar, crude oils, or creosotes

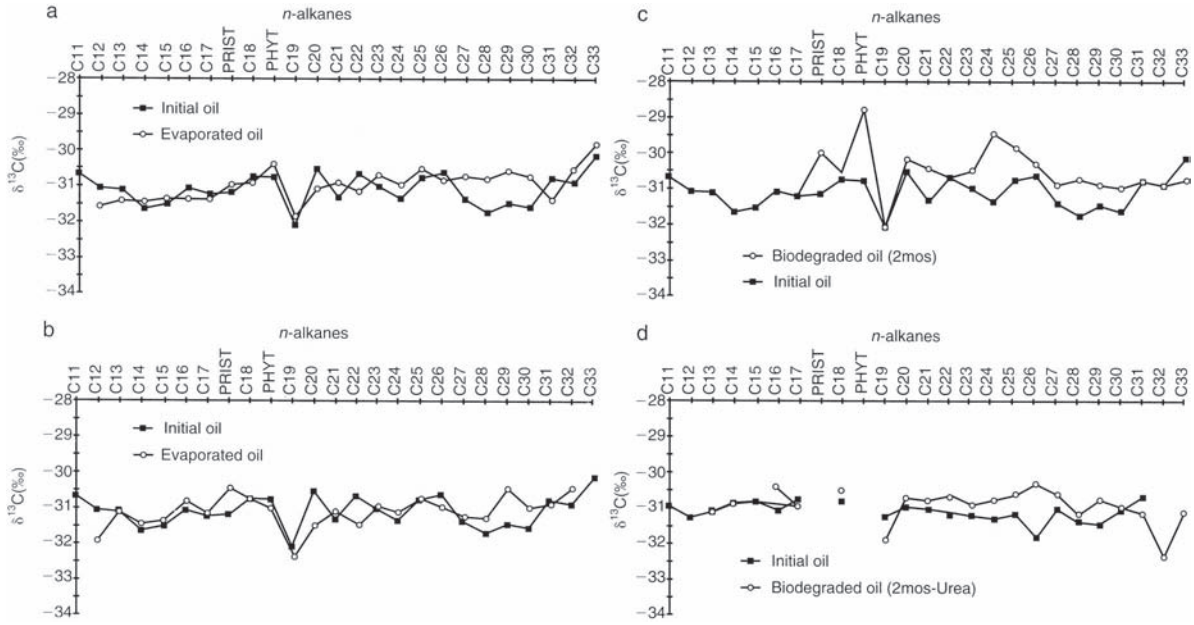


Figure 10.9

Characterization of the extensively biodegraded oils in Figure 10.8 by GCIRMS has shown that, with the exception of the most extensively biodegraded sample, all of these samples can be correlated on the basis of the isotopic composition of the individual n-alkanes. This approach works for samples that have been weathered as a result of evaporation and water-washing, not only biodegradation. (Reprinted with permission from Mansuy et al. *(1997), American Chemical Society.)*

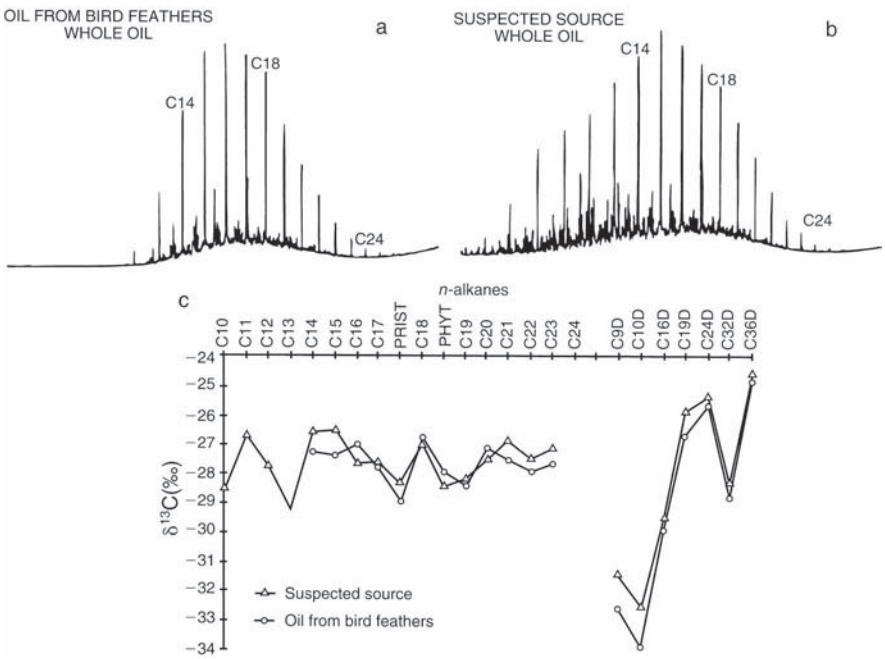


Figure 10.10

***a**, **b**. Bird feathers coated with a light fuel oil were not readily correlated by conventional techniques such as GC and GCMS due to the absence of the biomarkers. **c**. GCIRMS successfully showed a relationship between the samples. (Reprinted with permission from Mansuy* et al. *(1997), American Chemical Society.)*

for example. In many cases the chromatographic distributions of these compounds for different sources may be very similar and be of little use in terms of distinguishing one source from another. In addition, weathering can often result in additional ambiguities. However, with the development of GCIRMS, the possibility of determining the isotopic composition of the individual components provides a potential method for differentiating these sources of the PAHs. On a regional scale, source apportionment of polycyclic aromatic hydrocarbons (PAHs) both in the atmosphere and in sediment records has been studied intensely using δ^{13}C analysis. Interestingly, δD analysis of individual PAHs has not been reported to date, although this is partly due to the fact that PAH mixtures are very complex and if good separation is not obtained between individual compounds it becomes exceedingly difficult to get good reproducible isotope values particularly for hydrogen. A combination of concentration measurements and δ^{13}C isotopic analysis of individual PAHs in sediments from Lake Erie permitted three areas with different contamination histories to be distinguished. Furthermore, it could be shown that the main emission pathway for PAHs was fluvial input (Smirnov *et al.*, 1998). A significant study to evaluate the use of stable isotopes to differentiate PAHs derived from urban backgrounds versus those derived from coal gasification plants has been recently completed (Saber *et al.*, 2003, 2005; Craig *et al.*, 2005; Mauro *et al.*, 2005). The results of these works demonstrated the added value of the isotopic data to such a study as illustrated in Figures 10.11 and 10.12.

First, PAH samples that may appear chromatographically similar may be differentiated on the basis of their isotopic signatures, and similarly those that are isotopically different may be chromatographically similar. Further, the isotopic signatures of samples derived from manufactured gas plants are isotopically distinct from those typically derived from the urban background (see Figure 10.13). Hammer *et al.* (1998) studied the PAHs from two creosote-contaminated sites and showed that at these two sites there was a suite of compounds that had very similar isotopic values, which agreed within 1% of each other and which may be useful for identification of PAHs of creosote origin. In a recent study, various sources of PAHs were distinguished in sediments along the St. Lawrence River and, relatively heavy δ^{13}C values were found for three-ring PAHs originating from aluminum smelting in one area (Stark *et al.*, 2003). McRae *et al.* (1999, 2000) showed it was possible to relate coal-derived PAHs released during different thermal conversion processes (combustion, pyrolysis, gasification) with their isotopic compositions conserved in soil. Extremely isotopically light PAHs (δ^{13}C from −31 to −62%) in lagoon sediments near Ravenna (Italy) led to the conclusion that emissions were dominated by a former plant that used biogenic methane (δ^{13}C from −69 to −73%) as feedstock rather than by operating plants using petrogenic

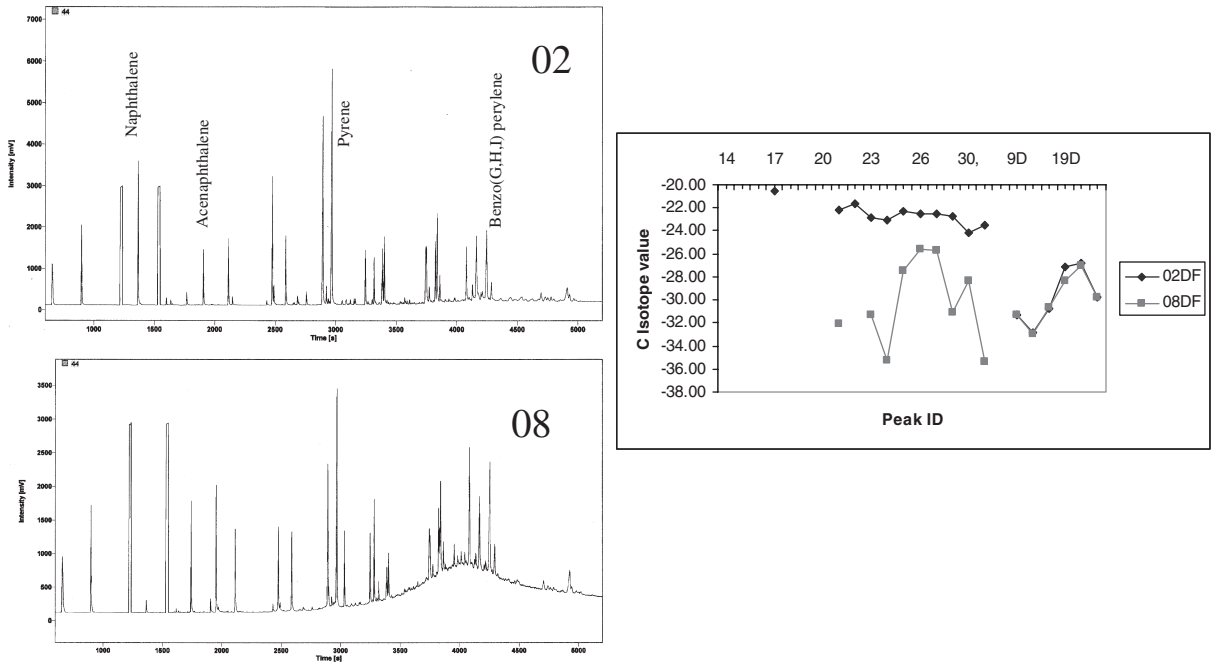


Figure 10.11

PAH samples with similar chromatograms may be from different sources but can be differentiated on the basis of their isotopic compositions as illustrated above. (In the isotope plots for Figs. 10.11 and 10.12 the numbers on the X axis correspond to compound numbers and 9D, 16D, etc to deuterited standards.)

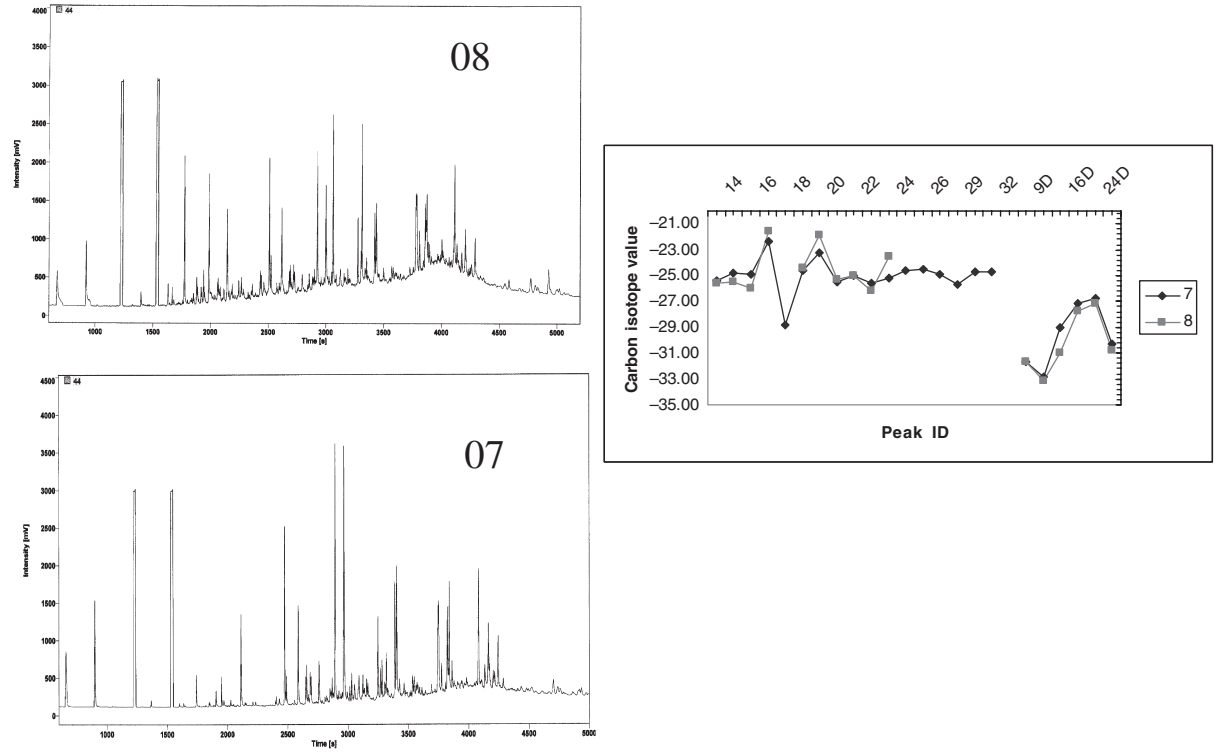


Figure 10.12

PAH samples with chromatograms may be from similar sources as shown by their isotopic compositions.

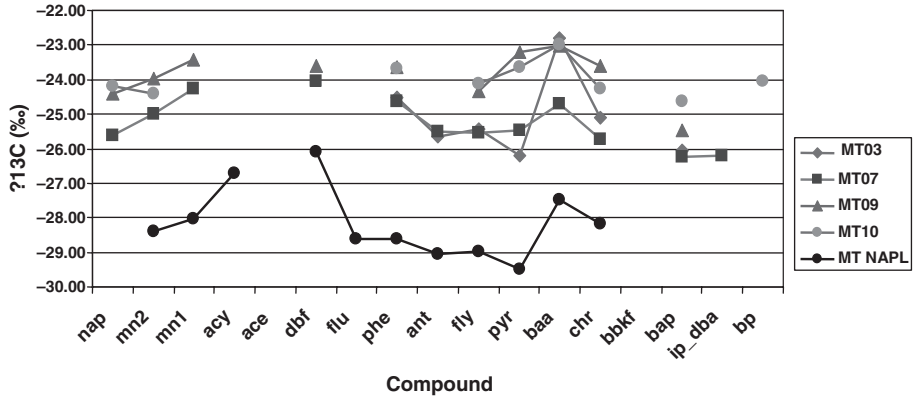


Figure 10.13

The carbon isotope values of individual PAHs can be used to differentiate those coming from manufactured gas plants (MGP) vs. those simply derived from the urban background. On this diagram MT NAPC is a coal tar sample and other samples are soil extracts.

feedstocks that typically are isotopically heavier (McRae *et al.*, 2000). Wilcke *et al.* (2002) used $\delta^{13}C$ analysis of perylene to substantiate their earlier hypothesis that in tropical environments recent biological sources of PAHs related to termites are important, whereas pyrolytic sources dominate in temperate climate. PAHs in atmospheric particles resulting from natural burning processes could be distinguished from those stemming from various anthropogenic combustion processes by using fingerprinting and $\delta^{13}C$ analysis of individual compounds (Norman *et al.*, 1999; Okuda *et al.*, 2002a). In Chinese urban areas, PAH $\delta^{13}C$ analysis was used successfully to identify either vehicle exhaust or coal combustion as major PAH source (Okuda *et al.*, 2002b). In contrast, PAH fingerprinting did not yield equivalent information. Many of these studies show the necessity to combine chemical fingerprinting techniques and compound-specific isotope analysis.

Mead *et al.* (2005) demonstrated that the isotopic ratios of $\delta^{13}C$ on *n*-alkanes from sediments in the Everglades, Florida, allowed them to distinguish between two natural sources of higher plants: terrestrial and aquatic. Shi *et al.* (2001) showed that the isotopic ratios of fatty acids and sterols can be used to differentiate aquatic sources from terrestrial sources of organic matter in estuaries.

10.3.5 SOURCES AND SINKS OF ATMOSPHERIC GASES

Isotopic measurements can be used as a tool to characterize various sources and sinks of atmospheric gases and estimate their relative importance. This approach has been used since the early 1980s, but nowadays the availability of the combined GCIRMS permits a much higher sampling frequency (Rice *et al.*, 2001). A detailed comparison between dual-inlet IRMS and GCIRMS for measuring $\delta^{13}C$ and $\delta^{2}H$ of methane was undertaken by Rice *et al.* (2001), who found no systematic deviations between the methods. On a Pacific

transect $\delta^{13}C$ values ranging from −47.0‰ to −47.3‰ and $\delta^{2}H$ values of −85.8 ± 1.2 and −91.7 ± 2.0‰ were observed for the southern and northern hemispheres, respectively, indicating that $\delta^{2}H$ analysis was more sensitive to spatial variations than $\delta^{13}C$ analysis. For chloromethane, Thompson *et al.* (2002) found a global $\delta^{13}C$ average of −36.2‰. Budgeting of atmospheric chloromethane based on isotopic composition of known sources suggested the existence of additional emission sources with an average $\delta^{13}C$ value of −41.9 ± 7.8‰. There are several natural sources of chloromethane, and isotopic composition can clearly play an important role in distinguishing natural sources from anthropogenic sources (Gribble, 1998, 2003). Chloromethane emitted by polypore fungi was found to be highly depleted in $\delta^{13}C$ (−43.3‰) compared to the growth substrate and even more depleted in two higher plant species (Harper *et al.*, 2001). Chloromethane produced by two tropical fern species was found to be significantly depleted in ^{13}C ($\delta^{13}C$ −69.3‰ and −72.7‰, respectively) compared to chloromethane that resulted from biomass burning and industrial emissions (Harper *et al.*, 2003). The capability of discerning sources of dichloromethane and other volatile atmospheric components will play an important role in evaluating the contributions of atmospheric contaminants and their potential role, if any, in future climate change and global warming or global cooling.

10.3.6 USE OF GCIRMS FOR MTBE

Methyl tert-butyl ether (MTBE) is rapidly becoming a major environmental problem but one that may be ideally suited to study using carbon and hydrogen isotope compositions. The recent controversy over addition of MTBE to gasoline to boost oxygen content and decrease carbon monoxide emissions to the atmosphere led to a proposed phase-out of this compound by 2002, although this has not happened on a large scale nationwide. MTBE has been blended with some gasolines at low levels (2%) to improve octane performance. It was known to be a carcinogen since 1979 but was thought to be less of a health risk than many of the other constituents of gasoline (Hanson, 1999). MTBE has become a major contaminant of water (i.e., when gasoline leaks from underground storage tanks (UST) and makes contact with groundwater) and, to a lesser extent, air (i.e., when dispensed at a gasoline pump, although it is washed out relatively rapidly following the first rain). Furthermore in air, MTBE is relatively rapidly photodegraded. Examination of MTBE-contaminated groundwater samples show that in many cases they are dominated by MTBE with only minor concentrations of other components present in the chromatograms (see Figure 10.14). Determining the precise origin of MTBE can be an extremely difficult question since many sites may

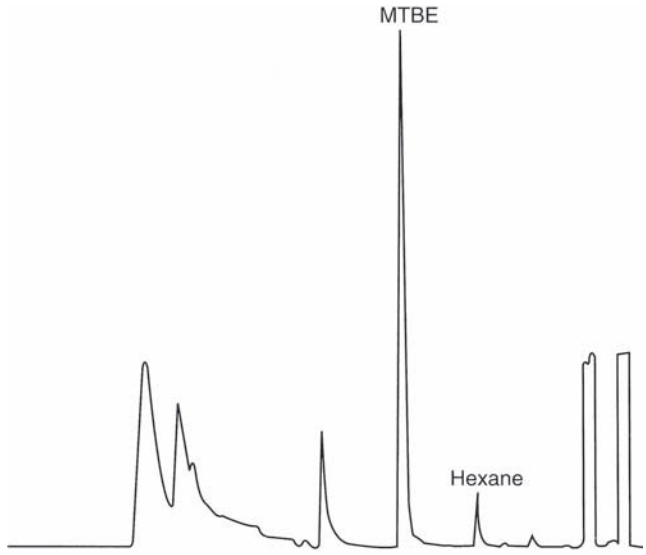


Figure 10.14
Examination of MTBE-contaminated groundwater samples shows that in many cases they are dominated by MTBE with only minor concentrations of other components present in the chromatograms.

have multiple sources of MTBE from different gas stations in the area. Early studies indicated the recalcitrant nature of MTBE provided a potential opportunity for the source/spill correlation of gasoline or MTBE using isotope-ratio mass spectrometry (Smallwood *et al.*, 2001). Although not the ultimate solution, GCIRMS is probably the only viable technique that has any chance to discriminate various sources of MTBE in the groundwater samples.

Since these samples typically are dominated by MTBE, GC and GCMS will be of little use in correlating or distinguishing samples of different origins. It was shown that a combination of carbon and hydrogen isotope values GCIRMS can be used to distinguish MTBE from different sources, and these values do not appear to be affected by evaporation or water-washing (Smallwood *et al.*, 2001). Four pure MTBE samples showed little variation in their $\delta^{13}C$ values but could be discriminated on the basis of their δD values (Sessions *et al.*, 1999; see Figure 10.15). When a larger number of gasoline-containing MTBE samples were examined, a much greater variation in the carbon isotope compositions was observed, as a result of the MTBE coming from different suppliers (Smallwood *et al.*, 2001).

The carbon isotope composition of MTBE can be determined reproducibly at concentrations in the low ppb levels (Kolhatkar *et al.*, 2002; Zwank *et al.*, 2003) using a purge and trap (P&T) system in conjunction with the GCIRMS technique. P&T GC-IRMS is a reliable technique for these kinds of compounds in water, which is very significant since concentrations of 15–20 ppb of MTBE in drinking water is the threshold for taste and odor. This is also the limit set by most states within the United States, for the maximum

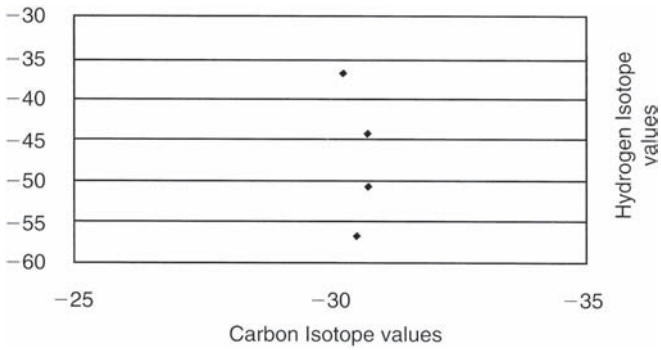


Figure 10.15
Four pure MTBE samples showed little variation in their $\delta^{13}C$ values but these samples could be discriminated on the basis of their δ-D values.

concentration of MTBE allowed before remediation is enforced or drinking water wells are shut down.

However, as will be described later, the major interest in utilization of isotopes in the study of MTBE has moved away from source determination to the use of the isotopes as a tool for evaluating the onset of natural attenuation. Remediation of contaminated sites is always an expensive proposition but natural attenuation can greatly reduce clean-up costs. A combination of concentration changes and isotopic enrichment can be used as evidence for the onset of natural attenuation (Kuder *et al.*, 2005).

In summary, these results have provided a number of illustrations concerned with the application of both bulk isotopic compositions and isotopic compositions of individual compounds for the purposes of correlating and identifying samples of various origins. It is essential to stress that although these data are extremely useful, particularly in undertaking correlations that may not be possible by other approaches, the approach will be most powerful when used in conjunction with currently existing techniques, such as GC and GCMS. Isotope values can provide that additional piece of information that may be used to confirm conclusions based on these other data sets. In the situation where the contaminant is a single component, determination of the isotopic composition, preferably carbon and hydrogen, is probably the only approach with any chance of providing a successful answer.

10.4 OTHER ISOTOPES

Lead was added to gasolines in the United States for many decades in the form of tetraethyl lead (TEL) and other forms of alkylate derivatives. In recent times, health-related effects of lead in gasoline became a major issue. The concentrations of lead were greatly reduced in the early 1980s, and the use of TEL was phased out and completely eliminated in all U.S. gasolines by 1996. If the lead content of a groundwater sample containing gasoline is measured

and found to be nonexistent or negligible, the normal conclusion is that the sample must have been an unleaded gasoline. If lead is present then typically the assumption is made that the spill must have been of a leaded gasoline or a mixture of leaded and unleaded gasolines. A number of combustion methods can be pursued to measure lead content or alternatively GC methods that are available to determine the distribution of alkyl lead derivatives in the extracts. An alternative approach is to measure the lead isotope content of the extract from the groundwater. There are two reasons for doing this; first the ability to measure the lead isotope ratio will simply indicate the presence of lead in the sample; second is the possibility of obtaining an approximate age for the sample release. There are four naturally occurring lead isotopes: ^{204}Pb is the stable isotope and ^{206}Pb, ^{207}Pb, and ^{208}Pb are stable radiogenic isotopes (though the latter three are actually radiogenic, they have extremely long half-lives and can be thought of as being stable in the time frame under consideration in the current context). The lead isotope ratios are measured relative to ^{204}Pb or ^{206}Pb and similar isotope ratios should imply a common source (Hurst *et al.*, 1996). The age dating application here is not directly related to the decomposition of any of these isotopes but rather due to the fact that the lead isotope ratios reflect the source of the lead used in the preparation of the TEL. Furthermore there were basically only two major manufacturers of TEL from 1960 through 1990, namely DuPont and Ethyl Corporation. In a plot of δ^{206}Pb (^{206}Pb/^{207}Pb) in gasoline vs. time it can be seen that the isotopic value for the lead becomes progressively heavier as it gets closer to the present time (see Figure 10.16). However, the number of sources of lead was somewhat

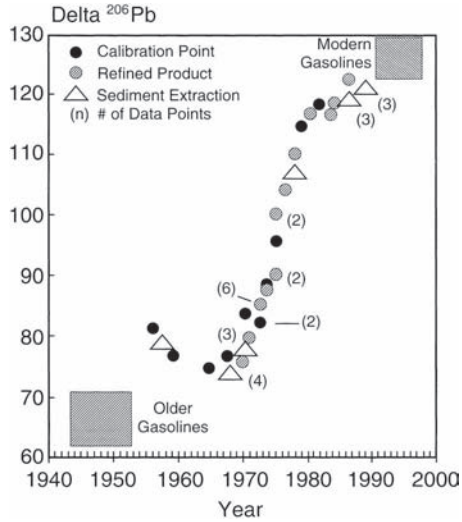


Figure 10.16

ALAS model calibration curve. In a plot of δ^{206}Pb in gasoline vs. time it can be seen that the isotopic value for the lead becomes progressively heavier as it gets closer to the present. (Reprinted with permission from Hurst et al. *(1996).)*

limited and divided between Mississippi Valley (MV) type ores and imported ore, largely from Australia.

The $^{206}Pb/^{207}Pb$ ratio from the former was approximately 1.30, whereas for imported ore, the $^{206}Pb/^{207}Pb$ ratio was less than 1.1. The plot of the $^{206}Pb/^{207}Pb$ against time shows a steady increase with time for lead in gasolines from a number of different sources (Hurst *et al.*, 1996). From this plot came the Anthropogenic (i.e., gasoline-derived) Lead Archeostratigraphy Model (ALAS) to distinguish between multiple sources of gasoline. The model plots one lead isotope ratio against the other or one isotope ratio against concentration and it is proposed by Hurst that patterns emerge that can distinguish multiple sources of gasoline plus time of formulation within a five-year time period (Hurst, 2000). If correct, the lack of scatter within the data over a 40-year period clearly demonstrates the potential viability of using this approach to provide an approximate indication of the age or date of manufacture of the TEL added to the gasoline. However, the gradual shift in isotopic ratios for individual spills is puzzling since this implies lead from Mississippi Valley was being mixed with lead from Australia or elsewhere, prior to making the tetraethyl lead and the relative proportions were changing systematically with time. In all probability as with any commodity, lead was being purchased on the spot market at the best possible price and the variations in the amounts of lead being used from different sources probably varied significantly from month to month. Second, if we assume that MV lead has an isotopic value of –1.32 and the other source is –1.04, on the basis of the ALAS model it can be computed that MV lead constituted 46% in 1965 and 64% in 1987, where Hurst claims that over the same period MV ore use rose from 9% to 82%. Other problems include the potential for additional sources, such as lead-based paints or pipeline contributions to the groundwater; this could be significant and therefore one must exercise some caution while applying this method at the current level of knowledge. The problem is also exacerbated by the fact that mixing of gasolines of several vintages at the same site could greatly distort the age obtained on the basis of the lead isotope ratio. It may be possible to unravel the situation but this could be a complex problem if end-members (pure samples of the individual components that comprise the mixture) are unavailable.

Though not related to the ALAS model just described, Vilomet *et al.* (2003) provided evidence of the capability of stable lead isotopes to trace landfill leachate in a shallow groundwater. Stable lead isotopes were used to characterize a landfill leachate signature ($^{206}Pb/^{207}Pb = 1.189 \pm 0.004$), which was shown to be clearly different from that of the local atmospheric ($^{206}Pb/^{207}Pb = 1.150 \pm 0.006$) and crustal lead ($^{206}Pb/^{207}Pb = 1.200 \pm 0.005$). The results showed that the complexity of pollutant plume dispersion in the shallow groundwater

system requires several independent tracers to clearly resolve the origin and transport pathways for contaminants.

As far as can be ascertained, the bulk of isotope studies to date that are concerned with environmental problems have been based on carbon and hydrogen isotope values. O, N, S, and Cl isotopes are starting to see more widespread use in specific areas. For example, a study of chlorinated solvents used a plot of $^{13}C/^{12}C$ against $^{37}Cl/^{35}Cl$ to illustrate a relationship between isotopic content and manufacturer for a variety of solvent types such as PCE, TCE, and TCA (Beneteau *et al.*, 1999). Other studies on chlorinated compounds have demonstrated the combined use of $\delta^{37}Cl$ and $\delta^{13}C$ to differentiate sources of chlorinated species (Warmerdarm *et al.*, 1995; Drenzek *et al.*, 2002; Shouakar-Stash *et al.*, 2003; Sturchio *et al.*, 2003). The specific work on chlorinated compounds will be presented in detail later in this chapter (see Section 10.5).

Sulfur has four naturally occurring isotopes, ^{32}S, ^{33}S, ^{34}S, and ^{36}S. The ^{32}S and ^{34}S isotopes are most commonly used for isotope ratio analysis. Even though reports have been presented on the use of sulfur isotopes to differentiate anthropogenic and natural sources of sulfur dioxide, it would not appear that the sulfur isotopes have been used extensively for the purposes of identifying contaminants in the environment. One recent application published by Becker and Hirner (1998) examined the combined use of bulk carbon and sulfur isotopes as a means of discriminating crude oils from various sources. It was noted that since many carbon isotope values are very similar, the additional use of the sulfur values permits separation of oils from different sources.

Coffin *et al.* (2001) evaluated the $\delta^{13}C$ and $\delta^{15}N$ isotope analyses in field and laboratory experiments to trace the sources of TNT in contaminated soils and groundwater. Their results allowed them to distinguish five different sources of TNT on the basis of the carbon and nitrogen isotope ratios, highlighting the potential of this kind of tool. Once again it should be noted that with the capability of determining the N isotope composition of individual compounds, in the near future one could expect to see an increase in the use of the N isotopes to study the fate of N-containing pesticides and other compounds in the environment. N isotopes also have been used to determine sources (animal waste, fertilizer, sewage) of excess nitrate (the U.S. standard is 10 mg/L) in groundwater. Similarly O isotopes again have yet to see any significant applications in forensic studies of contemporary environments although they have been used extensively in paleoclimate studies. It should be noted of course that O, N, Cl, S, and certain radioisotopes have been used extensively in many hydrogeology studies for such purposes as dating groundwater, characterizing groundwater quality, and examining changes in hydro-

geological cycles and studying precipitation (Clark and Fritz, 1997). However, the majority of these applications are not strictly forensic as discussed in the context of the current monograph.

10.5 CHARACTERIZATION OF CHLORINATED COMPOUNDS

The widespread occurrence, and stability, of chlorinated contaminants in the environment, PCBs, dioxins, DDT, and a wide variety of other compounds, has led to an increased interest in the possible use of the chlorine isotope ratios as a means of studying the origin, transport, and fate of these compounds in the environment. A paper by Reddy *et al.* (2000) studied a number of PCBs from a variety of sources and noted that the δ^{37}Cl values fell within a narrow range from −3.4 to 2.1‰ but found no correlation between the mass percent of chlorine and δ^{37}Cl values. It should be noted that the samples used in this particular study were obtained from environmental standard suppliers who in turn had obtained the materials from Bayer (Germany) and Caffaro (Italy). In the United States there was only one manufacturer of PCBs, namely Monsanto. The variability observed in the standards was suggested to arise from differences in the starting materials, synthetic methods, purification, or postproduction storage and handling of these products or some combination of all these factors. PCBs extracted from sediments ranged from −4.54 to −2.25‰. These isotopic variations suggest that δ^{37}Cl could be a useful tool for tracing the sources and fate of PCBs. In a limited number of experiments there did not appear to be any significant differences between the δ^{37}Cl values for individual congeners.

In another slightly different approach with PCBs, Jarman *et al.* (1998) examined the carbon isotope variations on individual congeners in a variety of commercial PCBs. The bulk carbon isotope values for individual PCBs varied from −22.34 to −26.95‰, whereas individual congeners showed values ranging from −18.65 to −27.98‰. Large differences were observed for individual congeners between mixtures and the δ^{13}C values for individual congeners showed a depletion of ^{13}C with increasing chlorine content, probably as the result of a kinetic isotope effect caused by the position of the chlorine atom on the biphenyl molecule. Such a wide diversity of carbon isotope values could prove to be very powerful in source determination for PCBs in the environment. Recent developments in GCIRMS capabilities plus the results described in the preceding papers would suggest that a combination of both the carbon and chlorine isotope compositions could prove to be a powerful tool in distinguishing PCBs, and other chlorinated compounds, from different sources. Difficulties in obtaining authentic standards of PCBs manufac-

tured in the 1960–1970 period could make it difficult to assign absolute sources of the PCBs in sediments. However, variations in the carbon and chlorine isotopic compositions should be sufficient to determine whether the presence of PCBs in the sediments resulted from single or multiple inputs to a particular site.

In the past few years attention has turned to the more volatile chlorinated compounds such as PCE, TCE, and related degradation products. These chlorinated hydrocarbons have been used in a variety of industries since the 1940s including electronic and instrument manufacturing, aerospace, machine, printing, and dry cleaning industries (Beneteau *et al.*, 1999; Hunkeler *et al.*, 1999; Jendrzejewski *et al.*, 2001). They have been, and continue to be, the cause of major environmental contaminant in water as a result of spills, leaks, or improper disposal. As these compounds are DNAPLs (dense nonaqueous phase liquids), they tend to persist for a long time in groundwater systems (Feenstra and Cherry, 1988). As an example, TCE is detected in approximately 19% of groundwater samples worldwide, at mean concentrations of 27.3 ppb (US EPA, 1989, cited in IARC 1995). In contaminated sites, the complex dispersal of these compounds, the presence of different contaminant sources and their ability to migrate long distances often leads to complex plumes of dissolved chlorinated solvents and related degradations products, which are difficult to trace to source (Huang *et al.*, 1999). As with some other organic groundwater contaminants like MTBE, GCIRMS has been recognized as a valuable technique in the investigation of the sources and behavior of chlorinated compounds. Most of the work to date with the chlorinated compounds has been related to utilization of the carbon isotopes and, to a lesser extent, hydrogen isotopes.

With the abundance of chlorinated hydrocarbons in the environment there is a growing interest in utilizing chlorine isotopes as an additional tool for source apportionment. At the present time there is still no continuous flow measurement method available because conversion of the chlorine into a measurable gas is a complex process, and has not been carried out successfully online yet. The off-line methods currently convert chlorinated compounds into methyl chloride (Warmerdam *et al.*, 1995; Holt *et al.,* 1997; Jendrzejewski *et al.*, 2001). The determination of the chlorine isotopic composition is based on the mass 52/50 of methyl chloride ($CH_3{}^{37}Cl/CH_3{}^{35}Cl$) and several studies in dual inlet measurements for pure chlorinated solvents have been reported (Warmerdam *et al.*, 1995; Jendrzejewski *et al.*, 2001; Numata *et al.*, 2002; Shouakar-Stash *et al.*, 2003). Measured $\delta^{37}Cl$ values for the chlorinated compounds are in the range of −3.54 to 4.08‰ (Shouakar-Stash *et al.*, 2003). $\delta^{37}Cl$ analysis has also been used in the investigation of reductive dechlorination of

chlorinated ethylenes (Sturchio *et al.*, 1998; Heraty *et al.*, 1999; Numata *et al.*, 2002). Numata *et al.* (2002) reported chlorine fractionation in anaerobic degradation of PCE and TCE by testing various bacteria. Heraty *et al.* (1999) reported smaller isotope effects for the aerobic degradation of DCM and TCE and noted the δ^{37}Cl values of the residual DCM increased as a function of biodegradation.

The availability of online δ^{37}Cl measurements would be an asset in the degradation studies of chlorinated solvents that, with the few exceptions mentioned earlier, have been limited to online δ^{13}C measurements (e.g., see Hunkeler *et al.*, 1999; Bloom *et al.*, 2000; Salter *et al.*, 2001; Sherwood Lollar *et al.*, 2001). The development of appropriate means to convert chlorine online into methyl chloride or another species suitable for subsequent isotope analysis therefore remains one of the most challenging research areas in CSIA. A successful approach certainly would be adopted worldwide and commercialized in a short time. Some studies dealt with the continuous GC-IRMS measurements of δ^{37}Cl of inorganic chlorine in environmental samples (Wassenaar and Koehler, 2004; Shouakar-Stash *et al.*, 2005). Their results on inorganic species, particularly perchlorate species (Sturchio *et al.*, 2003; Shouakar-Stash *et al.*, 2005), are promising and illustrate the potential of on-line chlorine assays in environmental samples.

10.6 BIODEGRADATION

Carbon, hydrogen, and chlorine isotopes are now being used in a fairly regular and systematic manner for evaluating the origin of contaminants and the correlation of these contaminants with their suspected sources; however there are a number of other applications in environmental chemistry based on changes in isotopic compositions, particularly carbon values. The most important application for these values, in addition to correlations, would be monitoring the rate and/or extent of biodegradation of individual compounds in the environment. The potential of this approach is enormous, particularly in the case of groundwater samples and specifically for compounds that are very soluble in water such as MTBE (Gray *et al.*, 2002). One of the problems with compounds such as MTBE is that they are mobile and concentration changes at the monitoring wells do not necessary reflect changes in concentrations resulting from biodegradation. The changes in concentration at a specific water well may simply reflect movement of the samples within the plume. However, the potential of this approach has been demonstrated in a number of field and laboratory studies, both published and unpublished. For example, in a number of studies from our laboratory and others, it has been shown that

with increasing biodegradation MTBE progressively becomes isotopically heavier (Gray *et al.*, 2002; Mancini *et al.*, 2003; Chartrand *et al.*, 2005; Kuder *et al.*, 2005). This effect has now been observed at many sites and appears to be reproducible.

Similar isotopic changes have also been observed for other compounds undergoing biodegradation, particularly in laboratory studies. Attempts are being made to try and quantify these changes, which if successful, could provide an excellent tool to demonstrate progressive mineralization of the MTBE. There are some potential limitations since biodegradation anywhere depends upon a number of factors such as oxicity, presence of nutrients, nature of the microbial community, and the salinity of groundwater. All processes that work well in the laboratory may not work as well or in such a predictable manner in the field. Ahad *et al.* (2000) noticed a small but reproducible isotopic enrichment of approximately 2‰ during the degradation of toluene in the laboratory. As with the MTBE, this introduces the possibility of using the approach in the field although such a relatively small fractionation could be quite difficult to detect at a complex field site. However, with the current generation of GCIRMS systems, a precision of 0.5‰ is easily attainable.

Another problem that must be taken into consideration is the fact that rates and changes in these isotopic compositions may not be the same at different sites, meaning that each site would need to be calibrated independently. Microbial degradation of toluene by Meckenstock *et al.* (1999) produced a shift in the $^{13}C/^{12}C$ values of the residual fraction under different environmental conditions. Many biological reactions are known to produce a $^{13}C/^{12}C$ isotopic fractionation of the substrate. In general the ^{12}C isotopes are used preferentially and the naturally occurring ^{13}C is discriminated against. As a result of the preferential consumption of the lighter isotope, the residual substrate will show a corresponding increase in its ^{13}C content.

In their papers Meckenstock *et al.* (1999, 2002) demonstrated, using toluene as a substrate, that shifts of up to 10‰ in favor of the heavier isotope could be observed depending on the particular strain of bacterium being utilized. Similar experiments in a nonsterile soil column induced changes of about 2 to 3‰ in the toluene. In a study of the degradation of C_1–C_5 alkylbenzenes in crude oils, Wilkes *et al.* (2000) observed that for *o*-xylene a shift in the $\delta^{13}C$ of 4‰ was observed after loss of 70% of the original material whereas a shift of 6‰ was observed for *o*-ethyltoluene. Similar results were obtained for sites contaminated with benzene, toluene, ethylbenzene, and xylenes (BTEX) from a tanker truck accident (Meckenstock *et al.*, 2002), in a landfill leachate plume (Richnow *et al.*, 2003), and at a former gas works site (Griebler *et al.*, 2004). Calculation of the residual substrate concentrations along such monitoring

transects relies on R_0, the isotope ratio of the substrate in the source. This parent value (usually the most depleted in ^{13}C) may be replaced by the isotope ratio of the substrate in the most upstream monitoring well near the source if the source itself is not available without changing the result of the calculation.

A similar approach has been successfully used to estimate the extent of biodegradation for benzene at a contaminated site (Mancini *et al.*, 2003) based on laboratory-derived carbon and hydrogen fractionation factors (Mancini *et al.*, 2002) with mostly rather good agreement between findings from the two elements. At a former gas works site, carbon isotope fractionation of BTEX and naphthalenes during anaerobic degradation was observed (Griebler *et al.*, 2004). The observed fractionation factors for naphthalenes, however, were much smaller than for BTEX. It is not known yet if this is due to different rate-limiting steps during degradation or to a larger number of carbon atoms in the naphthalene structure that dilute the measured isotope fractionation.

Methyl tert-butyl ether (MTBE) is rapidly mineralized by a variety of aerobic bacteria but from a practical point of view, the bulk of MTBE plumes are anaerobic and anaerobic biodegradation will be more significant for natural attenuation. A number of studies have shown MTBE degradation in anaerobic microcosms and in the field, under a range of electron acceptor conditions (Bradley *et al.*, 1999; Finneran and Lovley, 2001). Degradation under anaerobic conditions (Somsamak *et al.*, 2001) is slower and in most cases may be restricted to biotransformation of MTBE to tert-butyl alcohol (TBA). As a consequence, natural attenuation of anaerobic plumes tends to be relatively slow, resulting in the accumulation of tert-butyl alcohol (TBA) as MTBE disappears.

Microcosm studies offer easily interpretable results, but they are expensive and tend to be time-consuming. Microcosm studies may take from several months to over a year to complete. Microcosm studies can show only that the aquifer harbored microorganisms that were capable of degrading MTBE under the conditions that pertained at the time the aquifer material was sampled. The stable isotope approach provides direct information on the extent to which MTBE has been degraded in the groundwater.

The application of monitored natural attenuation to MTBE at gasoline spill sites is limited by the difficulty to convincingly demonstrate the removal of MTBE. Trends in the concentration of MTBE over time remain the chief line of evidence for attenuation. Evidence of biodegradation based on concentrations of electron acceptors and degradation products is typically inconclusive. The presence of TBA in ground water is not conclusive evidence of biodegradation of MTBE, because TBA is also a trace component of commercial MTBE

added to gasoline. TBA is miscible with water and will partition from gasoline, potentially reaching high concentration in groundwater in the absence of MTBE biodegradation (Zwank *et al.*, 2002).

Anaerobic biodegradation of MTBE using stable carbon and hydrogen isotopes was documented by the work of Kuder *et al.*, 2005. An initial study reported enrichment of carbon isotopes during anaerobic biodegradation of MTBE (Kolhatkar *et al.*, 2002), but progress in method optimization also allowed analysis of hydrogen isotope ratios in MTBE in aqueous samples at environmentally relevant concentrations. Having data on both carbon and hydrogen was particularly powerful, since the concurrent fractionation of C and H can prove that a field sample was degraded through the anaerobic pathway. If the pathway was anaerobic, then the isotope data can be used to calculate the extent of biodegradation. The calculated extent of biodegradation can be compared to overall attenuation in concentrations to evaluate the relative contribution of biodegradation.

Clearly a great deal of work remains to be done in terms of isotopic changes that may be expected during biodegradation since such changes depend on a variety of factors including the compound itself as well as all the associated environmental factors and strains of bacteria responsible for the degradation reactions. Furthermore the possibility of abiogenic degradation must also be considered. At this stage of development it certainly does not appear that any universal set of observations or guidelines have started to appear for individual compounds that can be used in a wide variety of environments. To complicate the matter further, not all compounds show any degree of isotopic fractionation upon biodegradation.

O'Malley *et al.* (1994) noted that naphthalene showed no isotopic fractionation as a result of biodegradation and in similar studies, Kelly *et al.* (1995) and Trust *et al.* (1995) showed that there was no change in the isotopic composition of acenaphthene, fluorene, phenanthrene, and fluoranthene. The extent of fractionation will decrease with increasing molecular size. During the degradation process only one carbon is attacked and if this is one carbon out of four, the isotopic effect will be much more observable than if it is one carbon out of 12 (Elsner *et al.*, 2005). The extent of isotopic fractionation may also depend on the enzymatic mechanism (Drenzek *et al.*, 2001; Mazeas *et al.*, 2002b; Morasch *et al.*, 2002) and degradation of aromatic hydrocarbons by a ring dioxygenase was observed to yield no significant carbon isotope fractionation. Several recent studies (Ward *et al.*, 2000; Hunkeler *et al.*, 2001; Gray *et al.*, 2002; Mancini *et al.*, 2002; Morasch *et al.*, 2002), utilizing both carbon and hydrogen isotope measurements, concluded that the much greater hydrogen isotope fractionation frequently observed is a more powerful tool for

providing evidence of biodegradation at contaminated field sites, particularly for low levels of biodegradation, than the carbon fractionation. However, one should be aware that isotopic fractionation for a specific element depends very much on the rate-limiting step of the biochemical reaction mechanism. As was found for source apportionment studies, a combination of the hydrogen and carbon isotopic analysis has the potential to improve the assessment of biodegradation.

An alternative approach to the biodegradation problem was described several years ago when Aggarwal and Hinchee (1991) proposed measuring the CO_2 produced by hydrocarbon degradation. It was anticipated that the carbon isotopic composition of the CO_2 produced in this way from petroleum hydrocarbons could be distinguished from that produced from other sources and other mechanisms. In the analysis of CO_2 from three sites contaminated with jet fuel it was noted that $\delta^{13}C$ in the CO_2 was about 5‰ lower than the values for CO_2 at the uncontaminated locations. Hence it would appear that in areas undergoing aerobic biodegradation this approach could provide viable evidence for *in situ* biodegradation of the hydrocarbons. Extensive levels of biodegradation could lead to anaerobic conditions for the biodegradation and under these conditions degradation of the CO_2 could affect the $\delta^{13}C$ values for the residual CO_2; clearly this is an area where additional studies are required.

The isotopic composition of the dissolved inorganic carbon (DIC) can also be used as a means to monitor mineralization of contaminants in the environment. For example, at a site where a ground water aquifer was contaminated with jet fuel, the $\delta^{13}C$ values for the DIC ranged from −28 to +11.9‰. The $\delta^{13}C$ value for the jet fuel was −27‰, and following degradation under aerobic conditions produced DIC with values of −26‰ or −18‰ under sulfate reducing conditions. In periods of low rainfall or lack of oxygen when methanogenesis was the major terminal electron acceptor process, the DIC values ranged up to +11.9‰. The results obtained in this study demonstrated that the $\delta^{13}C$ values of the DIC could be used to indicate zonation of biodegradation processes under the influence of hydrologically controlled electron acceptor availability (Landmeyer *et al.*, 1996).

In a similar approach, stable isotope ratios of C and H of dissolved CH_4 and the DIC in the groundwater in the vicinity of a crude oil spill near Bemidji, Minnesota were used to support the concept of CH_4 production by acetate fermentation (Revesz *et al.*, 1995). Two biogenic CH_4 production pathways are commonly recognized—the reduction of CO_2 by H_2 and the fermentation of acetate. Acetate fermentation refers to methanogenesis that involves transfer of a CH_3^- group from a substrate (e.g., acetate, methylamines, etc.).

The fractionation factor between coexisting CO_2 and CH_4 can be expressed as:

$$\alpha_{CO_2\text{–}CH_4} = \delta^{13}C_{CO_2}/\delta^{13}C_{CH_4} \tag{10.1}$$

For reduction of CO_2 this value is in the range 1.05–1.09 and for fermentation in the range 1.04–1.06. If the pathways cannot be distinguished solely by use of the carbon values, it is possible to utilize the H/D values to confirm the suspected pathway. In freshwater environments methanogenesis usually occurs by both acetate fermentation and CO_2 reduction pathways, with fermentation being the source of about 70% of the CH_4 in freshwater environments, and the remainder coming from CO_2 reduction. In the case of the Bemidji samples, the CH_4 production by acetate fermentation caused an increase in concentration and $\delta^{13}C$ values for the dissolved inorganic carbon in the saturated zone. In the unsaturated zone, the CH_4 is oxidized at shallower depths as indicated by an increase in $\delta^{13}C$ values and decrease in DIC concentration. The $\delta^{13}C$ values for the CO_2 increase with depth in accordance with oxidation of CH_4.

Biological degradation of organic refuse is another prolific source of both CO_2 and CH_4. Landfill methane or biogenic methane typically will have $\delta^{13}C$ values ranging from −42‰ to −61‰. Nonlandfill environmental values of $\delta^{13}C$ in methane range from −60‰ to −95‰ for biogenic reduction from CO_2 and from −47‰ to −63‰ via acetate formation (Whiticar *et al.*, 1986; Schoell, 1988). A very practical application for these variations in the isotopic composition of the methane is to distinguish natural gas sources of methane from biogenic methane (Kaplan *et al.*, 1997). There have been numerous reported cases of anomalously high concentrations of methane in populated areas. The first reaction is to place responsibility with the operators of any natural gas pipeline operators in the area. In many cases this may indeed be the source of the gas. However, there have also been cases where these anomalous concentrations simply result from a buildup of methane coming from degradation of organic refuse, typically from landfills. Source identity can be established very quickly on the basis of the isotopic composition of the methane. Natural gas methane will be relatively heavy with $\delta^{13}C$ values in the −35 to −45‰ range. Natural gas will also typically contain C_{2+} components, which will also be isotopically heavy. Biogenic methane as noted earlier will be isotopically much lighter than the natural gas methane. Furthermore it should also be noted that unlike biogenic gas, thermogenic methane will be devoid of ^{14}C.

Hence in summary it would appear that in addition to the various correlation approaches just mentioned, the most significant additional use of isotope values would have to be monitoring the effects, rate, and possible mechanisms of biodegradation. The approach should be particularly useful in the case of

single components that are very water soluble and whose rates of degradation are difficult to measure by other methods.

10.7 COMBINING ISOTOPE METHODS WITH OTHER METHODS

Stable isotopes have many powerful applications in environmental forensics, but in most cases they should be used in conjunction with other analytical techniques, such as GC and GCMS. In almost all applications the isotope values should be used as a preliminary tool to provide guidance on possible relationships between samples. This is true of both bulk isotope numbers and values for individual components, although for single components, isotope values may be the only tool available for correlation purposes.

The techniques that will be used most commonly in conjunction with the carbon isotopes, either the bulk isotopes or those obtained via GCIRMS, will be GC and GCMS or GCMSMS. These techniques are not discussed in detail here since they are well documented in the literature and here in Chapter 9. However, to illustrate the value of such combinations, two examples will be provided. The first of these involves a number of tar balls collected from the Indian Ocean washing onto the beaches of the Seychelles Islands. The question of interest here was whether or not there was any relationship between various samples, and second, to determine whether these samples were derived from passing tanker traffic or from underwater seeps. The chromatograms showed the samples had been weathered to varying degrees in the ocean as a result of evaporation, water-washing, and biodegradation (see Figure 10.17). Varying amounts of the n-alkanes had been removed from the samples, hence the bulk numbers for the samples would not be very useful since these values will change as a result of the n-alkane removal. Since these samples were crude oils, the sterane and terpane biomarker distributions can be used to demon-

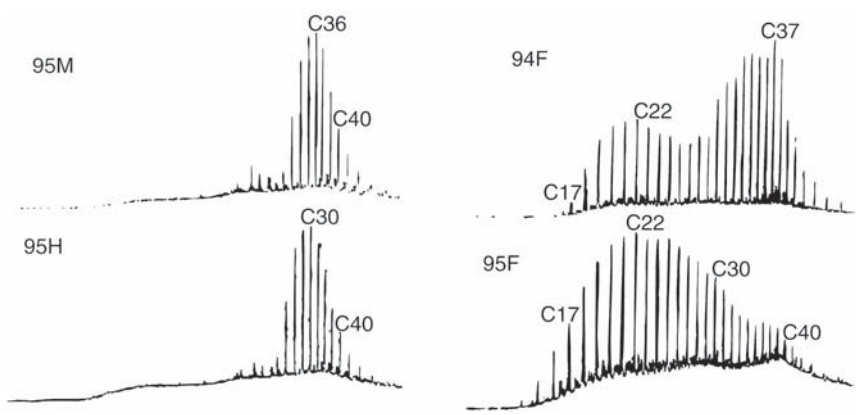


Figure 10.17

Gas chromatograms of the heavily biodegraded Seychelles Islands tar ball samples. Samples could not be correlated on the basis of the bulk isotope numbers due to removal of varying amounts of n-alkanes.

strate the relationships between the four samples. The biomarker data showed the two samples on the left-hand side of the diagram were clearly related to each other and different from the two on the right-hand side of the diagram (see Figure 10.17). Detailed interpretation reveals that the two samples on the left-hand side were derived from carbonate source rocks and had many characteristics similar to oils from the Middle East. Hence it is suspected that these samples were indeed derived from tanker traffic. The other two samples were not derived from carbonate source rocks but had characteristics similar to other samples known to be derived from lacustrine source rocks in the Seychelles area. A cursory examination of the two chromatograms on the right-hand side could easily suggest that these two samples are virtually identical to each other. However, carbon isotope values for the individual n-alkanes of the two samples showed the samples to be isotopically distinct (see Figure 10.18). This finding prompted closer examination of the biomarker chromatograms, which subsequently revealed the presence of a rather unique biomarker, oleanane, in one sample but not the other. The presence of this compound plus the isotopic differences permitted the conclusion to be made that these two seep samples originated from different geological facies of the same source

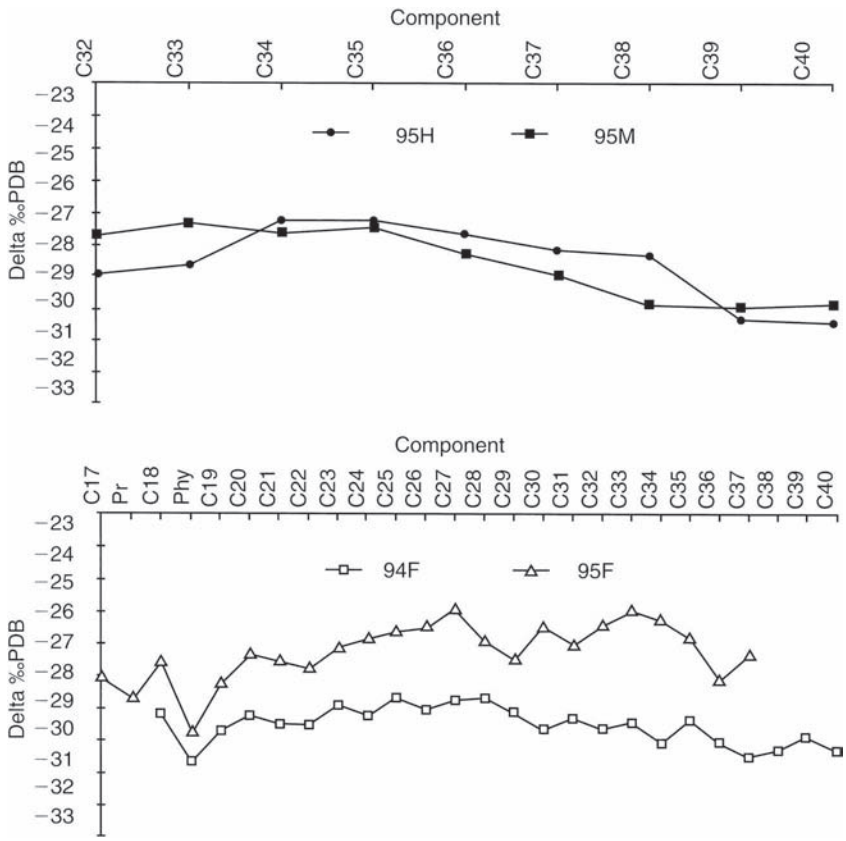


Figure 10.18
GCIRMS could be used to establish a relationship between the Seychelles Islands tar balls samples on the basis of their sterane and terpane distributions.

rock. Hence this is an excellent example of where the combination of GC, GCMS, and GCIRMS were combined to differentiate origins of these samples. This example illustrates differences between two naturally occurring seeps, but such small differences could also be responsible in the differentiation of two sources of crude oil responsible for contamination of the environment.

Another classic example of this isotope approach is Kvenvolden *et al.* (1995), which described a bulk carbon isotope study on oil residues collected from the beaches of Prince William Sound after the Exxon Valdez spill. It was noted that basically there were two populations of residues: one with an isotopic composition similar to the Exxon Valdez oil of around −29‰, and one with a value of around −23‰—quite different from any Alaskan oil. So the carbon isotope data suggested two populations of crude oil residues. Once this information became available, more detailed analyses of the various residues by GC and GCMS revealed that the heavier residues were actually derived from California crudes sourced from Monterey Shale.

Despite increased specificity, the GCIRMS technique still suffers from a number of problems and limitations related to the correlation of moderate to severely biodegraded oils. In order to be successful, GCIRMS requires the presence of well-resolved components in the chromatogram. Severely weathered, particularly biodegraded samples typically have lost their n-alkanes and the resulting chromatogram will be dominated by a hump of unresolved components. However, it has been shown that isolation and pyrolysis of the asphaltenes from these types of samples provides an alternative method for undertaking correlations (Behar *et al.*, 1984). Previous studies have demonstrated the use of bulk isotopic composition of asphaltenes as a correlation parameter for severely weathered oils (Hartman and Hammond, 1981; Macko *et al.*, 1981). Hartman and Hammond (1981) did not see any change in the isotopic composition of these compounds although Stahl (1977) observed a slight ^{12}C enrichment with increasing artificial biodegradation. In general the asphaltene pyrolysis is performed off-line (Eglinton, 1994) and the pyrolysates collected fractionated into saturates + unsaturates, aromatics, and polar fractions. The saturate + unsaturate fraction containing the n-alkene/n-alkane doublets can be subsequently analyzed by GCIRMS. This approach was applied to the asphaltenes from the oils that were artificially biodegraded for two and four months as described earlier (see Figure 10.8). The pyrograms of the biodegraded oil samples showed a good correlation with the original oil and had a predominance of the lighter hydrocarbons, maximizing at C_{15}–C_{17} (see Figure 10.19). However, the biodegraded samples did have a higher contribution of n-alkanes around C_{28} and C_{38} compared to the original oil. Analyses of these fractions by GCIRMS showed the isotopic compositions of individual n-alkanes of these fractions from the original and degraded samples to be quite similar (see Figure 10.20). Standard deviations calculated between the

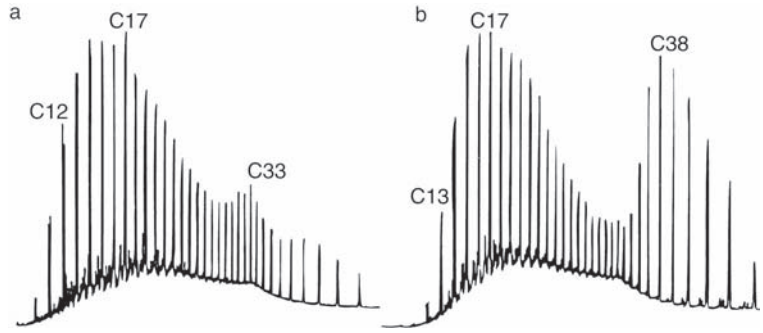


Figure 10.19

*The pyrograms of the asphaltenes from the two biodegraded oils (**a**. after two months; **b**. after four months) are similar to that of the initial oil with a predominance of the light ends, maximizing at C_{15}–C_{17}. These pyrolysis fractions were used to correlate the degraded and nondegraded samples.*

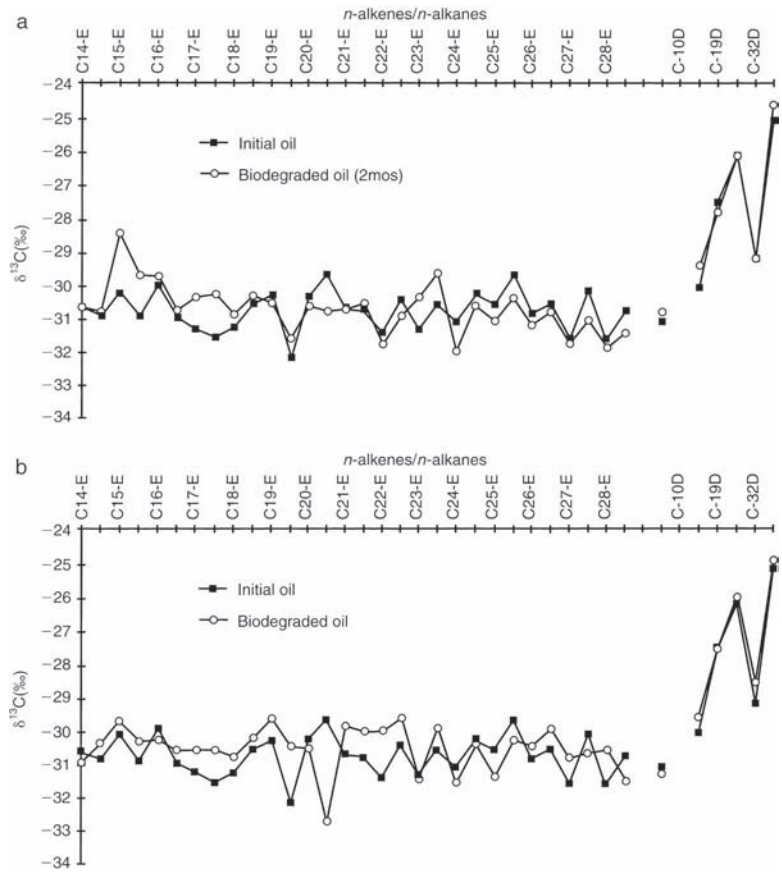


Figure 10.20

Analyses of the asphaltene pyrolysis fractions by GCIRMS showed the isotopic compositions of individual n-alkanes from these three fractions (see Figure 10.19) to be very similar and capable of correlating degraded and nondegraded samples. (Reprinted with permission from Mansuy et al. *(1997), American Chemical Society.)*

original and biodegraded oils were relatively low considering the complexity of the mixtures analyzed. The n-alkene/n-alkane doublets and the high background (see Figure 10.19) increase the chance of coelutions and affect the reproducibility of the analyses, which ranges between 0.33 and 0.40‰. The isotopic composition of the n-alkanes generated by the asphaltene pyrolysis is very close to the isotopic composition of the n-alkanes of their respective oils and of the initial oil.

These results confirm the ability of GCIRMS and the asphaltene pyrolysates to correlate severely biodegraded oils with their unweathered counterparts with an analytical error still acceptable despite the chromatographic problems encountered during the analyses of such complex mixtures. Although it may be argued that the number of steps involved in this process may render it ineffective as a correlation tool, it should be remembered that there are many situations where oil samples have been recovered from old storage tanks or tar balls in the oceans that have been weathered over many years, severely altering the biomarker fingerprints. In this situation, isolation of the asphaltenes, off-line pyrolysis, fractionation of the pyrolysates, followed by GCIRMS may prove to be extremely valuable when other techniques provide ambiguous or inconclusive results.

A final example to illustrate the combined use of isotopes and biomarkers to resolve an environmental problem is illustrated in Figures 10.21 to 10.23. Samples from the two monitoring wells illustrated in these figures were at the

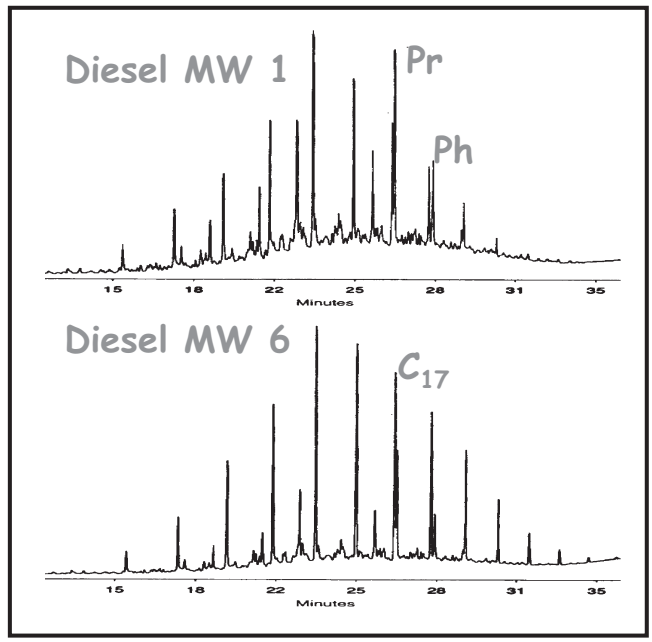


Figure 10.21

Gas chromatograms of the products in two monitoring wells from the same site. The uppermost sample showed signs of biodegradation and it cannot be ascertained from these chromatograms alone whether there is any relationship between the two samples.

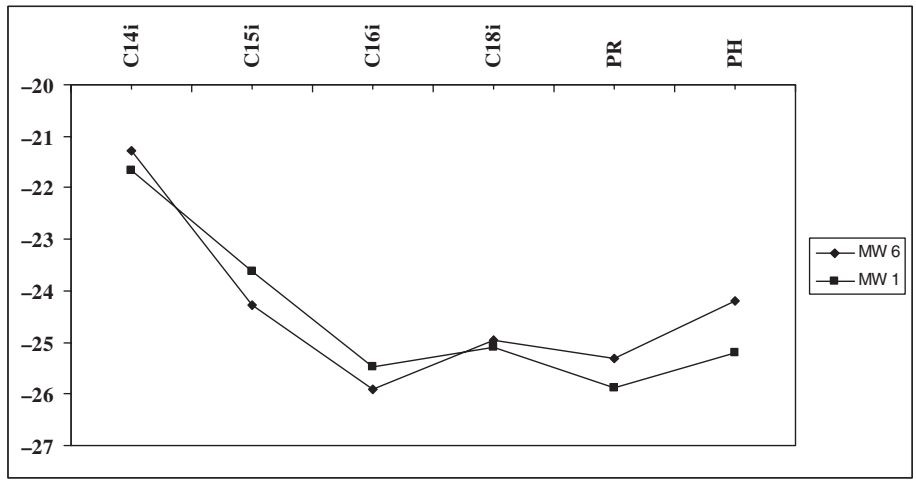


Figure 10.22

The isotopic composition of the individual isoprenoids in the two samples suggested the two samples were derived from the same or very similar sources.

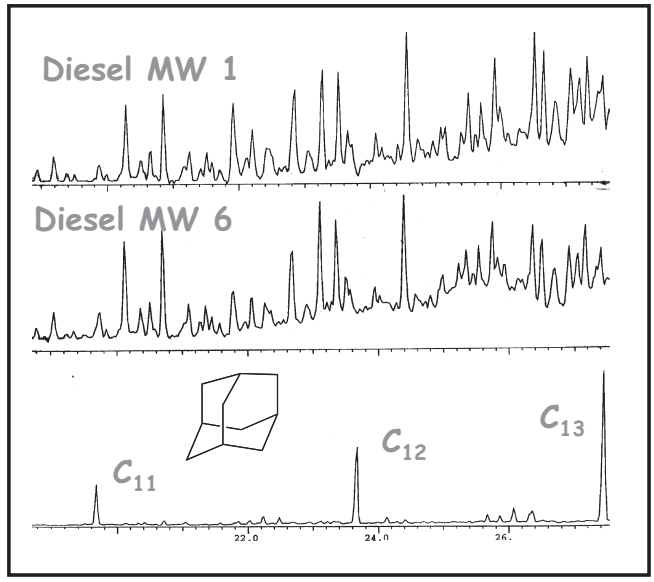


Figure 10.23

Additional evidence to support the concept that the two samples in Figures 10.21 and 10.22 are derived from the same source can be obtained from GCMS and detection of the adamantanes, shown here, and other families of biomarkers.

center of a dispute as to whether or not they were related and hence part of the same plume of hydrocarbons. Both samples were identified as a refined product, namely diesel, but the sample in monitoring well 1 appeared different although these differences are simply due to biodegradation, as can be observed from the much smaller n-C_{17}/Pr value in this sample compared to monitoring well 6 (see Figure 10.21). The samples were analyzed by GCIRMS and carbon values were determined for only the isoprenoids; these values are shown in Figure 10.22, suggesting a similarity in origins for the two samples.

However, one should not depend solely upon one piece of data to suggest a relationship between two samples. Hence the two samples were analyzed by GCMS and in particular the distribution of the adamantanes was determined and shown in Figure 10.23. The similarity in these distributions along with the isotope data and other biomarker fingerprints not shown here clearly established a relationship between these two samples.

Several new developments have been made in LC/IRMS with the new interfaces by Thermo-Electron and have shown good precision for the $\delta^{13}C$ measurement of various analyte classes, including carbohydrates, chlorophyll, and proteins. Some applications in environmental sciences have been reported (Abramson *et al.*, 2001). Some other studies have also been reported. For example, Krummen *et al.* (2004) tested with success the on-line coupling of a liquid chromatograph to a stable isotope ratio mass spectrometer for $^{13}C/^{12}C$ determination on amino acids, carbohydrates, and drugs. Godin *et al.* (2005) used LC/IRMS for isotopic measurement of biomolecules and showed that for compounds with molecular weights ranging from 131 to 66390 Da, the precision was better than 0.3‰, and accuracy varied from 0.1 to 0.7‰. Sessions *et al.* (2005) described a moving-wire analyzer for measuring ^{13}C in dissolved, involatile organic materials. The results show that the accuracy is better than 0.5‰ for nearly all dissolved analytes tested, including lipids, proteins, nucleic acids, sugars, halocarbons, and hydrocarbons. Future work will certainly benefit from these latest developments in LC/IRMS interfaces.

REFERENCES

Abrajano, J. and Sherwood Lollar, B. (1999) Compound-specific isotope analysis: Tracing organic contaminant sources and processes in geochemical systems. *Organic Geochemistry* 30(8B), v–vii.

Abramson, F.P., Black, G.E., and Lecchi, P. (2001) Application of high-performance liquid chromatography with isotope-ratio mass spectrometry for measuring low levels of enrichment of underivatized materials. *Journal of Chromatography A* 913(1–2), 269–273.

Aggarwal, P.K. and Hinchee, R.E. (1991) Monitoring *in-situ* biodegradation of hydrocarbons by using stable carbon isotopes. *Environment Science and Technology* 25(6), 1176–1180.

Ahad, J.M.E., Sherwood Lollar, B., Edwards, E.A., Slater, G.F., and Sleep, B.E. (2000) Carbon isotope fractionation during anaerobic biodegradation of toluene: Implications for intrinsic bioremediation. *Environment Science and Technology* 34(5), 892–896.

Aislabie, J.M., Balks, M.R., Foght, J.M., and Waterhouse, E.J. (2004) Hydrocarbon spills on Antarctic soils: Effects and management. *Environment Science and Technology* 38(5), 1265–1274.

Alzaga, R., Montuori, P., Ortiz, L., Bayona, J.M., and Albaigés, J. (2004) Fast solid-phase extraction–gas chromatography–mass spectrometry procedure for oil fingerprinting: Application to the Prestige oil spill. *Journal of Chromatography A* 1025(1), 133–138.

Becker, S. and Herner, A.V. (1998) Characterization of crude oils by carbon and sulfur isotope ratio measurements as a tool for pollution control. *Isotopes in Environmental Health Studies* 34, 255–264.

Behar, F., Pelet, R.J., and Roucache, J. (1984) Geochemistry of asphaltenes. In *Advances in Organic Geochemistry 1983*, P.A. Schenck, J.W. de Leeuw, and G.W.M. Lijmbach, eds. Pergamon Press, Oxford, 587–595.

Beneteau, K.M., Aravena, R., and Frape, S.K. (1999) Isotopic characterization of chlorinated solvents-laboratory and field results. *Organic Geochemistry* 30(8), 739–753.

Bloom, Y., Aravena, R., Hunkeler, D., Edwards, E., and Frape, S.K. (2000) Carbon isotope fractionation during microbial dechlorination of trichloroethene, cis-1,2-dichloroethene, and vinyl chloride: Implications for assessment of natural attenuation. *Environment Science and Technology* 34(13), 2768–2772.

Boreham, C.J., Dowling, L.M., and Murray, A.P. (1995) Biodegradation and maturity influences on n-alkane isotopic profiles in terrigenous sequences. *Abstracts 17th International Meeting EAOG*, San Sebastian, 539–541.

Bradley, P.M., Landmeyer, J.E., and Chapelle, F.H. (1999) Aerobic mineralization of MTBE and tert-butyl alcohol by stream-bed sediment micro-organisms. *Environmental and Science Technology* 33(11), 1877–1879.

Brenner, R.C., Magar, V.S., Ickes, J.A., Abbott, J.E., Stout, S.A., Crecelius, E.A., and Bingler, L.S. (2002) Characterization and fate of PAH-contaminated sediments at the Wyckoff/Eagle Harbor superfund site. *Environment Science and Technology* 36(12), 2605–2613.

Chapman, P.M. and Riddle, M.J. (2005) Toxic effects of contaminants in polar marine environments. *Environmental Science and Technology* 39(9), 200A–206A.

Chartrand, M.M.G., Morrill, P.L., Lacrampe-Couloume, G., and Sherwood Lollar, B. (2005) Stable isotope evidence for biodegradation of chlorinated ethenes at a fractured bedrock site. *Environmental Science and Technology* 39(13), 4848–4856.

Chen, S.-J., Luo, X.-J., Mai, B.-X., Sheng, G.-Y., Fu, J.-M., and Zeng, E.Y. (2006) Distribution and mass inventories of polycyclic aromatic hydrocarbons and organochlorine pesticides in sediments of the Pearl River estuary and the Northern South China Sea. *Environmental Science and Technology* 40(3), 709–714.

Coffin, R.B., Miyares, P.H., Kelley, C.A., Cifuentes, L.A., and Reynolds, C.M. (2001) $\delta^{13}C$ and $\delta^{15}N$ isotope analysis of TNT; two dimensional source identification. *Environment Toxicology and Chemistry* 20(12), 2676–2680.

Craig, D.R., Mauro, D., Saber, D., Sirivedhin, T., Philp, R.P., and Allen, J.A. (2005) Carbon isotope ratios of PAHs in urban background soil. 3rd International Conference on Remediation of Contaminated Sediments. 57th Annual Meeting American Academy of Forensic Sciences, Feb. 21–26th, New Orleans.

Czuczwa, J.M. and Hites, R.A. (1984) Environmental fate of combustion-generated polychlorinated dioxins and furans. *Environmental Science and Technology* 18(6), 444–449.

Davis, A., Curnou, P.D., and Eary, L.E. (1997) Discriminating between sources of arsenic in the sediments of a tidal waterway in Tacoma, Washington. *Environmental Science and Technology* 31(7), 1985–1991.

Dempster, H.S., Sherwood Lollar, B., and Feenstra, S. (1997) Tracing organic contaminants in goundwater: A new methodology using compound-specific isotopic analysis. *Environmental Science and Technology* 31(11), 3193–3197.

Dincer, T., Al-Mugrin, A., and Zimmermann, U. (1974) Study of the infiltration and recharge through the sand dunes in arid zones with special reference to the stable isotopes and thermonuclear tritium. *Journal of Hydrology* 23(1–2), 79–109.

Dowling, L.M., Boreham, C.J., Hope, J.M., Murray, A.P., and Summons, R.E. (1995) Carbon isotope composition of hydrocarbons in ocean transported bitumens from the coastline of Australia. *Organic Geochemistry* 23(8), 729–737.

Drenzek, N.J., Eglinton, T.I., May, J.M., Wu, Q.Z., Sowers, K.R., and Reddy, C.M. (2001) The absence and application of stable carbon isotopic fractionation during the reductive dechlorination of polychlorinated biphenyls. *Environmental Science and Technology* 35(16), 3310–3313.

Drenzek, N.J., Tarr, C.H., Eglinton, T.I., Heraty, L.J., Sturchio, N.C., Shiner, V.J., and Reddy, C.M. (2002) Stable chlorine and carbon isotopic compositions of selected semi-volatile organochlorine compounds. *Organic Geochemistry* 33(4), 437–444.

Eganhouse, R.P., Baedecker, M.J., Cozzarelli, I.M., Aiken, G.R., Thorn, K.A., and Dorsey, T.F. (1993) Crude oil in a shallow sand and gravel aquifer-II. Organic Geochemistry. *Appied Geochemistry* 8(6), 551–567.

Eglinton, T.I. (1994) Carbon isotopic evidence for the origin of macromolecular aliphatic structures in kerogen. *Organic Geochemistry* 21(6–7), 721–735.

Elsner, M., Zwank, L., Hunkeler, D., and Schwarzenbach, R.P.A. (2005) New concept linking observable stable isotope fractionation to transformation pathways of organic pollutants. *Environmental Science and Technology* 39(18), 6896–6916.

Fang, J., Kawamura, K., Ishimura, Y., and Matsumoto, K. (2002) Carbon isotopic composition of fatty acids in the marine aerosols from the western North Pacific: Implication for the source and atmospheric transport. *Environmental Science and Technology* 36(12), 2598–2604.

Farran, A., Grimalt, J., Albaiges, J., Botello, A.V., and Macko, S.E. (1987) Assessment of petroleum pollution in a Mexican river by molecular markers and carbon isotope ratios. *Marine Pollution Bulletin* 18(6), 284–288.

Feenstra, S. and Cherry, J.A. (1988) Subsurface contamination by dense non-aqueous phase liquid (DNAPL) chemicals. International Groundwater Symposium. International Association of Hydrology, Halifax, Nova Scotia.

Finneran, K.T. and Lovley, D.R. (2001) Anaerobic degradation of methyl tert-butyl ether (MTBE) and tert-butyl alcohol (TBA). *Environmental Science and Technology* 35(9), 1785–1790.

Franco, M.A., Viñas, L., Soriano, J.A., de Armas, D., González, J.J., Beiras, R. *et al.* (2005) Spatial distribution and ecotoxicity of petroleum hydrocarbons in sediments from the Galicia continental shelf (NW Spain) after the *Prestige* oil spill. *Marine Pollution Bulletin*, in press.

Freeman, K.H., Hayes, J.M., Trendel, J.M., and Albrecht, P. (1990) Evidence from carbon isotope measurements for diverse origins of sedimentary hydrocarbons. *Nature* 343, 254–256.

Fuentes, H.R., Jaffe, R., and Shrotriya, R.V. (1996) Weathering of crude oil spills in seawater under simulated laboratory conditions. *5th Latin American Congress on Organic Geochemistry*, Cancun, Mexico, Abstracts, Mexican Petroleum Institute, Mexico City, Mexico, 310–312.

Fuex, A.N. (1997) The use of stable carbon isotopes in hydrocarbon exploration. *J Geochem Expl* 7(2), 155–188.

Gallon, C., Tessier, A., Gobeil, C., and Carignan, R. (2006) Historical perspective of industrial lead emissions to the atmosphere from a Canadian smelter. *Environmental Science and Technology* 40(3), 741–747.

Galt, J.A., Lehr, W.J., and Payton, D.L. (1991) Fate and transport of the Exxon Valdez oil spill. *Environmental Science and Technology* 25(2), 202–209.

Glaser, B., Dreyer, A., Bock, M., Fiedler, S., Mehring, M., and Heitmann, T. (2005) Source apportionment of organic pollutants of a highway-traffic influenced urban area in Bayreuth (Germany) using biomarker and stable carbon isotope signatures. *Environmental Science and Technology* 39(11), 3911–3917.

Godin, J.P., Hau, J., Fay, L-B., and Hopfgartner, G. (2005) Isotope ratio monitoring of small molecules and macromolecules coupled to isotope ratio mass spectrometry. *Rapid Communications in Mass Spectrometry* 19(18), 2689–2698.

Gray, R., Lacrampe-Couloume, G., Gandhi, D., Scow, K.M., Wilson, R.D., Mackay, D.M., and Sherwood Lollar, B. (2002) Carbon and hydrogen isotopic fractionation during biodegradation of methyl tert-butyl ether. *Environmental Science and Technology* 36(9), 1931–1938.

Gribble, G.W. (1998) Naturally occurring organohalogen compounds. *Accounts of Chemical Research* 31(3), 141–152.

Gribble, G.W. (2003) The diversity of naturally produced organohalogens. *Chemosphere* 52(2), 289–297.

Griebler, C., Safinowski, M., Vieth, A., Richnow, H.H., and Meckenstock, R.U. (2004) Combined application of stable carbon isotope analysis and specific metabolites determination for assessing *in situ* degradation of aromatic hydrocarbons in a tar oil-contaminated aquifer. *Environmental Science and Technology* 38(2), 617–631.

Gürgey, K., Philp, R.P., Clayton, C., Emiroğlu, H., and Siyako, M. (2005) Geochemical and isotopic approach to maturity/source/mixing estimations for natural gas and associated condensates in the Thrace Basin, NW Turkey. *Applied Geochemistry* 20(11), 2017–2037.

Hammer, B.T., Kelley., C.A., Coffin, R.B., Cifuentes, L.A., and Mueller, J.G. (1998) ^{13}C values of polycyclic aromatic hydrocarbons collected from two creosote-contaminated sites. *Organic Geochemistry* 152(1–2), 43–58.

Hanson, D. (1999) MTBE—Villain or victim. *Chemical and Engineering News*, October 19th, 49.

Harper, D.B., Kalin, R.M., Hamilton, J.T.G., and Lamb, C. (2001) Carbon isotope ratios for chloromethane of biological origin: Potential tool in determining biological emissions. *Environmental Science and Technology* 35(18), 3616–3619.

Harper, D.B., Hamilton, J.T.G., Ducrocq, V., Kennedy, J.T., Downey, A., and Kalin, R.M. (2003) The distinctive isotopic signature of plant-derived chloromethane: Possible application in constraining the atmospheric chloromethane budget. *Chemosphere* 52(2), 433–436.

Hartman, B. and Hammond, D. (1981) The use of carbon and sulfur isotopes as a correlation parameter for source identification of beach tars in southern California. *Geochimica Cosmochimica Acta* 45, 309–319.

Haws, N.W., Ball, W.P., and Bouwer, E.J. (2006) Modeling and interpreting bioavailability of organic contaminant mixtures in subsurface environments. *Journal of Contaminant Hydrology* 82(3–4), 255–292.

Hayes, J.M., Freeman, K.H., Popp, B.N., and Hoham, C.H. (1990) Compound specific isotope analysis: A novel tool for reconstruction of ancient biogeochemical processes. *Organic Geochemistry* 16(4–6), 1115–1128.

Heraty, L.J., Fuller, M.E., Huang, L., Abrajano, T., and Sturchio, N.C. (1999) Isotopic fractionation of carbon and chlorine by microbial degradation of dichloromethane. *Organic Geochemistry* 30(8), 739–753.

Hilkert, A.W., Douthitt, C.B., Schluter, H.J., and Brand, W.A. (1999) Isotope ratio monitoring gas chromatography/mass spectrometry of D/H by high temperature conversion isotope ratio mass spectrometry. *Rapid Communication in Mass Spectrometry* 13, 1226–1230.

Hillaire-Marcel, C. (1986) Isotopes and food. Chapter 12. In *Handbook of Environmental Isotope Geochemistry, vol. 2, The Terrestrial Environment, B*, P. Fritz and J.Ch. Fontes, eds. Elsevier, Amsterdam, 507–548.

Hillaire-Marcel, C. and Ghaleb, B. (1997) Thermal ionization mass spectrometry measurements of ^{226}Ra and U-isotopes in surface and groundwaters-porewater/matrix interactions revisited and potential dating implications. Isotope Techniques in the Study of Past and Current Environmental Changes in the Hydrosphere and Atmosphere. IAEA Symposium 349, Vienna.

Holt, B.D., Sturchio, N.C., Abrajano, T., and Heraty, L.J. (1997) Conversion of chlorinated volatile compounds to carbon dioxide and methyl chloride for isotopic analysis of carbon and chlorine. *Analytical Chemistry* 69(14), 2727–2733.

Holt, B.D., Heraty, L.J., and Sturchio, N.C. (2001) Extraction of chlorinated aliphatic hydrocarbons from groundwater at micromolar concentrations for isotopic analysis of chlorine. *Environmental Pollution* 113(3), 263–269.

Hong, S.H., Yim, U.H., Shim, W.J., and Oh, J.R. (2005) Congener-specific survey for polychlorinated biphenlys in sediments of industrialized bays in Korea: Regional characteristics and pollution sources. *Environmental Science Technology* 39(19), 7380–7388.

Horii, Y., Kannan, K., Petrick, G., Gamo, T., Falandysz, J., and Yamashita, N. (2005) Congener-specific carbon isotopic analysis of technical PCB and PCN mixtures using two-dimensional gas chromatography-isotope ratio mass spectrometry. *Environmental Science and Technology* 39(11), 4206–4212.

Hough, R.L., Whittaker, M., Fallick, A.E., Preston, T., Farmer, J.G., and Pollard, S.J.T. (2006) Identifying source correlation parameters for hydrocarbon wastes using compound-specific isotope analysis. *Environmental Pollution*, in press.

Huang, L., Sturchio, N.C., Abrajano, T., Heraty, L.J., and Holt, B.D. (1999) Carbon and chlorine isotope fractionation of chlorinated aliphatic hydrocarbons by evaporation. *Organic Geochemistry* 30(8), 777–785.

Hunkeler, D., Aravena, R., and Butler, B.J. (1999) Monitoring microbial dechlorination of tetrachloroethene (PCE) in groundwater using compound-specific stable carbon isotope ratios: Microcosm and field studies. *Environmental Science and Technology* 33(16), 2733–2738.

Hunkeler, D., Andersen, N., Aravena, R., Bernasconi, S.M., and Butler, B.J. (2001) Hydrogen and carbon isotope fractionation during aerobic biodegradation of benzene. *Environmental Science and Technology* 35(17), 3462–3467.

Hurst, R.W., Davis, T.E., and Chinn, B.D. (1996) Lead fingerprints of gasoline contamination. *Environmental Science and Technology* 30, 304A–307A.

Hurst, R.W. (2000) Applications of anthropogenic lead archaeostratigraphy (ALAS model) to hydrocarbon remediation. *Environmental Forensics* 1(1), 11–24.

IARC. (1995) Dry cleaning, some chlorinated solvents and other industrial chemicals. *IARC Monographs on the Evaluation of Carcinogenic Risk of Chemicals to Humans, vol. 63.* Lyon, France: International Agency for Research on Cancer-IARC, 558.

Jarman, W.M., Hilkert, A., Bacon, C.E., Collister, J.W., Ballschmitter, P., and Risebrough, R.W. (1998) Compound specific carbon isotope analysis of Aroclors, Clophens, Kaneclors, and Phenoclors. *Environmental Science and Technology* 32(6), 833–836.

Jendrzejewski, N., Eggenkamp, H.G.M., and Coleman, M.L. (2001) Characterization of chlorinated hydrocarbons from chlorine and carbon isotopic compositions: Scope of application to environmental problems. *Applied Geochemistry* 16(9–10), 1021–1031.

Juttner, I., Henkelmann, B., Schramm, K., Steinberg, C.E.W., Winkler, R., and Kettrup, A. (1997) Occurrence of PCDD/F in dated lake sediments of the Black Forest, Southwestern Germany. *Environmental Science and Technology* 31(3), 806–812.

Kaplan, I.R., Galperin, Y., Lu, S-T., and Lee, R-P. (1997) Forensic Environmental Geochemistry: Differentiation of fuel types, their sources and release time. *Organic Geochemistry* 27(5–6), 289–317.

Kelley, C.A., Coffin, R.B., Cifuentes, L.A., Lantz, S.E., and Mueller, J.G. (1995) The use of GC/C/IRMS coupled with GC/MS to monitor biodegradation of petroleum hydrocarbons. Platform Abstracts of the Third International Symposium in Situ and On Situ Bioreclaimation, San Diego CA. April 24–27, 1995.

Kelley, C.A., Hammar, B.T., and Coffin, R.B. (1997) Concentrations and stable isotope values of BTEX in gasoline contaminated groundwater. *Environmental Science and Technology* 31(9), 2469–2472.

Kolhatkar, R., Kuder, T., Philp, P., Allen, J., and Wilson, J.T. (2002) Use of compound-specific stable carbon isotope analyses to demonstrate anaerobic biodegradation of MTBE in groundwater at a gasoline release site. *Environmental Science and Technology* 36(23), 5139–5146.

Krummen, M., Hilkert, A.W., Juchelka, D., Duhr, A., Schluter, H-J., and Pesch, R. (2004) A new concept for isotope ratio monitoring liquid chromatography/mass spectrometry. *Rapid Communications in Mass Spectrometry* 18(19), 2260–2266.

Kuder, T., Wilson, J.T., Kaiser, P., Kolhatkar, R., Philp, P., and Allen, J. (2005) Enrichment of stable carbon and hydrogen isotopes during anaerobic biodegradation of MTBE: Microcosm and field evidence. *Environmental Science and Technology* 39(1), 213–220.

Kvenvolden, K.A., Hostettler, F.D., Carlson, P.R., Rapp, J.B., Threlkeld, C.N., and Warden, A. (1995) Ubiquitous tar balls with a California-source signature on shorelines of Prince William Sound, Alaska. *Environmental Science and Technology* 29(10), 2684–2694.

Lafargue, E. and Barker, C. (1988) Effect of water-washing on crude oil compositions. *American Association of Petroleum Geologists Bulletin* 72, 263–276.

Landmeyer, J.E., Vroblesky, D.A., and Chapelle, F.H. (1996) Stable carbon isotope evidence of biodegradation zonation in a shallow jet fuel contaminated aquifer. *Environmental Science and Technology* 30(4), 1120–1128.

Lee, B.-C., Shimizu, Y., Matsuda, T., and Matsui, S. (2005) Characterization of polycyclic aromatic hydrocarbons (PAHs) in different size fractions in deposited road particles (DRPs) from Lake Biwa area, Japan. *Environmental Science and Technology* 39(19), 7402–7409.

Li, M., Huang, Y., Obermajer, M., Jiang, C., Snowdon, L.R., and Fowler, M.G. (2001) Hydrogen isotopic compositions of individual n-alkanes as a new approach to petroleum correlation: Case studies from the Western Canada Sedimentary Basin. *Organic Geochemistry* 32(12), 1387–1399.

Lieser, K.H., and Steinkopff, T. (1989) Chemistry of radioactive Cesium in the hydrosphere and in the geosphere. *Radiochimica Acta* 46, 39–47.

Macko, S.A., Parker, P.L., and Botello, A.V. (1981) Persistence of spilled oil in a Texas salt marsh. *Environmental Pollution Bulletin* 2, 119–128.

Macko, S.A. and Parker, P.L. (1983) Stable nitrogen and carbon isotope ratios of beach tars on South Texas barrier islands. *Marine Environmental Research* 10, 93–103.

Mancini, S.A., Lacrampe-Couloume, G., Jonker, H., van Breukelen, B.N., Groen, J., Volkering, F., and Sherwood Lollar, B. (2002) Hydrogen isotopic enrichment: An indicator of biodegradation at a petroleum hydrocarbon contaminated field site. *Environmental Science and Technology* 36(11), 2464–2470.

Mancini, S.A., Ulrich, A.C., Lacrampe-Couloume, G., Sleep, B., Edwards, E.A., and Sherwood Lollar, B. (2003) Carbon and hydrogen isotopic fractionation during anaerobic biodegradation of benzene. *Applied Environmental Microbiology* 69, 191–198.

Mansuy, L., Philp, R.P., and Allen, J. (1997) Source identification of oil spills based on the isotopic composition of individual components in weathered oil samples. *Environmental Science and Technology* 31(12), 3417–3425.

Mazeas, L. and Budzinski, H. (2002a) Molecular and stable carbon isotopic source identification of oil residues and oiled bird feathers sampled along the Atlantic Coast of France after the Erika oil spill. *Environmental Science and Technology* 36(2), 130–137.

Mazeas, L., Budzinski, H., and Raymond, N. (2002b) Absence of stable carbon isotope fractionation of saturated and polycyclic aromatic hydrocarbons during aerobic bacterial biodegradation. *Organic Geochemistry* 33(11), 1259–1272.

McRae, C., Sun, C.G., Snape, C.E., Fallick, A.E., and Taylor, D. (1999) $\delta^{13}C$ values of coal-derived PAHs from different processes and their application to source apportionment. *Organic Geochemistry* 30(8), 881–889.

McRae, C., Snape, C.E., Sun, C-G., Fabbri, D., Tartari, D., Trombini, C., and Fallick, A.E. (2000) Use of compound-specific stable isotope analysis to source anthropogenic natural gas-derived polycyclic aromatic hydrocarbons in a lagoon sediment. *Environmental Science and Technology* 34(22), 4684–4686.

Mead, R., Xu, Y., Chong, J., and Jaffé, R. (2005) Sediment and soil organic matter source assessment as revealed by the molecular distribution and carbon isotopic composition of n-alkanes. *Organic Geochemistry* 36(3), 363–370.

Meckenstock, R.U., Morasch, B., Warthmann, R.W., Schink, B., Annweller, E., Michaelis, W., and Richnow, H.H. (1999) $^{13}C/^{12}C$ Isotope fractionation of aromatic hydrocarbons during microbial degradation. *Environmental Microbiology* 1(5), 409–414.

Meckenstock, R.U., Morasch, B., Kastner, M., Vieth, A., and Richnow, H.H. (2002) Assessment of bacterial degradation of aromatic hydrocarbons in the environment by analysis of stable carbon isotope fractionation. *Water Air Soil Pollution Focus* 2, 141–152.

Meier-Augenstein, W. (1999) Applied gas chromatography coupled to isotope ratio mass spectrometry. *Journal of Chromatography* 842(1–2), 351–371.

Michel, F.A., Kubasiewicz, R.J., Patterson, R.J., and Brown, R.M. (1984) Ground water flow velocity derived from tritium measurements at the Glouster landfill site, Glouster, Ontario. *Water Pollution Research Journal of Canada* 19, 13–22.

Milkov, A.V. (2005) Molecular and stable isotope compositions of natural gas hydrates: A revised global dataset and basic interpretations in the context of geological settings. *Organic Geochemistry* 36(5), 681–702.

Milner, C.W.D., Rogers, M.A., and Evans, C.R. (1977) Petroleum transformations in reservoirs. *Journal of Geochemical Exploration* 7, 101–153.

Morasch, B., Richnow, H.H., Schink, B., Vieth, A., and Meckenstock, R.U. (2002) Carbon and hydrogen stable isotope fractionation during aerobic bacterial degradation of aromatic hydrocarbons. *Applied Environmental Microbiology* 68, 5191–5194.

Norman, A.L., Hopper, J.F., Blanchard, P., Ernst, D., Brice, K., Alexandroi, N., and Klouda, G. (1999) The stable carbon isotope composition of atmospheric PAHs. *Atmospheric Environment* 33(17), 2807–2814.

Numata, M., Nakamura, N., Koshikawa, H., and Terashima, Y. (2002) Chlorine isotope fractionation during reductive dechlorination of chlorinated ethenes by anaerobic bacteria. *Environmental Science and Technology* 36(20), 4389–4394.

Ohe, T., Watanabe, T., and Wakabayashi, K. (2004) Mutagens in surface waters: A review. *Mutation Research/Reviews in Mutation Research* 567(2–3), 109–149.

Okuda, T., Kumata, H., Zakaria, M.P., Naraoka, H., Ishiwatari, R., and Takada, H. (2002a) Source identification of Malaysian atmospheric polycyclic aromatic hydrocarbons nearby forest fires using molecular and isotopic compositions. *Atmospheric Environment* 36, 611–618.

Okuda, T., Kumata, H., Naraoka, H., Ishiwatari, R., and Takada, H. (2002b) Vertical distributions and $\delta^{13}C$ isotopic compositions of PAHs in Chidorigafuchi Moat sediment, Japan. *Organic Geochemistry* 33(7), 843–848.

O'Malley, V.P., Abrajano, T., and Hellou, J. (1994) Determination of the $^{13}C/^{12}C$ ratios of individual PAH from environmental samples: Can PAH sources be apportioned? *Organic Geochemistry* 21(6–7), 809–822.

O'Malley, V.P., Abrajano, T., and Hellou, J. (1996) Stable isotopic apportionment of individual polycyclic aromatic hydrocarbons in St. Johns Harbour, Newfoundland. *Environmental Science and Technology* 30(2), 634–639.

Oros, D., Aboul-Kassim, T.A., Simoneit, B.R.T., and Collier, R.C. (1996) V Latin American Congress on Organic Geochemistry, Cancun, Mexico, Abstracts, 330–332.

Palmer, S.E. (1993) Effect of biodegradation and water-washing on crude oil composition. In *Organic Geochemistry*, M.H. Engel and S.E. Macko, eds, Plenum Press, New York, 511–534.

Pennington, W. and Cambray, R.S. (1973) Observations on lake sediments using fallout ^{137}Cs as a tracer. *Nature* 242, 324–326.

Peters, K.E. and Fowler, M.G. (2002) Applications of petroleum geochemistry to exploration and reservoir management. *Organic Geochemistry* 332(1), 5–36.

Philp, R.P., Allen, J., and Kuder, T. (2002) The use of the isotopic composition of individual compounds for correlating spilled oils and refined products in the environment with suspected sources. *Environmental Forensics* 3(3–4), 341–348.

Pond, K.L., Huang, Y., Wang, Y., and Kulpa, C.F. (2002) Hydrogen isotopic composition of individual n-alkanes as an intrinsic tracer for bioremediation and source identification of petroleum contamination. *Environmental Science and Technology* 36(4), 724–728.

Reddy, C.M., Heraty, L.J., Holt, B.D., Sturchio, N.C., Eglinton, T.I., Drenzek, N.J. *et al.* (2000) Stable carbon isotope compositions of Aroclors and Aroclor contaminated sediments. *Environmental Science and Technology* 34(13), 2866–2870.

Revesz, K., Coplen, T.B., Baedecker, M.J., and Glynn, P.D. (1995) Methane production and consumption monitored by stable C and H ratios at a crude oil spill site, Bemidji, Minnesota. *Applied Geochemistry* 10, 505–516.

Rice, A.L., Gotoh, A.A., Ajie, H.O., and Tyler, S.C. (2001) High-precision continuous-flow measurement of δ^{13}C and δD of atmospheric CH_4. *Analytical Chemistry* 73(17), 4104–4110.

Ricking, M., Koch, M., and Rotard, W. (2005) Organic pollutants in sediment cores of NE-Germany: Comparison of the marine Arkona Basin with freshwater sediments. *Marine Pollution Bulletin* 50(12), 1699–1705.

Richnow, H.H., Meckenstock, R.U., Reitzel, L.A., Baun, A., Ledin, A., and Christensen, T.H. (2003) *In situ* biodegradation determined by carbon isotope fractionation of aromatic hydrocarbons in an anaerobic landfill leachate plume (Vejen, Denmark). *Journal of Contaminant Hydrology* 64(1–2), 59–72.

Ritchie, J.C. and McHenry, J.R. (1990) Application of radioactive fallout cesium –137 for measuring soil-erosion and sediment accumulation rates and patterns—A review. *Journal of Environmental Quality* 19, 215–233.

Rogers, K.M. and Savard, M.M. (1999) Detection of petroleum contamination in river sediments from Quebec City region using GC–IRMS. *Organic Geochemistry* 30(12), 1559–1569.

Saber, D., Mauro, D., Philp, R.P., Allen, J., and Kuder, T., (2003) Carbon isotope ratios of PAHs in urban background soil. Battelle Meeting, Orlando, June 2003.

Saber, D., Mauro, D., Philp, R.P., and Allen, J.A. (2005) Chemical fingerprinting of sediments and soils near a former MPG site. *3rd International Conference on Remediation of Contaminated Sediments.* New Orleans, January 2005. Abs. only.

Samara, F., Tsai, C.W., and Aga, D.S. (2006) Determination of potential sources of PCBs and PBDEs in sediments of the Niagara River. *Environmental Pollution* 139(3), 489–497.

Schirmer, M. and Butler, B.J. (2004) Transport behavior and natural attenuation of organic contaminants at spill sites. *Toxicology* 205(3), 173–179.

Schlosser, P., Stute, M., Sonntag, C., and Munnich, K.O. (1989) Tritogenic ^{3}He in shallow ground water. *Earth and Planetary Science Letters* 94, 245–254.

Schmidt, T.C., Zwank, L., Elsner, M., Berg, M., Meckenstock, R.U., and Haderlein, S.B. (2004) Compound-specific stable isotope analysis of organic contaminants in natural environments: A critical review of the state of the art, prospects, and future challenges. *Analytical and Bioanalytical Chemistry* 378, 283–300.

Schoell, M. (1988) Multiple origins of methane in the earth. *Chemical Geology* 71, 1–10.

Schüth, C., Taubald, H., Bolaño, N., and Maciejczyk, K. (2003) Carbon and hydrogen isotope effects during sorption of organic contaminants on carbonaceous materials. *Journal of Contaminant Hydrology* 64(3–4), 269–281.

Sessions, A.L., Burgoyne, T.W., Schimmelmann, A., and Hayes, J.M. (1999) Fractionation of hydrogen isotopes in lipid biosynthesis. *Organic Geochemistry* 30(9), 1193–1200.

Sessions, A.L., Sylva, S., and Hayes, J.M. (2005) Moving-wire device for carbon isotopic analyses of nanogram quantities of nonvolatile organic carbon. *Analytical Chemistry* 77(20), 6519–6527.

Sherwood Lollar, B., Slater, G.F., Sleep, B., Witt, M., Kledka, G.M., Harkness, M., and Spivack, J. (2001) Stable carbon isotope evidence for intrinsic bioremediation

of tetrachloroethene and trichloroethene at area 6, Dover Air Force Base. *Environmental Science and Technology* 35(2), 261–269.

Shi, W., Sun, M.-Y., Molina, M., and Hodson, R.E. (2001) Variability in the distribution of lipid biomarkers and their molecular isotopic composition in Altamaha estuarine sediments: Implications for the relative contribution of organic matter from various sources. *Organic Geochemistry* 32(4), 453–467.

Shouakar-Stash, O., Frape, S.K., and Drimmie, R.J. (2003) Stable hydrogen, carbon and chlorine isotope measurements of selected chlorinated organic solvents. *Journal of Contaminant Hydrology* 60(3–4), 211–228.

Shouakar-Stash, O., Drimmie, R.J., and Frape, S.K. (2005) Determination of inorganic chlorine stable isotopes by conditions flow isotope ratio mass spectrometry. *Rapid Communications in Mass Spectrometry* 19, 121–127.

Slater, G.F., Sherwood Lollar, B., Sleep, B.E., and Edwards, E.A. (2001) Variability in carbon isotopic fractionation during biodegradation of chlorinated ethenes: Implications for field applications. *Environmental Science and Technology* 35(5), 901–907.

Smallwood, B.J., Philp, R.P., Burgoyne, T.W., and Allen, J. (2001) The use of stable isotopes to differentiate specific source markers for MTBE. *Journal of Environmental Forensics* 2, 215–221.

Smallwood, B.J., Philp, R.P., and Allen, J. (2002) Stable carbon isotopic composition of gasolines determined by isotope ratio monitoring gas chromatography mass spectrometry. *Organic Geochemistry* 33(2), 149–159.

Smirnov, A., Abrajano, T.A., Smirnov, A., and Stark, A. (1998) Distribution and sources of polycyclic aromatic hydrocarbons in the sediments of Lake Erie, Part 1. Spatial distribution, transport, and deposition. *Organic Geochemistry* 29(5–7), 1813–1828.

Somsamak, P., Cowan, R.M., and Haggblom, M.M. (2001) Anaerobic biotransformation of fuel oxygenates under different anoxic conditions. *FEMS Microbiology Ecology* 37, 259–264.

Stahl, W.J. (1977) Carbon and nitrogen isotopes in hydrocarbon research and exploration. *Chemical Geology* 20, 121–149.

Stark, A., Abrajano, T., Hellou, J., and Metcalf-Smith, J.L. (2003) Molecular and isotopic characterization of polycyclic aromatic hydrocarbon distribution and sources at the international segment of the St. Lawrence River, *Organic Geochemistry* 34(2), 225–237.

Stout, S.A., Uhler, A.D., and Emsbo-Mattingly, S.D. (2004) Comparative evaluation of background anthropogenic hydrocarbons in surficial sediments from nine urban waterways. *Environmental Science and Technology* 38(11), 2987–2994.

Sturchio, N.C., Clausen, J.L., Heraty, L., Huang, L., Holt, B.D., and Abrajano, T. (1998) Chlorine isotope investigation of natural attenuation of trichloroethene in an aerobic aquifer. *Environmental Science and Technology* 32(20), 3037–3042.

Sturchio, N.C., Hatzinger, P.B., Arkins, M.D., Suh, C., and Heraty, L.J. (2003) Chlorine isotope fractionation during microbial reduction of perchlorate. *Environmental Science and Technology* 37(17), 3859–3863.

Tamura, T. and Jacobs, D.G. (1960) Structural implications in cesium sorption. *Health Physics* 2, 391–398.

Thompson, A.E., Anderson, R.S., Rudolph, J., and Huang, L. (2002) Stable carbon isotope signatures of background tropospheric chloromethane and CFC113. *Biogeochemistry* 60(2), 191–211.

Thornton, S.F., Lerner, D.N., and Banwart, S.A. (2001) Assessing the natural attenuation of organic contaminants in aquifers using plume-scale electron and carbon balances: Model development with analysis of uncertainty and parameter sensitivity. *Journal of Contaminant Hydrology* 53(3–4), 199–232.

Trust, B.A., Coffin, R.B., Cifuentes, L.A., and Mueller, J.G. (1995) In *Monitoring and Verification of Bioremediation*, R.E. Hinchee, G.S. Douglas, and S.K. Ong, eds. Battelle Press, Colombia, OH, 233–239.

van Metre, P.C., Callendar, E., and Fuller, C.C. (1997) Historical trends in organochlorine compounds in river basins identified using sediment cores from reservoirs. *Environmental Science and Technology* 31(8), 2339–2344.

Vilomet, J.D., Veron, A., Ambrosi, J.P., Moustier, S., Bottero, J.Y., and Chatelet-Snidaro, L. (2003) Isotopic tracing of landfill leachates and pollutant lead mobility in soil and groundwater. *Environmental Science and Technology* 37(20), 4586–4591.

Von Gunten, H.R., Strum, M., and Moser, R.N. (1997) 200-year record of metals in lake sediments and natural background concentrations. *Environmental Science and Technology* 31(8), 2193–2197.

Wang, X-C., Altabet, M.A., Callahan, J., and Chen, R.F. (2004) Stable carbon and nitrogen isotopic compositions of high molecular weight dissolved organic matter from four U.S. estuaries. *Geochimica and Cosmochimica Acta* 68(12), 2681–2691.

Wang, Y., Huang, Y., Huckins, J.N., and Petty, J.D. (2004) Compound-specific carbon and hydrogen. *Environmental Science and Technology* 38(13), 3689–3697.

Wang, Z. and Christensen, J.H. (2005) Crude oil and refined product fingerprinting: Applications, Chapter 17 in *Environmental Forensics: A Contaminant Specific*

Guide, Robert D. Morrison and Brian L. Murphy, eds. Academic Press, Burlington, MA.

Wang, Z., Fingas, M., and Sergy, G. (1994) Study of 22-year old Arrow oil samples using biomarker compounds by GCMS. *Environmental Science and Technology* 28, 1733–1746.

Wang, Z., *et al*. (2005) Crude oil and refined product fingerprinting: Principles, Chapter 16 in *Environmental Forensics: A Contaminant Specific Guid*e, Robert D. Morrison and Brian L. Murphy, eds. Academic Press, Burlington, MA.

Ward, J.A.N., Ahad, J.M.E., Lacrampe-Couloume, G., Slater, G.F., Edwards, E.A., and Sherwood Lollar, B. (2000) Hydrogen isotope fractionation during methanogenic degradation of toluene: Potential for direct verification of bioremediation. *Environmental Science and Technology* 34(21), 4577–4581.

Warmerdam, E.M., Frape, S.K., Aravena, R., Drimmie, R.J., Flatt, H., and Cherry, J.A. (1995) Stable chlorine and carbon isotope measurements of selected chlorinated organic solvents. *Applied Geochemistry* 10(5), 547–552.

Wassenaar, L.I. and Hobson, K.A. (2000) Improved method for determining the stable-hydrogen isotopic composition (δD) of complex organic materials of environmental interest. *Environmental Science and Technology* 34(11), 2354–2360.

Wassenaar, L. and Koehler, G. (2004) On-line technique for the determination of the ^{37}Cl of inorganic and total organic Cl in environmental samples. *Analytica Chemistry* 76, 6384–6388.

Wen, S., Feng, Y., Yu, Y., Bi, X., Wang, X., Sheng, G., Fu, J., and Peng, P. (2005) Development of a compound-specific isotope analysis method for atmospheric formaldehyde and acetaldehyde. *Environmental Science and Technology* 39(16), 6202–6207.

Whiticar, M.J., Faber, E., and Schoell, M. (1986) Biogenic methane formation in marine and freshwater environments: CO_2 reduction vs. acetate fermentation-isotopic evidence. *Geochimica and Cosmochimica Acta* 50, 693–709.

Wiesenberg, G.L.B., Schwarzbauer, J., Schmidt, M.W.I., and Schwark, L. (2004) Source and turnover of organic matter in agricultural soils derived from n-alkane/n-carboxylic acid compositions and C-isotope signatures. *Organic Geochemistry* 35(11–12), 1371–1393.

Wilcke, W., Krauss, M., and Amelung, W. (2002) Carbon isotope signature of polycyclic aromatic hydrocarbons (PAHs): Evidence for different sources in tropical and temperate environments? *Environmental Science and Technology* 36(16), 3530–3535.

Wilkes, H., Boreham, C., Harms, G., Zengler, K., and Rabus, R. (2000) Anaerobic degradation and carbon isotopic fractionation of alkylbenzenes in crude oil by sulphate-reducing bacteria. *Organic Geochemistry* 31(1), 101–115.

Wong, C.S., Sanders, G., Engstrom, D.R., Long, D.T., Swackhamer, D.L., and Eisenreich, S.J. (1995) Accumulation, inventory, and diagenesis of chlorinated hydrocarbons in Lake Ontario sediments. *Environmental Science and Technology* 29(10), 2661–2672.

Yanik, P.J., O'Donnell, T.H., Macko, S.A., Qian, Y., and Kennicutt, II M.C. (2003) Source apportionment of polychlorinated biphenyls using compound specific isotope analysis. *Organic Geochemistry* 34(2), 239–251.

Zwank, L., Schmidt, T.C., Handerlein, S.B., and Berg, M. (2002) Simultaneous determination of fuel oxygenates and BTEX using direct aqueous injection gas chromatography/mass spectrometry (DAI-GC/MS). *Environmental Science and Technology* 36, 2054–2059.

Zwank, L., Berg, M., Schmidt, T.C., and Haderlein, S.B. (2003) Compound-specific carbon isotope analysis of volatile organic compounds in the low-microgram per liter range. *Analytical Chemistry* 75, 5575–5583.

Zwank, L., Berg, M., Elsner, M., Schmidt, T.C., Schwarzenbach, R.P., and Haderlein, S.B. (2005) New evaluation scheme for two-dimensional isotope analysis to decipher biodegradation processes: Application to groundwater contamination by MTBE. *Environmental Science and Technology* 39(18), 7344–7344.

CHAPTER 11

FORENSIC APPLICATIONS OF CONTAMINANT TRANSPORT MODELS IN THE SUBSURFACE

Ashok Katyal and Robert D. Morrison

11.1 Introduction 514
11.2 Contaminant Chemistry 514
11.2.1 Henry's Law Constant (K_H) 514
11.2.2 Liquid Density 515
11.2.3 Viscosity 515
11.2.4 Vapor Pressure and Density 517
11.2.5 Solubility and Nonaqueous Phase Liquids (NAPLs) 517
11.3 Processes Affecting Contaminant Chemistry 520
11.3.1 Hydrolysis 520
11.3.2 Sorption and Retardation 520
11.3.3 Biodegradation 521
11.4 Application of Forensic Models 522
11.4.1 Fluid Transport through a Paved Surface and Compacted Soil 522
11.4.2 Vapor Transport through a Paved Surface 523
11.4.3 Liquid Transport through a Paved Surface 526
11.4.4 Fluid Transport through Soil 529
11.4.4.1 Vapor Transport 529
11.4.4.2 Liquid Transport 530
11.4.4.3 SESOIL 531
11.4.4.4 VLEACH 532
11.4.5 Contaminant Transport Models in Soil 533
11.5 Contaminant Transport Models for Groundwater 536
11.5.1 Inverse Models 538
11.6 Conclusions 558
References 558

11.1 INTRODUCTION

Identifying when a contaminant release occurred and its origin is often pivotal for determining liability and for apportioning costs between responsible parties (Morrison and Murphy, 2005). Contaminant transport models for fluids (vapor and liquid) often provide the only basis to provide this information.

This chapter provides an overview of the types of contaminant transport models available to estimate the timing and age of a contaminant migrating through a paved or unpaved surface, the soil column, and/or in groundwater. For a detailed examination of the mathematics associated with these various approaches, we encourage you to examine the foundational texts on this subject (Carslaw and Jaeger, 1959; Scheidegger, 1960; Bear, 1979; Sposito *et al.*, 1979; Neuman, 1990; Anderson and Woessner, 1992).

11.2 CONTAMINANT CHEMISTRY

An understanding of the physicochemical properties of a contaminant is necessary when applying and/or evaluating environmental forensic techniques for age dating and source identification. In cases of multiple releases, a measurement of physical and chemical properties evaluated via multivariate analysis may provide a straightforward basis for source discrimination (Trevizo *et al.*, 2000). Key contaminant properties evaluated and/or measured include the Henry's law constant, liquid density, solubility, viscosity, vapor pressure, vapor density, and boiling point. Physical and chemical reactions affecting contaminant chemistry include hydrolysis, sorption (adsorption and absorption), and biodegradation. An understanding of these physicochemical properties and contaminant reactions is important for evaluating age dating and source identification techniques that rely on these processes.

11.2.1 HENRY'S LAW CONSTANT (K_H)

The *Henry's law constant* (K_H) (also called the *air–water partition coefficient*) is the ratio of a compound's partial pressure in air to the concentration of the compound in water at a given temperature. Values for Henry's law constants are expressed in units of atmospheres for air to moles per cubic meter for water (atm-m^3/mol) or in a dimensionless unit described as $K_{H'} = K_H/(RT)$ where $K_{H'}$ is the dimensionless Henry's law constant, K_H is the Henry's law constant (atm-m^3/mol), R is the ideal gas constant (8.20575×10^{-5} atm-m^3/mol-K) and T is the water temperature (K). As a rule of thumb, compounds

with a Henry's law constant greater than 10^{-3} atm-m^3/mol and a molecular weight less than 200 grams per mole are considered volatile (United States Environmental Protection Agency, 1996).

A compound with a Henry's law constant less than about 5×10^{-5} atm-m^3/mol is considered soluble and tends to remain in water (Olson and Davis, 1990). Ranges of Henry's law constant values in (atm-m^3/mol) for selected chlorinated solvents are trichloroethylene (1.2×10^{-2} to 9.9×10^{-3}), tetrachloroethylene (1.3×10^{-2} to 2.9×10^{-3}), carbon tetrachloride (2.3×10^{-2} to 3.0×10^{-2}), methylene chloride (2.0×10^{-3} to 2.9×10^{-3}), and 1,1,1-trichloroethane (1.3×10^{-2} to 1.8×10^{-2}) (Hine and Mookerjee, 1975; United States Environmental Protection Agency, 1980; Lyman *et al.*, 1982; Roberts and Dandliker, 1983; Lincoff and Gossett, 1984; Warner *et al.*, 1987; Pankow and Rosen, 1988; Schwille, 1988; Mercer and Cohen, 1990; Montgomery, 1991).

11.2.2 LIQUID DENSITY

The *specific gravity* of a liquid is the ratio of its specific weight to distilled water at 4°C. The fluid density of most organic compounds is greater than 1 g cm^{-3}. Compounds with densities greater than 1.0, relative to water (e.g., perchloroethylene, trichloroethylene, polychlorinated biphenyls, bromoform) have a greater probability of penetrating into the groundwater when introduced as a phase separate liquid. Table 11.1 lists selected contaminants, their chemical formulas, and liquid densities (Morrison *et al.*, 1998; Ramamoorthy and Ramamoorthy, 1998; Morrison, 1999a).

Compared to other physical properties used in contaminant models, water density changes little. Under a pressure of 0.1 MPa, the density of water is at its maximum at 3.98°C and has a density of 1000 kgm^{-3}, which is most often assumed in contaminant transport models even though water at 25°C under atmospheric pressure has a density of 997.02 kgm^{-3}, a relative difference of 2.98% from its maximum value (Grant, 2000). The density of a contaminant as it is transported through the subsurface can change. The density of a solvent mixture, for example, can change due to the selective dissolution of the compounds into soil or rock. In most cases, the mixture density increases.

11.2.3 VISCOSITY

Viscosity characterizes a fluid's resistance to flow. The viscosity of water increases exponentially with decreasing temperature and is affected by the type and concentration of solutes (Vand, 1948). *Kinematic viscosity* describes the absolute viscosity of the substance divided by its density. A high-density

Table 11.1
Chemical formula and liquid density for selected chlorinated solvents.

Chlorinated Solvent	Formula	Liquid Density (g/cm^3) at 20°C
1,2-Dibromomethane (EDB)[a]	$BrCH_2CH_2Br$	2.18
Tetrachloroethylene (PCE)	CCl_2CCl_2	1.63
1,1,2,2-Tetrachloroethane	$C_2H_2Cl_4$	1.60
Carbon tetrachloride (CT)	CCl_4	1.59
1,1,1,2-Tetrachloroethane	$C_2H_2Cl_4$	1.54
Trichlorofluoromethane (Freon 11)	CCl_3F	1.49
Chloroform[b]	$CHCl_3$	1.49
Trichloroethylene (TCE)	$CHCl:CCl_2$	1.46
1,1,2-Trichloroethane	$C_2H_3Cl_3$	1.44
1,1,1-Trichloroethane (TCA)	$C_2H_3Cl_3$	1.35
Dichloromethane (DCM)[c]	CH_2Cl_2	1.34
Dichlorodifluoromethane (Freon 12)	CCl_2F_2	1.33
cis-1,2-Dichloroethylene	$C_2H_2Cl_2$	1.28
trans-1,2-Dichloroethylene	$C_2H_2Cl_2$	1.26
1,2-Dichloropropane	$CH_3CHClCH_2Cl$	1.16
1,1-Dichloroethane	$C_2H_4Cl_2$	1.1
Chloroethane	C_2H_5Cl	0.92
Vinyl chloride (VC)	C_2H_3Cl	0.91

[a]Also ethylene bromide and ethylene dibromide.
[b]Also trichloromethane.
[c]Also methylene chloride.

Table 11.2
Kinematic viscosity of selected solvents.

Chemical	Kinematic Viscosity (centistokes)
1,2-Dibromomethane	0.79
1,2-Dichloropropane	0.75
1,1,1-Trichloroethane	0.62
Carbon tetrachloride	0.61
Tetrachloroethylene	0.54
Trichloroethylene	0.39
cis-1,2-Dichloroethylene	0.38
Chloroform	0.38
Methylene chloride	0.32
trans-1,2-Dichloroethylene	0.32

liquid with a low viscosity has a low kinematic viscosity; such a fluid flows quickly through a porous medium compared with one with a higher kinematic viscosity.

Fluid velocity through porous media is approximated as inversely proportional to the kinematic viscosity. A decrease in viscosity therefore increases the velocity of a compound through porous media. The kinematic viscosity of selected chlorinated solvents is summarized in Table 11.2 (Huling and Weaver, 1991).

Compound	Vapor Pressure (mm at 20°C)	Vapor Density (g/L)
Freon 11	687.0	5.61
Methylene chloride	348.9	3.47
Freon 113	270.0	7.66
1,1,1-Trichloroethane (TCA)	90.0	5.45
Trichloroethene (TCE)	57.8	5.37
Tetrachloroethene (PCE)	14.0	6.86

Table 11.3
Vapor pressure and density of selected chlorinated solvents.

Compound	Literature Solubility (mg/L)
Methylene chloride (MC)	20,000
cis-1,2-Dichloroethylene	3,500
1,2-Dibromomethane (EDB)	4,200
trans-1,2-Dichloroethylene	6,300
Trichloroethylene (TCE)	1,100
1,1,1-Trichloroethane (TCA)	1,300
Tetrachloroethylene (PCE)	200
Carbon tetrachloride (CTC)	825

Table 11.4
Solubility of selected chlorinated solvents.

11.2.4 VAPOR PRESSURE AND DENSITY

The *vapor pressure* and *density* of a contaminant, especially chlorinated solvents, are important when modeling solvent transport as a vapor cloud. The vapor pressure and density of selected chlorinated solvents are summarized in Table 11.3 (Morrison, 1999a).

11.2.5 SOLUBILITY AND NONAQUEOUS PHASE LIQUIDS (NAPLs)

The *solubility* of a compound is its saturated concentration in water at a known temperature and pressure. In general, the higher the water solubility, the more likely it is for the compound to be mobile while being less accumulative, bio-accumulative, volatile, and persistent. The solubility of several pure phase chlorinated solvents in water at 25°C is summarized in Table 11.4 (Pankow *et al.*, 1996; Morrison, 1999a).

The *effective solubility* of a multicomponent solvent in water depends on the composition of the mixture. A constituent's solubility within this multicomponent mixture may be orders of magnitude lower than the aqueous solubility of the pure chemical in water (Odencrantz *et al.*, 1992). The vapor pressure and composition in equilibrium with a solution can yield valuable information regarding the thermodynamic properties of the liquids involved. Raoult's law

relates the vapor pressure of components to the composition of the solution and assumes ideal behavior. It gives a simple picture of the situation just as the ideal gas law does. As the interactive forces between molecules and the volume of the molecules approach zero, the behavior of gases approaches the behavior of the ideal gas. Raoult's law assumes that the physical properties of the components are identical. The more similar the components, the more their behavior approaches the ideal gas described by Raoult's law.

An issue when developing a conceptual contaminant transport model is whether the contaminant migrated as a *nonaqueous phase liquid* (NAPL) as this will influence the appropriateness of the model selection. Techniques available to provide this information include direct and indirect methods. A direct indication is the reaction of a soil sample containing a NAPL to a hydrophobic dye powder such as Sudan IV or inferentially using induced fluorescence measurements of coconstituents dissolved in the NAPL (Kram, 1998; Rossabi and Nave, 1998; Balshaw-Biddle *et al.*, 1999). The visual detection of a DNAPL has also been accomplished using a high-resolution video microscope mounted on a cone penetrometer rig (CPT) (Lieberman and Knowles, 1998). This technology can image colored NPALs in the soil pores next to the CPT rod. Another direct technique is the use of a Membrane Interface Probe (MIP), which is mounted on the push rod of a drilling rig. The MIP approach uses a selectively permeable membrane that permits the entry of only volatile organic compounds, which can be used to provide a continuous profile of VOCs through depth (Christy, 1998). An optical probe installed in a hardened window of a cone penetrometer test push rod had been used to detect DNAPLs either directly or by Raman spectroscopy.

Another direct approach is the Ribbon NAPL Sampler, which is used in conjuction with direct push or other drilling methods (Keller *et al.*, 2000; Riha *et al.*, 2000). The Ribbon Sampler consists of a hydrophobic sorbent liner with an impregnated indicator dye that is deployed in a borehole and directly contacts the soil. DNAPL is wicked into the liner and causes a color change in the dye. The liner is then retrieved from the borehole and colored spots and their depths are logged to identify the presence and depth of the NAPL.

A method relying upon direct measurements assumes that if soil properties and contaminant concentrations are known and equilibrium partitioning is assumed, the presence of a NAPL can be determined. An illustration of this approach is to use PCE, TCE, and TCA and assume a soil with a total porosity of 40%, a water-filled porosity of 15%, a soil organic content ranging from 0.1 to 2%, and a soil dry bulk density of 1.59 g/cm^3. Given this information, PCE, TCE, and TCA concentrations indicative of a NAPL are in the range 117–744, 302–1300, and 366–1408 mg/kg, respectively, for different K_{oc} values.

When the solvent concentration exceeds these values, the contaminant mass in the soil exceeds the capacity of the soil, thereby indicating the presence of a NAPL. If the concentration is significantly less than the estimated concentration range, there is an indication that the contaminant did not enter the soil as a NAPL.

Absent direct observation of a NAPL, several indirect methods are available. One technique is to use the concentration of a dissolved phase single-component NAPL at a threshold concentration as an indicator. In the 1980s, the rule of thumb was that if a dissolved concentration of 10% of the solubility of the compound was detected, the presence of a NAPL was inferred (Feenstra and Cherry, 1998). In the early 1990s, research indicated that concentrations of 1% or more of a compound's solubility constitute a high likelihood of the presence of a NAPL (United States Environmental Protection Agency, 1992, 1993; Cohen *et al.*, 1993). The current convention is consistent with the 1% rule (Newell and Ross, 1991; Pankow *et al.*, 1996; McNab *et al.*, 1999).

Soil vapor concentrations can provide an indidrect means to identify the presence of a NAPL. A soil gas concentration of 100 parts per million (ppm) by volume commonly is considered to be indicative of the presence of NAPL. Another field technique for NAPL identification is the partitioning tracer method originally developed to measure the residual oil in the subsurface by petroleum engineers (Nelson *et al.*, 1999). The partitioning tracer method compares the transport of one or more partitioning tracers, such as calcium bromide, with a nonreactive tracer. Given that conservative tracers have a partition coefficient of zero relative to a NAPL whereas partitioning tracers have a nonzero partition coefficient relative to the NAPL, judgments regarding the presence of a NAPL between the injection and recovery wells are possible (Deeds *et al.*, 2000). The retardation and residual NAPL saturation is inferred from breakthrough curves at a monitoring well downgradient of the injected tracer if the NAPL/water partition coefficient for the partitioning tracer is known (Jin *et al.*, 1995, 1997).

The measurement of naturally occurring radon-222 in groundwater is proposed as a naturally occurring tracer for detecting and quantifying NAPL contamination (Hunkeler *et al.*, 1997; Semprini *et al.*, 2000). In the absence of a NAPL, radon concentrations in groundwater reach a site-specific equilibrium value based on the mineral fraction of the aquifer solids. In the presence of NAPL, the radon concentrations are preferentially reduced due to the preferential partitioning of radon in the organic NAPL phase. At the Borden test site in Ontario, Canada, a mixture of chloroform, TCE, and PCE (perchloroethylene) was injected into the shallow sand aquifer. Groundwater samples were collected from monitoring wells upgradient, within and downgradient of the NAPL and analyzed for radon (Feenstra and Cherry, 1998).

The radon concentrations in the NAPL zone decreased by a factor of two to three times compared to values observed in upgradient and downgradient monitoring well samples. Precautions when using the radon method include the significant impact of heterogeneities in aquifer porosity and radon emanations. In the absence of a NAPL, a decrease in porosity results in higher radon concentrations in the pore fluid, thus complicating data interpretation.

Geophysical techniques used to identify NAPLs include electrical impedance tomography (EIT) (Daily *et al.*, 1998), seismic reflection profiling (Brewster *et al.*, 1995; Waddell *et al.*, 1996), cross radar technology (Wright *et al.*, 1998), negative ion sensors (United States Environmental Protection Agency, 1998), and complex resistivity techniques. Geophysical techniques when combined with other methods can be applied to provide a higher confidence level regarding the presence of a NAPL. At Hill Air Force Base in Ogden, Utah, for example, electrical impedance tomography was used with fiber-optic chemical sensors and neutron logs to verify the presence of NAPLs (United States Environmental Protection Agency, 1998).

11.3 PROCESSES AFFECTING CONTAMINANT CHEMISTRY

11.3.1 HYDROLYSIS

Hydrolysis is the chemical breakdown of substances by water and depends on the chemistry, solubility, pH, and the oxidation–reduction (E_h or redox) potential of compound. Hydrolysis rates in soil may be different than rates reported in water, depending on the effects of pH, redox, sorption, and surface catalyzed reactions (Dilling *et al.*, 1975; Mabey and Mill, 1978). The degradation of 1,1,1-TCA (trichloroethane) to 1,1-DCE (dichloroethene) is a well-known hydrolysis reaction that provides a potential means to age date the release of the TCA (Gauthier and Murphy, 2003).

11.3.2 SORPTION AND RETARDATION

Retardation values for a compound used in contaminant transport models can exercise a significant impact on the results of the simulations. Retardation is a generic term that includes sorption of a compound to the material through which it migrates. Sorption encompasses adsorption, absorption, ion exchange, and chemisorption, and can be broadly defined as the process that effects the transient development of concentration profiles in the subsurface (Ham, 2006). Absorption (the penetration of substances into the bulk of a solid or liquid) and adsorption (the surface retention of a solid, liquid, or gas mole-

cules by a solid or liquid) are the most important of these sorption processes. The sorption capacity of a compound is described by its *sorption coefficient*. The sorption coefficient is the ratio of an adsorbed chemical per unit weight of organic carbon to its concentration in water. The sorption coefficient usually is estimated via batch isotherm experiments. Predicted correlations are also available that relate the distribution coefficient of a compound to the soil organic content, as expressed by the octanol–water partition coefficient (Fetter, 1994). In one study, predicted retardation coefficients are two to five times lower than measured values (Ball and Roberts, 1991).

Retardation values generally increase with increasing fractions of organic carbon and/or clay. A range of retardation values for PCE in a sand and gravel aquifer, for example, is between 1 (no retardation) and 5 (Schwarzenbach *et al.*, 1983; Barber *et al.*, 1988). The retardation value for TCE is reported as less than 10; values between 1 and 2.5 are commonly used in contaminant transport models.

Apparent retardation values can be biased as a function of well design and sampling technique. Apparent retardation rates in one study were found to be inconsistent between monitoring wells depending on the saturated screen length, the degree of screen desaturation during purging, and the longitudinal distance from the contaminant source (Robbins, 1989). The selection of a retardation value for an entire well field may therefore be inappropriate due to differences in well construction. A single retardation value may similarly be overly simplifying in cases where concentration averaging is used.

11.3.3 BIODEGRADATION

Biodegradation refers to the transformation of compounds, principally by bacteria, fungi, and yeast, into simpler substances. Bacteriological or reductive dechlorination is the principal process resulting in the degradation of chlorinated solvents in the subsurface (Gao *et al.*, 1995). For most biodegradation pathways, the final degradation products are carbon dioxide and water (after Vogel *et al.*, 1987a,b; McCarty, 1993, 1994; Butler and Hayes, 1999; Morrison *et al.*, 1999a; Morrison, 2000a,b,c; Morrison and Murphy, 2005).

The dichloroethene (DCE) isomers for TCE anaerobic degradation include 1,1-DCE, *cis*-1,2-DCE and *trans*-1,2-DCE. Of these isomers, *cis*-1,2-DCE is produced in the greatest abundance at a rate of about 30 times that of *trans*-1,2-DCE. The decay rate from TCE to *cis*-1,2-DCE is also faster than from TCE to 1,1-DCE. The decay rate from *cis*-1,2-DCE to vinyl chloride (chloroethene), however, is slower than from *trans*-1,2-DCE and 1,1-DCE to vinyl chloride. Under some conditions, vinyl chloride is degraded to acetate.

Another example of reductive dechlorination is the degradation pathway of trichlorofluoroethene (TCFE). Similarities in the degradation kinetics of TCFE, TCE, and PCE, when present at comparable initial concentrations, offer the opportunity of using TCFE and its degradation products as benign reactive tracers to measure *in situ* degradation rates of PCE and TCE. This relationship may provide evidence in support of a chlorinated solvent degradation model used for age dating and/or source identification.

Anaerobic degradation occurs via dehalogenation with the release of a chlorine atom. The rate of reductive dehalogenation decreases as more chlorine is removed (Woodbury and Ulrych, 1996). Strong reducing conditions with low dissolved oxygen levels are conducive to dehalogenation.

Aerobic degradation is initiated by enzymes called oxygenases (Tsien *et al.*, 1989). These enzymes initiate degradation after producing a specific organic compound that serves as a carbon and energy source for the bacteria. The rate and ability of microbes to degrade a compound are dependent on the ability of the subsurface environment to support a healthy community of microbes. The necessary nutrients, oxygen content, and soil and/or groundwater temperature required to sustain viable microbial communities are highly variable. Biodegradation may not occur if the concentration of the compound is low, although many can be degraded to some extent.

A recent contaminant of concern in the United States is methyl tertiary butyl ether (MTBE). Biodegradation studies of MTBE in shallow aquifers indicate that MTBE is biodegradable under anaerobic and aerobic environments and may account for the presence of TBA (tert-butyl alcohol) at many sites (Bradley *et al.*, 2001a,b; Kane *et al.*, 2001). A potential difficulty is that while tertiary butyl alcohol (TBA) is a primary metabolite of MTBE via atmospheric and microbial oxidation, it is also a byproduct of MTBE synthesis and is often present as an impurity (0.1–1%) (Drogos, 2000). TBA may also be added to MTBE up to 5% by volume (Brown, 2000). The detection of TBA with MTBE may therefore be independent of its degradation relationship with MTBE (Wilson *et al.*, 2000). The use of stable isotopes (^{13}C) may provide a means to distinguish between the TBA produced from different MTBE sources or its relationship as a degradation product of MTBE (Wilson *et al.*, 2005).

11.4 APPLICATION OF FORENSIC MODELS

11.4.1 FLUID TRANSPORT THROUGH A PAVED SURFACE AND COMPACTED SOIL

Models describing liquid transport through paved surfaces are highly specific to the chemical and physical composition of the paved surface and the physicochemical properties of the contaminant. In order to identify the most likely

contaminant transport mechanism (e.g., liquid advection, gas diffusion, liquid diffusion and/or evaporation) to be simulated, a realistic physical and mathematical model is required. The most commonly encountered models through paved surfaces in environmental litigation are liquid and vapor transport models.

A frequent area of inquiry is whether a liquid migrated through a paved or compacted surface (e.g., asphalt, concrete, crushed rock, or compacted soil), and if so, the time required (Morrison, 2000a). Absent direct measurements, contaminant transport equations are often the only method available to estimate the time required for a fluid to migrate through a paved or compacted soil.

11.4.2 VAPOR TRANSPORT THROUGH A PAVED SURFACE

The transport of a *vapor cloud* through a paved surface is a common area of inquiry, especially for volatile organic compounds originating from vapor degreasers. Vapor degreasers and dry cleaning equipment often are installed in a concrete catch basin or trough. Catch basins can exacerbate the potential for vapor transport because they accumulate the vapor and can minimize vapor dilution with the ambient air.

If vapor and liquid releases are probable, it may be difficult to determine the more realistic transport scenario. Existing test results may provide this information. One example indicative of vapor transport is the detection of a compound in soil vapor that is absent in the soil. An example is the detection of MTBE in soil vapor that is absent in soil at the same location but is present in the underlying groundwater.

Vapor diffusion through a paved surface can be more rapid than liquid transport. Ideally, the following information is available for developing a vapor transport model through a paved surface:

- Vapor density of the primary compound and/or liquid mixture
- Whether the vapor source is constant or transient above the pavement
- The Henry's law constant for the contaminant
- A measured gas permeability of the paved surface (American Petroleum Institute Method RP40)
- Paved surface thickness, porosity, and moisture content
- Vapor concentration above the pavement
- The vapor concentration within and below the paved surface prior to the spill

Vapor density is a key variable and is approximately equal to the molecular weight (MW) of the fluid divided by the molecular weight of air (<29). The

vapor density of common chlorinated solvents relative to air is TCA (4.6), TCE (4.5), PCE (5.7), and vinyl chloride (2.1).

Numerous diffusion equations are available to estimate the velocity of a vapor through paved surface for different boundary conditions (Crank, 1985; McCoy and Rolston, 1992; Choy and Reible, 1999). A general expression for the unsteady, diffusive radial flow of vapor from a contaminant source is described by:

$$\partial^2 C_a/\partial r^2 + [1/r(\partial C_a/\partial r)] = (R_a/D^*)(\partial C_a/\partial t) \tag{11.1}$$

where the air-filled porosity (n_a) of the paved surface is assumed to be constant, R_a is the vapor retardation coefficient, C_a is the computed concentration of the vapor in air, r is the source radius, and D^* is the effective diffusion coefficient (for TCE = 3.2×10^{-6} m^2 per second and 0.072 cm^2 per second for PCE) (Millington and Quirk, 1959) that is equal to:

$$D^* = D\tau_a \tag{11.2}$$

where $\tau_a = n_a^{3.333}/n_t^2$ and n_i is the total paved surface porosity, which is the sum of the air-filled porosity and water-filled porosity. The paved surface vapor retardation factor (R_a) in Equation 11.1 is equal to:

$$R_a = 1 + n_w/(n_a K_H) + \rho_b K_d/(n_a K_H) \tag{11.3}$$

where n_w is the water-filled porosity, n_a is the air-filled paved surface porosity, P_b is the paved surface bulk density, K_d is the distribution coefficient, and K_H is the dimensionless Henry's law constant.

Another conceptual option is to consider vapor diffusion through a medium (concrete) bounded by two parallel plates. This approach assumes a paved surface thickness (l) and a diffusion coefficient (D) whose surfaces $x = 0$ and $x = l$ are maintained at a constant concentration specified as C_1 and C_2, respectively.

After a time, steady state conditions are reached and concentrations are constant at all locations in the pavement. The diffusion equation in one dimension reduces to (Crank, 1985):

$$d^2C/dx^2 = 0 \tag{11.4}$$

which assumes that the diffusion coefficient (D) is constant. On integrating with respect to x, the following expression arises:

$$dC/dx = \text{constant} \tag{11.5}$$

and by introducing the conditions at $x = 0$, $x = l$, and integrating, then:

$$[C - C_1/C_2 - C_1] = x/l \tag{11.6}$$

Equation 11.6 describes a linear concentration change from C_2 to C_1 through the pavement. The rate of diffusion (F) through the paved surface is assumed to be the same and is described by:

$$F = -D\mathrm{d}C/\mathrm{d}x = -D(C_2 - C_1)/l \tag{11.7}$$

If the paved surface thickness (l) and the surface concentrations C_1 and C_2 are known, D is deduced from an observed value of F with Equation (11.7). A value for the diffusion coefficient can also be measured experimentally.

If the paved surface $x = 0$ is maintained at a constant concentration C_1 and $x = l$, evaporation into the atmosphere is assumed for which the equilibrium concentration immediately within the paved surface is C_2. Assuming a mass transfer coefficient h, then:

$$\mathrm{d}C/\mathrm{d}x + h(C - C_2) = 0, \quad x = l \tag{11.8}$$

and

$$(C - C_1)/(C_2 - C_1) = (hx)/(1 + hl) \tag{11.9}$$

and

$$F = \mathrm{D}h(C_1 - C_2)/(1 + hl) \tag{11.10}$$

If the surface conditions of the paved surface is defined as:

$$\mathrm{d}C/\mathrm{d}x + h_1(C_1 - C) = 0, \quad x = 0 \tag{11.11}$$

and

$$\mathrm{d}C/\mathrm{d}x + h_2(C - C_2) = 0, \quad x = l \tag{11.12}$$

then

$$C = [h_1C_1\{1 + h_2(l - x)\} + h_2C_2(1 + h_1x)]/[h_1 + h_2 + h_1h_2l] \tag{11.13}$$

and

$$F = \mathrm{D}h_1h_2(C_1 - C_2)/[h_1 + h_2 + h_1h_2l] \tag{11.14}$$

Although the most effective challenges to liquid or vapor transport models through paved surface are site specific, generic categories to these mathematics include the following (Morrison, 2000b):

- Variability of the model parameters (e.g., are they measured values or reasonable estimates?)
- The accuracy of the known circumstances of the spill event(s)
- The environmental conditions at the time of the release(s) (e.g., thickness of the liquid, composition of the liquid, duration of the release, ambient air temperature, etc.)
- Consistency of the modeled results with measured contaminant concentrations under the paved surface (if available)
- The impact of potential short circuiting pathways such as expansion joints and/or cracks on the conceptual model
- If a physical sample of the paved surface is obtained and the physical properties tested, the consistency of the input parameters used in the model with these direct measurements

11.4.3 LIQUID TRANSPORT THROUGH A PAVED SURFACE

Liquid transport through a paved surface is commonly believed to be a rapid process. This assumption is true if the paved surface is cracked, allowing unrestricted flow through the paved surface, or if the spill occurs over an expansion or isolation joint filled with a permeable such as wood, oakum, felt, or tar. Expansion joints usually are located at the junction of floors and walls, foundation columns, and footings. Isolation joints include premolded joint fillers (Kosmatka and Panarese, 1988) and extend the full depth of the slab. Given the sorptivity of expansion and isolation joint materials, testing these materials for the contaminant of interest can establish whether a contaminant was preferentially transported via this pathway.

Absent direct measurements, the presence of cracks, and/or the sampling and testing of expansion joint materials, a contaminant transport model can be used to estimate the time required for a liquid to permeate a paved surface. Ideally, information available to conceptually formulate a model includes:

- The temporal nature of the release (steady state or transient)
- The saturated and/or unsaturated hydraulic conductivity values of the pavement

- The physical properties of the contaminant (density, viscosity, vapor pressure)
- The chemical properties of the liquid (pure phase, mixed or dissolved in another fluid)
- Air temperature at the time of the release
- Volume released
- Evaporative flux

An understanding of the circumstances of a contaminant release and paved surface composition is essential so that reasonable assumptions regarding the release are incorporated into the model. For example, if the model does not account for evaporation and/or assumes that the liquid thickness on the paved surface is constant, the transport rate can be overestimated. If cleanup activities are performed coincident with the release (e.g., sawdust, green sand, absorbent socks, Sorball, crushed clay, pumping, etc.) or if the spill occurred in a building with forced air, these competing activities will result in less liquid available for transport through the paved surface and hence a slower transport rate. Additional information useful to develop the conceptual model includes whether the surface is treated with an epoxy coating or specialty coating to prevent corrosion from acid releases (common in plating shops, electronic and aerospace manufacturing facilities, and dry cleaners), or for a concrete pavement, whether the concrete was mixed with an additive to reduce its permeability to chemicals (e.g., Dow Latex No. 560), paved surface thickness, paved surface porosity, composition and slope, and the nature of the surface prior to the release (i.e., was it impregnated with oils and dirt, smooth vs. pitted, sloped toward a drain, etc.).

The liquid thickness residing on the paved surface and the duration of time that the liquid is in contact with the paved surface are key variables. If TCE (trichloroethylene), for example, is released onto a paved surface on a warm day or in a building with forced air, evaporation of the TCE is rapid. As a consequence, little liquid is available to migrate into the pavement. A description for the evaporative loss associated with a release of a compound for a semi-infinite region with a uniform initial concentration and mass transfer at the surface is given by (Choy and Reible, 1999):

$$\partial c_A/\partial t = (D_{A(\text{liquid})}/R_f)\partial^2 c_A/\partial z^2 \quad z \in [0, \infty] \tag{11.15}$$

and

$$-D_{A(\text{liquid})}\partial c_A/\partial z\big|_{z=0} + (k_a)\, c_A(z,t)\big|_{z=0} = 0, \quad t > 0 \tag{11.16}$$

and

$$c_A(z,t)\big|_{t=\infty} = c_{A0} \quad t > 0 \tag{11.17}$$

$$c_A(z,t)\big|_{t=0} = c_{A0} \quad z0[0, \infty] \tag{11.18}$$

The concentration profile is defined as:

$$c_A(z,t) = c_{A0}\left\{\mathrm{erf}\left[R_f z/(4D_{A(liquid)}R_f t)^{1/2}\right] + \exp(k_a z/D_{A(liquid)}R_f)\right.$$
$$\left.\mathrm{erfc}[R_f z/4D_{A(liquid)}R_f t)^{1/2} + k_a(t/D_{A(liquid)}R_f)^{1/2}]\right\} \tag{11.19}$$

where $z \in [0, \infty]$, $t > 0$.

The surface flux out of the system is described by:

$$j_A(t)\big|_{z=0} = (k_a)(c_{A0})\exp(k_a^2 t/D_{A(liquid)}R_f)\mathrm{erfc}[k_a(t/D_{A(liquid)}R_f)^{1/2}, \quad t > 0 \tag{11.20}$$

where k_a is the surface mass transfer coefficient; c_A is the concentration of species A in the air phase per unit volume; D_A is the effective diffusion coefficient of species A; R_f is the retardation factor; j_A is the mass flux of the component; z is depth of position; and t is time.

Conversely, if TCE accumulates in a blind concrete sump/neutralization pit or clarifier, the TCE may reside for a sufficient period to allow for the TCE to migrate through the concrete. Numerous equations and models can be modified to calculate the rate of transport of a liquid through paved surface (Ghadiri and Rose, 1992). For saturated flow, a simplified one-dimensional expression for the vertical transport of a liquid is Darcy's law where outflow (q) in time (t) from a paved column with an area (α) and thickness (l) subjected to a head difference (Δh) and the saturated hydraulic conductivity ($K_{saturated}$) is equal to:

$$q/t = [K_{saturated}\alpha/l](\Delta h) \tag{11.21}$$

This expression defines the downward velocity of the liquid to be equal to the downward flux (q/t) divided by the area and by the porosity of the pavement. The downward flux is the saturated hydraulic conductivity multiplied by the vertical gradient. Dividing the velocity into the paved surface thickness gives the transport time. Saturated hydraulic conductivity and porosity values for paved materials are measured directly or estimated from published values.

Fluid transport models using Darcy's law assume that the paved surface is saturated with liquid prior to the release. If the paved surface is unsaturated,

liquid transport may be dominated by unsaturated flow resulting in contaminant velocities several times slower than for saturated flow. In such cases, the use of unsaturated zone equations such as the Richards equation may be an appropriate model to estimate fluid velocity (Richards, 1931; Hillel, 1980). In order to develop the relationship between moisture content, capillary potential, and corresponding unsaturated hydraulic conductivity values, a soil moisture characteristic curve is required. The wetting fluid used to create a soil moisture characteristic curve for concrete or other material is ideally identical to the released liquid so that the extent of repulsion and/or retardation due to the hydrophobic or hydrophilic properties of the liquid can be quantified.

11.4.4 FLUID TRANSPORT THROUGH SOIL

A frequent inquiry in environmental litigation is identifying when a contaminant entered the groundwater or migrated onto another property. Absent direct measurements or testimony regarding the timing of a release, contaminant transport equations are relied upon to develop reasonable estimates. Equations for contaminant migration through soil range in complexity from one-dimensional infiltration to three-dimensional multiphase flow equations (Hughes *et al.*, 1992; Sanders, 1997; Selim and Ma, 1998). A determination is required prior to model selection to identify whether the fluid was transported as a vapor or as a liquid and whether transport occurred under saturated or unsaturated conditions through the soil column.

11.4.4.1 Vapor Transport

Vapor is transported in soil via advective or gaseous diffusion. Advective transport is important near the ground surface due to atmospheric pressure variations or near buildings that create pressure gradients due to differential heating. Gaseous diffusion describes transport via molecular processes and is the primary transport mechanism for contaminant migration through the soil. In 1855, Adolf Fick described gaseous diffusion as $F = -D(\mathrm{d}C/\mathrm{d}x)$, where F is the mass flux, D is the diffusion coefficient, and $\mathrm{d}C/\mathrm{d}x$ is the concentration gradient. This expression is known as Fick's first law (Fick, 1855).

For contaminant vapor transport through soil, the effective diffusion coefficient is the vapor diffusion coefficient corrected for soil porosity. For many vapors, the diffusion coefficient is approximately 0.1 cm^2 per second. A conservative estimate is that porosity reduces vapor diffusivity by a factor of 10. Thus for vapors, the effective diffusion coefficient D_e can be approximated as equal to about 0.01 cm^2 per second. A first order approximation illustrating

the distances that a contaminant can travel via vapor diffusion through soil in one year (assuming no adsorption) is (Morrison *et al.*, 1998):

$$D = [(2)(0.01\text{cm}^2\text{s}^{-1})(31{,}536{,}000\text{ sec})]^{1/2} \approx 800\text{ cm} \approx 25\text{ feet} \quad (11.22)$$

More complex expressions for vapor transport are available in the literature (Carslaw and Jaeger, 1959; Crank, 1985).

11.4.4.2 Liquid Transport

Models used to model the transport of a contaminant through the soil column rely upon similar mathematics. A common one-dimensional equation describing liquid transport via advection and diffusion in soil is (Jury *et al.*, 1986; Jury and Roth, 1990; Jury and Fuller, 1992):

$$R_l \partial C_l/\partial t = D_e \partial^2 C_l/\partial z^2 - V\partial C_l//\partial z - \lambda_\mu R_l C_l \quad (11.23)$$

where C_l is the pore water concentration in the vadose zone; λ_μ is the decay constant; R_l is the liquid retardation coefficient; D_e is the effective diffusion coefficient; and V is the infiltration rate.

The retardation coefficient (R_l) is estimated by:

$$R_l = 1 + (\rho_b K/\theta) + (\Phi - \theta/\theta) K_H \quad (11.24)$$

where ρ_b is the soil bulk density (for example determined by American Society of Testing Materials Method D3550), K is the distribution coefficient for the contaminant, Φ is soil porosity, θ is the soil moisture content (American Society of Testing Materials Method D2216), and K_H is the Henry's constant for the contaminant.

For most soils, adsorption occurs on the organic carbon particles or clay materials; the organic carbon partition coefficient (K_{ow}) for organic chemicals is either measured or estimated. The partition coefficient is estimated by Lyman *et al.* (1982) as:

$$K_d = 0.6 f_{oc} K_{ow} \quad (11.25)$$

where f_{oc} is the organic carbon fraction in the soil and K_{ow} is the octanol-partition coefficient of the contaminant of concern. Additional adsorption rates can also be coupled into this expression to incorporate other materials, such as soot.

The degradation rate constant is estimated as:

$$\lambda_\mu = \ln(2)/t_{1/2} \quad (11.26)$$

where $t_{1/2}$ is the degradation half-life of the contaminant of concern. The effective diffusion coefficient is described as:

$$D_e = \tau_L(D_{LM}) + K_H\tau_G D_{GM} \tag{11.27}$$

where D_{LM} is the molecular diffusion coefficient in water; τ_L is soil tortuosity, relative to the diffusion of water in soil; D_{GM} is the molecular diffusion coefficient in air; and τ_G is soil tortuosity to air diffusion.

The tortuosity of a compound in water and air is described as (Millington and Quirk, 1959):

$$\tau_L = \nu^{10/3}/\Phi^2 \tag{11.28}$$

and

$$\tau_G = (\Phi - \nu)^{10/3}/\Phi^2 \tag{11.29}$$

where ν is equal to the air-filled soil porosity.

In the United States, SESOIL (Seasonal Soil Compartment Model) (Odencrantz *et al.*, 1992; Bonazountas *et al.*, 1997) and VLEACH (One-dimensional Finite-difference Vadose Zone Leaching Model) (Rosenbloom *et al.*, 1993) are the most common models used to estimate the time required for a contaminant to migrate through the soil column. Given that these models are frequently encountered, each is briefly discussed.

11.4.4.3 SESOIL

SESOIL was developed as a risk screening-level model that uses soil, chemical, and meteorological input values. SESOIL is a one-dimensional vertical transport code for the unsaturated zone and simultaneously models water transport and contaminant sediment transport. SESOIL can simulate contaminant fate and migration on a monthly basis and can estimate the average concentration of the contaminant in the underlying groundwater. The soil column can be divided in up to four layers, with each layer possessing different soil properties. Each soil layer in turn can be subdivided into a maximum of 10 sublayers. Table 11.5 lists some of the climatic, soil, and chemical input data used in SESOIL (General Sciences Corporation, 1998).

SESOIL describes the depth reached by a chemical with a linear equilibrium partitioning between its vapor, liquid, and adsorbed phases (Jury *et al.*, 1984):

$$D = (V_w t_c)/[(\theta + \rho_b K_d) + (f_a H)/R(T + 273)] \tag{11.30}$$

Table 11.5
Key input parameters for SESOIL.

Data Category	Parameter Information
Climatic data	Air temperature; cloud cover fraction; relative humidity; short wave albedo; rainfall depth; mean storm duration; number of storms per month; length of rainy season within a month.
Soil	Bulk density; intrinsic permeability; soil disconnectedness index; effective porosity; organic carbon content; cation exchange coefficient; Freundlich equation exponent.
Chemical	Solubility in water, diffusion coefficient in air, Henry's law constant, molecular weight, valence, hydrolysis rate (if appropriate), biodegradation rate in liquid and soil; adsorption coefficient on organic carbon; adsorption coefficient to soil.

where D is depth; V_w is water velocity; t_c is convective time; θ is soil water content; ρ_b is soil bulk density; K_d is contaminant partitioning coefficient; f_a is air filled porosity; H is Henry's law constant; R is universal gas constant; and T is soil temperature (˚C).

SESOIL is especially sensitive to the saturated hydraulic conductivity value(s) used and the organic matter content. The selection of the organic matter content and a representative hydraulic conductivity value can result in considerable differences (years) in the estimated travel time of a contaminant to groundwater.

11.4.4.4 VLEACH

VLEACH is a one-dimensional finite difference model that describes the movement of an organic contaminant as a dissolved solute, as a gas, and as an adsorbed compound (Rong, 1999). The mathematical expression for transport in each phase and model assumptions are described. Liquid phase advection is expressed as:

$$\partial C_l/\partial t = -q/\theta(\partial C_l/\partial z) \tag{11.31}$$

where C_l is the liquid phase concentration, θ is the water-filled porosity of soil by volume, $-q$ is the infiltration rate, z is the vertical dimension, and t is time. Gaseous diffusion is defined as:

$$\partial C_g/\partial t = -D_e(\partial^2 C_g/\partial z^2) \tag{11.32}$$

where C_g is the gaseous phase concentration and D_e is the effective diffusion coefficient. The governing equation used to describe the adsorbed phase concentration (C_s) is:

$$C_s = (K_{oc} f_{oc})\, C_l \tag{11.33}$$

where C_s is the adsorbed phase concentration, K_{oc} is the organic carbon partition coefficient and f_{oc} is the soil organic carbon content. The governing equilibrium equation used for each vertical simulation cell is:

$$C_T = (\theta C_l) + (\Phi - \theta) C_g + (\rho_b C_s) \tag{11.34}$$

where C_T is the total concentration for all three phases at equilibrium, Φ is the soil porosity, and ρ_b is soil bulk density. Recent additions to the model provide the ability to incorporate Monte Carlo simulations so that a reasonable range of input values can be used (Rong and Wang, 2000). VLEACH assumptions include

- Local or instantaneous equilibrium between each phase
- Linear isotherms describe the partitioning of the contaminant between each phase
- The moisture content in the homogenous soil is constant
- Liquid phase dispersion is neglected causing higher dissolved concentrations and lower travel time estimates than what would occur in reality
- No *in situ* degradation or production occurs
- Volatilization from the soil boundaries is unimpeded or completely restricted
- Nonaqueous phase liquids and/or flow conditions derived from fluid variable density are not simulated

11.4.5 CONTAMINANT TRANSPORT MODELS IN SOIL

Challenges to contaminant transport models in soil include an examination of the presence and significance of preferential pathways and the validity of the model assumptions. Categories of preferential pathways impacting model results include natural and artificial features, colloidal transport, and cosolvent transport.

Natural and artificial (dry wells, cisterns, utility line backfill, etc.) preferential pathways provide a means for contaminants to be rapidly introduced at depth (Barcelona and Morrison, 1988; Morrison *et al.*, 1998). Natural preferential pathways include worm channels, decayed root channels, soil fractures, slickenslides, swelling and shrinking clays, highly permeable soil layers, and insect burrows. The term *preferential flow* encompasses a range of processes with similar consequences for contaminant transport. The term implies that infiltrating liquid does not have sufficient time to equilibrate with the slowly moving water residing in the soil (Jarvis, 1998). Preferential flow includes finger flow (also viscous flow) (Glass and Nicholl, 1996), funnel flow (Philip, 1975; Kung, 1990a,b), and macropore flow (White, 1985; Morrison and Lowry, 1990).

Finger flow (also dissolution fingering) is initiated by small- and large-scale heterogeneities in soil such as a textural interface between a coarse-textured sand and an underlying silt (Fishman, 1998; Miller *et al.*, 1998). The term *finger flow* refers to the splitting of an otherwise uniform flow pattern into fingers. These fingers are associated with soil air compression encountered where a finer soil overlies a coarse and dry sand layer. The contact interface between the contaminant and the water at the capillary fringe can result in an instability. The spacing and frequency of these fingers are difficult to predict although they are at the centimeter scale and are sensitive to the initial water content (Ritsema and Dekker, 1995; Imhoff *et al.*, 1996).

Numerical simulations of fingering suggest that transverse dispersion has a significant impact on the formation and anatomy of fingers. Aspects of fingering phenomena that introduce uncertainty when modeling contaminants such as nonaqueous phase liquids (NAPLs) include the effect of dispersion, the impact of heterogeneity on porous media properties and residual NAPL saturation, the validity of fingering when a NAPL solution is flushed with chemical agents such as surfactants and alcohols, and incorporation of the impacts of fingering on NAPL phase mass transfer models when the model is discretized at scales larger than the centimeter scale.

Funnel flow occurs in soils with lenses and admixtures of particle sizes. For a saturated soil, the coarsest sand fraction is the preferred flow region; for unsaturated flow, finer textured materials are more conductive. Examination of textural descriptions on boring logs and contaminant concentration depth profiles can provide insight to determine if contaminant transport via funnel flow is a viable transport mechanism.

A macropore is a continuous soil pore that is significantly larger than the intergranular or interaggregate soil pores (micropores). A macropore may constitute only 0.001 to 0.05% of the total soil volume, but it may conduct a majority of an infiltrating liquid.

Colloidal transport describes the sorption of a hydrophobic compound to a colloid particle in water. A colloid is defined as a particle ranging from 0.003 to 10 mm in diameter. Colloids exist as suspended organic and inorganic matter in soil or aquifers. In sandy aquifers, the predominant colloids that are mobile range in size from about 0.1 to 10 mm.

The contaminant mass associated with colloidal transport may be significant. Numerous studies have demonstrated that contaminants are transported at depth via colloidal transport (Sheppard *et al.*, 1979, 1980; Robertson *et al.*, 1984; Puls *et al.*, 1990; Pohlmann *et al.*, 1994). Buddemeier and Rego (1986) found that virtually all the activity of manganese, cobalt, antimony, caesium, cerium, and europium was associated with colloidal particles in groundwater samples collected from underground nuclear-test cavities at the Nevada Test

Site. An example of the transport of polycyclic aromatic hydrocarbons via colloidal transport was examined in two creosote-impacted aquifers on Zealand Island in Denmark (Villholth, 1999). The mobile colloids were composed predominately of clay, iron oxides, iron sulfides, and quartz particles. The researchers concluded that the sorption is hydrophobic and associated with the organic content of the colloids. Creosote associated contaminants (i.e., polycyclic aromatic hydrocarbons, phenols and nitrogen, sulfur or oxygen containing heterocycles (Mueller *et al.*, 1989)) were found to be associated primarily with colloids larger than 100 nm. These findings indicate that colloid-facilitated transport of polycyclic aromatic hydrocarbons may be significant. Although Sheppard *et al.* (1980) concluded over 25 years ago that the transport of radionuclides by colloidal clay particles should be considered in contaminant transport models, this transport mechanism is rarely considered and/or simulated in soil or groundwater models.

Hydrophobic compounds generally are considered to be immobile in the soil profile due primarily to their low water solubility and their tendency to be adsorbed by clay, organic matter, and mineral surfaces (Odermatt *et al.*, 1993). Soil contaminant transport models tend to predict low velocities for these compounds. Cosolvation of these compounds in a fluid that can introduce these contaminants at depth is rarely included in contaminant transport models. Contaminants such as polychlorinated biphenyls (PCBs) or DDT (dichlorodiphenyltrichloroethane) and its derivatives can be remobilized by the preferential dissolution of the PCBs into a solvent released into the same soil column (Morrison and Newell, 1999). A variation to this scenario is the preferential dissolution of an immobile chemical into a solvent prior to its release (e.g., PCBs dissolved in a dielectric fluid).

A similar transport mechanism occurs when a contaminant sorbed by a soil is washed with a liquid that remobilizes the compound. An example is the presence of copper bound in soil under a leaking neutralization pit. Low pH wastewaters leaking through an expansion joint and contacting the precipitated copper will remobilize and transport the copper with the low pH wastewaters to depth until the acidic wastewater and copper solution is buffered and the copper reprecipitates at a lower depth.

Potential challenges to contaminant transport models in soil include examination of the model assumptions, including the circumstances of the release and whether these are realistically simulated in the model. Most soil transport models, for example, do not use statistical methods to determine the precision of the release parameters (Opdyke and Loehr, 1999). Release parameters are commonly reported deterministically (i.e., as a point estimate) rather than probabilistically (i.e., a distribution of values and their associated probabilities). The reasonableness of the release parameter assumptions or

measurements therefore should be considered to determine whether the release assumptions are reasonable.

A simplifying assumption often encountered in environmental forensics and which is often valid at the laboratory scale is that the soil is homogeneous (Miller *et al.*, 1965). Large-scale field studies of solute transport through the unsaturated zone, for example, have revealed enormous variability in pore-water velocities and hydrodynamic dispersion coefficients (Quisenberry and Phillips, 1976; Schulin *et al.*, 1987; Andreini and Steenhuis, 1990). This field variability coupled with the uncertainties associated with estimating other processes relevant to contaminant transport can introduce serious difficulties in the simulation of heterogeneous field-scale transport. A number of approaches, including scaling theories, Monte Carlo simulations, stochastic-continuum models (Russo and Dagan, 1991), and stochastic convective models have been proposed to deal with these field-scale issues.

An approach used to establish a confidence level for the various uncertainties in the input parameters is to perform a probabilistic analysis (Rong, 1995; Duke *et al.*, 1998). Two probabilistic methods are the likelihood method that is effective in estimating the confidence of parameters in nonlinear models (Donaldson and Schnabel, 1987; Bates and Watts, 1988) and Monte Carlo simulations (Ang and Tang, 1990; Berthouex and Brown, 1994). Monte Carlo simulations provide a numerical method for generating simulated data using assumed probability distribution. This analysis identifies those parameters that influence the calculated groundwater contaminant concentration and provides an uncertainty analysis, including the most sensitive model variables (Hetrick and Pandey, 1999).

11.5 CONTAMINANT TRANSPORT MODELS FOR GROUNDWATER

There are more than 400 groundwater flow (advective) and contaminant transport models (Zeng and Bennett, 1995). Groundwater models based on a multitude of mathematical approaches are available, including probabilistic and geostatistical simulations (Ewing, 1975; Bagtzoglou, 1990; Bagtzoglou *et al.*, 1991, 1992; Sidauruk *et al.*, 1998), analytical solutions (MacDonald, 1995; Alapati and Kabala, 2000), heat transport inversion (Clark and Oppenheimer, 1994; Ames and Epperson, 1997; Neupauer *et al.*, 2000), and direct methods (Elden, 1982; Atmadja and Bagtzoglou, 2000, 2001a,b,c). In environmental forensic investigations, the most frequently used contaminant transport models are classified generically as:

- Instantaneous point source solutions for a contaminant injected into the groundwater

- Continuous injection of a contaminant into the groundwater
- The transient injection of a contaminant into groundwater
- The transient injection of multiple sources of contaminants into the groundwater
- Multiphase models
- Multiphase inverse models

Regardless of the selection of the model, the mathematical expression most appropriate for most conditions and from which special model applications are based is the standard two-dimensional convection-dispersion-adsorption model for solute transport. The foundational advection/dispersion equation used in traditional groundwater models is also used for backward extrapolation modeling (Wilson *et al.*, 1981; Wilson and Linderfelt, 1991; Liu, 1995; Wilson and Liu, 1995). Governing assumptions are that the porous medium is homogeneous and isotropic, the pore space is saturated with a fluid, and that Darcy's law is valid (Bear *et al.*, 1992). If these assumptions are violated, the applicability or precision of the model is compromised to some degree.

For the flow of a conservative solute in a homogeneous, isotropic medium, this model is described as:

$$\partial/\partial t(\mathrm{R(x,y,t)c}) + \mathrm{V(x,y,t)}\partial \mathrm{c}/\partial \mathrm{x} - \partial/\partial \mathrm{x}(\mathrm{D_L(x,y,t)}\partial \mathrm{c}/\partial \mathrm{x}) \\ \partial/\partial \mathrm{y}(\mathrm{D_T(x,\,y,\,t)}\partial \mathrm{c}/\partial \mathrm{y}) = \mathrm{q} \quad (11.35)$$

where

q = volumetric solute source or sink term
c = vertically averaged solute concentration
V = $\mathrm{U_x(x,y,t)}/\varphi$ (x,y) and is the linear solute velocity
U_x = Darcy velocity for one-dimensional flow along the x axis
D_L = longitudinal dispersion coefficient
D_T = transverse dispersion coefficient
t = time
x,y = two dimensional coordinates

and

$$\mathrm{R} = 1 + \rho_\mathrm{b}\mathrm{K_d(x,y,t)}/\varphi\mathrm{(x,y)} \quad (11.36)$$

where

R = retardation factor
ρ_b = $(1 - \varphi)\,\rho_{\mathrm{grain}}$
$\mathrm{K_d}$ = distribution coefficient

and

$$K_d = K_{oc} f_{oc} \tag{11.37}$$

where

K_{oc} = organic carbon partition coefficient
f_{oc} = weight fraction of organic carbon in the soil

and

$$D_L = D^* + \alpha_L(x,y)V(x,y,t); \tag{11.38}$$

$$D_T = D^* + \alpha_T(x,y)V(x,y,t); \tag{11.39}$$

where

D^* = coefficient of molecular diffusion of the contaminant in water
α_L = longitudinal dispersivity
α_T = transverse dispersivity

11.5.1 INVERSE MODELS

The term *backward extrapolation model* was reported in 1991; other authors have used the terms *reverse or inverse modeling*, *hindcasting*, and *backward random walk* to describe the same approach (Kezsbom and Goldman, 1991; Kornfeld, 1992; Erickson and Morrison, 1994; Woodbury and Ulrych, 1996; Morrison, 1998, 1999a,b; Parker, 2000; Aral *et al.*, 2001; Mahinthakumar and Sayeed, 2006). In its simplest application, inverse modeling relies upon measured properties or contaminant concentrations to extrapolate the age and/or location of a contaminant release to some point in the past. Most techniques rely upon a geostatistical or optimization approach (Gorelick and Remson, 1982; Gorelick *et al.*, 1983; Khachikian and Harmon, 2000). A detailed mathematical description of the use of inverse models are described by Sun (1994). Inverse models are used in environmental litigation to reconstruct complex, multiparameter release histories with known source locations from spatially distributed plume concentration data (Skaggs and Kabala, 1994, 1995, 1998; Parker, 2000).

An early application of inverse modeling for contaminant transport in groundwater used a least squares regression technique to identify the locations and discharge rates of contaminants that were simultaneously migrating from five disposal sites over a five-year period (Gorelick *et al.*, 1983). Other

applications used backward probability models to establish the most likely location of a contaminant release. Bagtzoglou *et al.* (1992) described a reverse time random particle method to estimate the probability that a particular site was the source of a contaminant plume. These probability approaches usually assume a single, point source release (Birchwood, 1999). In another application, the vertical concentrations of PCE and TCE in an aquitard were evaluated to reconstruct a release history; only diffusion and sorption were assumed to affect mass transport in the aquitard (Liu, 1995; Ball *et al.*, 1997; Liu and Ball, 1999). Dimov *et al.* (1996) used the adjoint formulation of the one-dimensional advection–dispersion equation to identify a contaminant source and estimate the contaminant loading rates at different nodes. Another approach modeled contaminants from a constant source concentration using two-dimensional diffusion to estimate source characteristics and transport parameters from head and concentration data (Sonnenborg *et al.*, 1996). Sidauruk *et al.* (1998) developed a set of analytical procedures for contaminant source identification that relied upon contaminant concentrations measured within a two-dimensional contaminant plume. Contaminant flow was assumed to be homogeneous and groundwater velocity was uniform for instantaneous and continuous releases. The authors noted the limitations of certain parameters that can only be uniquely determined by sampling at a minimum of two different times.

Skaggs and Kabala (1994) and others have used Tikhonov regularization to transform an algebraically underdetermined inverse problem into a minimization problem with a unique solution as well as using the method of quasi-reversibility to solve the problem (Alifanov and Artyukhin, 1976). The location of the source was assumed to be known. Wagner (1992) proposed a deterministic approach for the combined estimation of model parameters and source characteristics. Bagtzoglou *et al.* (1992) used backward location probabilities to identify contaminant sources. Probability maps were subsequently developed using a random walk method by reversing the flow field and assuming that dispersion was constant.

One expression used to simulate a one-dimensional advection–dispersion model backward in time is described by Wilson and Liu (1995):

$$\partial g/\partial t = D(\partial^2 g/\partial x^2) - V(\partial g/\partial x) \tag{11.40}$$

where g is a probability density or cumulative distribution function of interest, D is the dispersion coefficient, t is time, and x is the upgradient distance to the source. The authors hypothesized that by selecting appropriate initial and boundary conditions, Equation 11.40 can be solved for the desired probability function to create travel time and location probability

maps when the location of the release is unknown. Neupauer and Wilson (1999, 2004) describe the solutions for source and time probabilities for a pumping well as:

$$\partial f_x/\partial \tau = D(\partial^2 f_x/\partial x^2) + v(\partial f_x/\partial x) \tag{11.41}$$

where

$$f_x \to 0 \quad \text{as} \quad x \to -\infty \tag{11.42}$$

and

$$v f_x + D(\partial f_x/\partial x) = 0 \text{ at } x = 0 \quad \text{and} \quad f_x(x, 0) = \delta(x - x_d) \tag{11.43}$$

where τ is backward time ($\tau = t_d - t$, where t_d is the detection time), x_d is the detection location at a pumping well ($x_d = 0$), D is the dispersion coefficient, v is the groundwater velocity, f_x is the location probability, and δ is the Dirac delta function. The solution for Equation 11.43 provides the backward probability for a detection at $x_d = 0$, as:

$$f_x(x, \tau) = 1/(\pi D_\tau)^{1/2} \exp\{-x + v\tau)^2/4D\tau\} - (v/2D)\exp\{-(vx)/D\} \\ \text{erfc}[(-x + v\tau)/(4D\tau)^{1/2}] \tag{11.44}$$

For a contaminant particle removed at the pumping well ($x_d = 0$), Equation 11.44 describes the probability density function of the location of the contaminant parcel at time t before it was observed at the pumping well. The backward model for the travel time probability map is then:

$$\partial f_\tau/\partial \tau = D(\partial^2 f_\tau/\partial x^2) + v(\partial f_\tau/\partial x) \tag{11.45}$$

where

$$f_T \to 0 \quad \text{as} \quad x \to -\infty \tag{11.46}$$

and

$$v f_\tau + D(\partial f_\tau/\partial x) = v\delta(\tau) \text{ at } x = 0 \quad \text{and} \quad f_\tau(x, 0) = 0 \tag{11.47}$$

where D and n are constant and the pumping well is assumed to be at $x_d = 0$. The solution to Equation 11.47 is the backward-in-time time probability from location x to a detection at $x_d = 0$ (Wilson and Liu, 1995):

$$f_\tau(x,\tau) = 1/(\pi D\tau)^{1/2}\exp\{-(x+v\tau)^2/4D\tau\} - (v^2/2D)\exp\{-(vx)/D\}$$
$$\text{erfc}[(-x+v\tau)/(4D\tau)^{1/2}] \qquad (11.48)$$

For a contaminant detected at the pumping well, Equation 11.48 describes the probability density function of the travel time from an upgradient source location (x) to the pumping well at $x_d = 0$.

A difficulty inherent in using inverse modeling for reconstructing a groundwater release scenario is the computational difficulty and ill-posedness in simulating different release histories that provide a close match with observed concentrations. To address and/or minimize these difficulties, various methods have been employed, including the following (Mahinthakumar and Sayeed, 2006):

- A minimium relative entropy approach (Skaggs and Kabala, 1994; Woodbury and Ulrych, 1996; Woodbury *et al.*, 1998)
- Tikhonov methods (Liu and Ball, 1999; Neupauer *et al.*, 2000)
- Constrained nonliear optimization (Mahar and Datta, 2000, 2001; Singh *et al.*, 2004)
- Progressive genetic algorithmns (Aral *et al.*, 2001; Mahinthakumar and Sayeed, 2005)
- Nonlinear least-squares with Levenbberg-Marquadt minimization (Sciortino *et al.*, 2000)
- Nonlinear least-squares with Gauss-Newton (Alapati and Kabala, 2000)

Depending on the specifics of a particular problem, the proper selection of one of these methods may minimize the computational difficulties and model errors.

The successful evaluation of an inverse model includes an analysis of model parameters and conceptual framework. Examples of model parameters incorporated into many inverse models for a single contaminant that represent potential sources of uncertainty include (Erickson and Morrison, 1995, 1998):

- The reasonableness of the selected porosity and hydraulic conductivity value(s). Soil porosity and hydraulic conductivity within an aquifer and with distance can change dramatically. A representative porosity value, on a field scale, is usually a fitted parameter within a published range of values for the predominant soil type encountered. As a result, differences in the modeling results can be adjusted by slight variations in the selected porosity value. Typical porosity values range from 0.25 to 0.50 for unconsolidated soils.

- The consistency of the groundwater flow direction and velocity over time. The hydraulic gradient value selected is an important variable to scrutinize as its value can vary in time and distance. If regional- or vicinity-wide data are relied upon to define the hydraulic gradient versus site measurements, considerable differences can occur. Sources of localized variations in the gradient include pumping wells, rivers, streams and groundwater recharge or spreading basins. If pumping wells in the immediate vicinity of the site are present, collecting the extraction rates, representative transmissivity values for the aquifer, and calculating the radius of pumping influence is useful to demonstrate the potential impact of pumping wells on the local gradient. Historical variations in the hydraulic gradient can introduce significant uncertainty regarding the historical direction of groundwater flow and hence the time required for the contaminant plume to attain its measured leading edge geometry. Typical hydraulic gradient values range from 0.1 to 0.001.
- The validity of the selected hydraulic gradient(s) values over time and distance from the release.
- The number of data points and time interval during which the data were collected (Mahar and Datta, 1997).
- The nature of the release (steady versus nonsteady).
- The loading or mass flux rate.
- The value(s) selected for aquifer(s) thickness (model specific).
- The horizontal and transverse dispersivity values (Mercer, 1991).
- Contaminant retardation and/or degradation rates (Davidson, 2000).
- Identification of the leading edge of the contaminant plume (model specific).
- The effect of recharge/discharge rates (if applicable) of water into the system and its impact on plume geometry and contaminant velocity. Some inverse models require an accurate value for the contaminant plume length, the location of the release, or when a contaminant first entered a monitoring well downgradient from the source. If the source location and/or the plume length are assumed, significant variations in the estimated age of the release can occur.

Another issue is whether the compound or class of compounds used to determine the plume length biases the measurement. The selection of benzene, for example, may reflect a stabilized or decreasing length, whereas the use of MTBE for the same plume may indicate an expanding plume, depending on the source strength and effects of biodegradation. A similar dilemma exists for transformation products. Among likely parent–daughter compound combinations for chlorinated solvents, distances between the respective maximum concentration locations were found to be on the order of 10% to 25% of the maximum plume length at most sites (McNab *et al.*, 1999). For multiple releases, compound selection, release location, and potential cosolvation impacts require scrutiny (Morrison and Newell, 1999). Examples of possible

challenges or areas of inquiry when evaluating inverse models results include (Morrison, 1999a,b, 2000b):

- Adjust the model-input parameter, within a reasonable range, so that results are produced consistent with measured values but inconsistent with the alleged released date. For example, the distribution of simulated head measurements depends to some degree on the recharge rate to saturated hydraulic conductivity value ratio.
- Determine the confidence level associated with simulated or assumed model parameters (e.g., groundwater velocity, analytical bias, hydraulic conductivity, temporal and spatial variations of the data, time interval between sampling dates, etc.).
- Collect water level measurements coincident with an activity, such as a pumping well, that stresses the aquifer. Use the model to simulate this activity and compare the actual measurements with those predicted by the model.
- Identify whether the water mass balance is conserved. A well-calibrated model exhibits a small difference in the water balance, which is often included as output to the model.
- Evaluate model calibration via the preparation of scattergrams comparing values for measured heads with the simulated heads for each well. If the calibration is perfect, the data sets should fall along a 45-degree line. Data points farther from this line reflect a greater deviation from actual flow conditions. Plotting the residuals (the difference between the simulated and observed heads) can also provide insight regarding model calibration.
- Compare the input parameters and compare them with those used for other site models (e.g., models used for designing a remediation system, health risk models used for site closure, etc.). It is not uncommon that physical and chemical parameters used in a model in support of a risk-based corrective action (RBCA) study are different than those used for inverse models for the same site.

Inverse models rarely incorporate the time required for the contaminant to flow through a paved surface. If the depth to groundwater is shallow or the contaminant is introduced quickly into the groundwater via a preferential pathway (i.e., a dry well, cistern, well casing, etc.), age dating the release with an inverse model may be a reasonable approach for estimating the age and/or location of a well-defined, discrete release. If the release occurred at the ground surface on impermeable soils, and/or the depth to groundwater is significant (>50 ft), the cumulative time required for contaminant migration through the soil column can be significant. Ignoring this portion of the conceptual model with an inverse model can significantly underestimate age of the release.

Boundary conditions are required for inverse models. For programs such as MODFLOW, general head boundaries are used to define the lateral boundary conditions that determine the flux of water recharge or discharge along these boundaries. The boundary conditions are a function of the hydraulic conductivity, groundwater flow gradient, and the absolute difference in water level elevations between the block elements located on the lateral boundaries and those outside of the model grid. Examples of boundary conditions include no-flow, specified flux, and fixed-head boundaries. Model boundary conditions are fixed and cannot be changed during a single simulation, however, they can be adjusted between simulations. It is conceptually undesirable to alter the boundary flux conditions to assist in calibration of each stress period, versus accounting for these differences by adjustments in dynamic features such as pumping wells or recharge of surface water bodies located within the grid. The impact of a model boundary can be examined if all model input files and software are available to reproduce the modeling result using different boundary conditions.

Grid selection is an important model variable. For numerical models, finite difference and finite element grids are used. Block-centered and mesh-centered grids are used for finite element grids. Finite element grids are generally more versatile than finite difference grids. For a finite difference model, the grid density should be examined to ascertain whether the data support finer mesh nodes or whether higher grid densities are selected in areas of interest but which contain insufficient data to warrant a higher grid density.

In finite difference modeling, numerical dispersion is inherent due to errors associated with the computational algorithms, especially in areas of varying grid size. The three-dimensional block size selected must be examined to determine the relative horizontal and vertical element aspect ratio. Aspect ratios less than 4 are generally acceptable; the horizontal and vertical aspect ratios that are greater than 4 become more susceptible to numerical dispersion.

For multiple layered models, it is important to determine whether the vertical gradients between the layers are measured or estimated. To confirm measured vertical gradients, the head difference between a shallow and deep well should be divided by the vertical distance between the bottom of both well screens. A negative value indicates a downward flow component. If the vertical gradients are estimated, an attempt should be made to determine the level of uncertainty associated with these values and their overall impact on solute transport between layers.

The definition of the contaminant source loading rate and when the loading commences (e.g., at the commencement of facility operations, 10 years after commencement, etc.) in a groundwater modeling grid is often arbitrary. The analyst should remove as much arbitrariness as possible by examining the site

operating history and the characteristics of transport through the vadose zone. Issues regarding the validity of a particular loading rate and its location include:

- Whether the soil and groundwater chemistry justifies the selected location and input rate
- Whether the mass loading rate is continuous or is transient in response to groundwater fluctuations or remediation activities
- Whether the start date for the mass loading is consistent with the operational history of the contributing surface sources

Most contaminant transport models perform a mass balance; this output should be obtained and reviewed. A significant error in the mass balance calculation indicates that the solution is imprecise.

The selection of an appropriate contaminant transport model is critical to the development of a realistic fate and transport model. If the conceptual model does not represent the relevant flow and contaminant transport phenomena, the subsequent modeling effort is wasted. Model complexity is determined by the quality of the data available for its design and verification. An example of an inappropriate model selection is a simple two-dimensional model used for a complex three-dimensional system. Conversely, a complex three-dimensional model may be selected that is overpowered relative to the available data. Model selection considerations include the following issues (Cleary, 1995):

- It is advisable to select a model with the prestige of a state or federal governmental agency development or official use, a substantial published history in peer-reviewed journals, and/or one that has already been tested in court. If the model code is obscure, it should be thoroughly scrutinized.
- The model should include a user's manual listing its governing assumptions, advantages, and capabilities.
- The model should be validated against analytical solutions for comparison.
- The analytical solution should have the same number of space dimensions as the numerical model.
- The model should be benchmarked against a numerical code.
- A model that has an available source code should be selected.
- A three-dimensional model may better represent reality than a two-dimensional approximation.

The model simulation ultimately selected to create a trial still or animation should be evaluated in the context of all the simulations. It is not unusual that

hundreds of simulations are performed until a simulation is obtained for use as evidence for a particular allegation. Discarded simulations should be compared with the selected simulation so that the legitimacy of the selected simulation, as the best representation, can be evaluated.

Of the hydrogeologic parameters, the saturated hydraulic conductivity value generally introduces the most significant variability in the computer simulations. It is generally recognized that the most representative measurements for determining the saturated hydraulic conductivity of a high permeability formation is via a pump test. Hydraulic conductivity values that rely upon slug tests, sieve analysis, and laboratory measurements of soil cores are considered less reliable. Measurements from a slug test are generally reliable to about one or more orders of magnitude with its accuracy increasing with lower permeability materials. Values of hydraulic conductivity can vary by a factor of 10 to 1000 within an aquifer and are often scale-dependent and strongly affected by heterogeneity. Vertical saturated hydraulic conductivity values are frequently a factor of 10–100 times less than the horizontal saturated hydraulic conductivity values within the same formation.

Dispersivity describes the three-dimensional spreading of a contaminant plume in groundwater with distance (Domenico, 1987; Domenico and Schwartz, 1990). Most contaminant transport models require a horizontal (longitudinal), vertical, and transverse dispersivity value. Longitudinal dispersivity is caused primarily by differences in groundwater flow through aquifer pores at a scale less than that used to characterize values of saturated hydraulic conductivity. Longitudinal dispersivity increases with distance from the source (also known as macrodispersivity). If a model assumes that the contaminant plume length is due only to advective flow, the estimated date of the release will be earlier than if dispersivity is also considered (Baca, 1999).

In reverse modeling, dispersivities are chosen to match the shape of the simulated contaminant plume to the observed plume geometry. Longitudinal dispersivity values used with solute transport models are commonly in the range of 90 to 300 feet (27–91 m). There is little physical evidence for using such large numbers except to simulate contaminant concentrations that compare favorably with observed values. In cases where no data exist to estimate horizontal dispersivities, the United States Environmental Protection Agency recommends multiplying the length of the plume by 0.1 (Wilson *et al.*, 1981). Other authors have used probabilistic theory to estimate transverse and vertical dispersivity as 0.33 and 0.056 times the plume length, respectively (Gelhar and Axness, 1981, 1983; Gelhar *et al.*, 1992; Salhotra *et al.*, 1993). Dispersivity is considered to be hysteretic (nonrepeating) with distance from

the contaminant source. As a result, a simple linear expression for dispersivity will introduce some degree of bias, especially at greater distances from the contaminant source.

The length of the contaminant plume is closely related to the source term, groundwater velocity, and contaminant degradation, primarily via reductive dehalogenation. In an evaluation of 247 chlorinated solvent plumes from 65 sites located in the western United States, a statistically significant correlation was observed between plume length and maximum concentration, with the correlation improving when higher concentration contours were used to define the plume length (i.e., 100 and 1000 ppb defined plumes) (McNab *et al.*, 1999). This correlation may reflect the increasing proximity to the source area, where variability in the maximum chlorinated solvent concentration is expected to have a more pronounced relationship to the variability in plume length.

When the length of a contaminant plume is measured, it is usually assumed *a priori* that the contaminant plume has been expanding at a relatively constant rate prior to the point in time that it is measured. This assumption may be invalid and should be examined when evaluating a model.

Significant advances have been made in multiphase flow and multicomponent transport in the last 15 years (Kaluarachchi *et al.*, 1990; Katyal *et al.*, 1991; Katyal and Parker, 1992; Huyakorn *et al.*, 1994; Pandey, 1995; Parker, 2000). Most of the comprehensive models based on finite difference and finite element solution algorithms solve continuity equations of water, oil, and gas for flow and convection dispersion transport equation for each component of concern.

The solution of inverse problems involving multiphase and variably saturated flow, and multicomponent transport are more difficult than a saturated groundwater flow problem. For example, under similar conditions, the groundwater model calibration involves finding sets of representative spatial distribution of hydraulic conductivity and porosity values. In comparison a multiphase calibration involves spatial distribution of conductivity, porosity, three/four retention parameters, fluid-specific gravity, interfacial surface tensions/scaling parameters, and interphase mass transfer and hysteresis parameters. The uncertainty in the initial and boundary conditions and source/sink for each fluid (water, oil, gas) adds to the complexities. The release history and past fluid levels' (air-water, air-oil) fluctuations are highly crucial as they determine the NAPL volume trapped in the unsaturated and saturated zones. The current fluid levels do not reflect the presence or absence of historic trapped NAPL. However, the trapped oil can become free when fluid fluctuations reverse and will be reflected by the current fluid levels. The trapped and free NAPL become the sources of vadose zone and groundwater contamination.

The mass conservation equations for water (w) and NAPL (o) for incompressible liquids and compressible gas can be written as:

$$\phi \frac{\partial S_w}{\partial t} = -\frac{\partial q_{w_i}}{\partial x_i} \tag{11.49}$$

$$\phi \frac{\partial S_o}{\partial t} = -\frac{\partial q_{o_i}}{\partial x_i} \tag{11.50}$$

$$\phi \frac{\partial \rho_a S_a}{\partial t} = -\frac{\partial \rho_a q_{a_i}}{\partial x_i} \tag{11.51}$$

Darcy velocities in p-phase ($p = w, o, a$ for water, oil, and gas phases) are defined by

$$q_{p_i} = -K_{p_{ij}} \left(\frac{\partial \psi_p}{\partial x_j} + \rho_{rp} u_j \right) \tag{11.52}$$

where $K_{p_{ij}}$ is the p-phase conductivity tensor, $\psi_p = \dfrac{P_p}{g\rho_w^*}$ is the water equivalent pressure head of phase p, P_p is the p-phase pressure, g is gravitational acceleration, and ρ_w^* is the density of pure water, $\rho_{rp} = \dfrac{\rho_p}{\rho_w^*}$ is the p-phase specific gravity, ρ_p is the p-phase density, and $u_j = \dfrac{\partial z}{\partial x_j}$ is a unit gravitational vector measured positive upwards, z is elevation, and S_p is the p-phase saturation.

The p-phase conductivity tensor depends on the saturated hydraulic conductivity tensor, and the relative permeability of the respective phase (function of the effective phase saturation). The effective phase saturations are functions of the interphase capillary pressures, and interphase surface tensions. Multiphase models account for hysteresis to accurately estimate the spatial and temporal distribution of trapped NAPL; this adds two or three additional unknowns (such as maximum residual saturations in the unsaturated and saturated zones) to the calibration.

Defining the source(s) of contamination (organic compounds) is a difficult task. The spatial distribution, release time and rate, and composition (species' mole fractions) are generally not known accurately. When release time, rate/amount, and composition are known, the spatial distribution is a challenge; most of the uncertainty is due to soil heterogeneity and anisotropy. During the subsurface migration of oil (NAPL), pockets of residual NAPL may be formed. When this occurs in low permeability strata, the dissolution and vola-

tilization processes are extremely slow. Conventional single porosity convection dispersion fate and transport models can't mimic the physics accurately. A more general dual-porosity formulation (which can easily be reduced to a single porosity formulation; see later) described next is more desirable.

A typical subsurface transport media has five distinct regions:

- Voids filled with air
- Mobile water located inside the larger interaggregate pores or fractures
- Immobile water located mainly in the intra-aggregate pores or in the porous media surrounding fractures
- A dynamic soil region, in equilibrium with the mobile phase
- A stagnant soil region where mass transfer is diffusion limited

A general transport equation can be described as (van Genuchten and Wierenga, 1976):

$$\frac{\partial}{\partial t}(\theta_m C_{wm}) + \frac{\partial}{\partial t}(\theta_{im} C_{wim}) + \frac{\partial}{\partial t}(f \rho P_{wm}) + \frac{\partial}{\partial t}[(1-f)\rho P_{wim}] = \frac{\partial}{\partial x_i}\left(\theta_m D_{ij} \frac{\partial C_{wm}}{\partial x_i}\right) - \frac{\partial}{\partial x_j}(q_i C_{wm}) - q_s C_{ws} \quad (11.53)$$

where θ_m and θ_{im} are the fraction of the soil filled with mobile and immobile water, respectively; C_{wm} and C_{wim} are the concentration $[ML^{-3}]$ of species w in the mobile and immobile water, respectively; q_i is the Darcy velocity $[L/T]$; P_{wm} and P_{wim} are adsorbed phase concentration of species ω in the mobile and immobile phase $[M/M]$, respectively; f is the fraction of the sorption sites, which is in direct contact with the mobile liquid ($f = 1$ and $\theta_{im} = 0$ reduce equation 11.53 to a typical single porosity formulation); ρ is soil bulk density $[ML^{-3}]$; q_s is the volumetric flow rate of fluid injection (or withdrawal) per unit volume of the porous medium; and C_{ws} is the concentration of species w in the injected fluid and D_{ij} is the hydrodynamic dispersion tensor defined as

$$\theta_m D_{ij} = d_L |q| \delta_{ij} + (d_L - d_T)\frac{q_i q_j}{|q|} + \theta_m \tau D_c \delta_{ij} \quad (11.54)$$

where d_L and d_T are the longitudinal and the transverse dispersivities, respectively; δ_{ij} is the Kronecker delta; τ is tortuosity; D_c is the coefficient of molecular diffusion; and $|q|$ is the absolute value of the Darcy velocity.

The mobile and immobile phase's concentrations are related as

$$\frac{\partial C_{wim}}{\partial t}[\theta_{im} + (1-f)\rho k_d] = \chi (C_{wm} - C_{wim}) \quad (11.55)$$

where χ is a mass transfer coefficient [T^{-1}] for diffusive mass exchange between the mobile and immobile phases.

Incorporating decay losses λ_{wm} and contaminant loading from a hydrocarbon source to the mobile phase H_w, Equation (11.55) can be written as (see Resources and Systems International, 2000b for details)

$$\frac{\partial C_{wm}}{\partial t}(\theta_m + f\rho k_d) + \frac{\partial C_{wim}}{\partial t}[\theta_{im} + (1-f)\,\rho k_d] = \frac{\partial}{\partial x_i}\left(\theta_m D_{ij}\frac{\partial C_{wm}}{\partial x_j}\right) - q\frac{\partial C_{wm}}{\partial x_i} - q(C_{ws} - C_{wm}) - \lambda_{wm} + H_w \qquad (11.56)$$

Mass transfer due to the contact between the groundwater and the NAPL plume is computed using the mass transfer model (Pfannkuch, 1984)

$$H_w = K_{nw} q_v (C_{weq} - C_{wm}) \qquad (11.57)$$

where q_v is the resultant groundwater velocity, K_{nw} [L^{-1}] is the mass transfer coefficient between the water and NAPL phases.

If an inverse model is selected to simulate the migration of a phase-separate fluid, different governing equations and parameters are needed (Parker and Lenhard, 1990; Butcher and Gauthier, 1994). The solution of the inverse problem for identifying NAPL locations involves estimating dissolution parameters in addition to conventional flow and transport modeling parameters (Khachikian and Harmon, 2000). Descriptions of this approach are discussed in the literature (Kool *et al.*, 1987).

Multiphase flow theory is complex. In addition to spatial distribution of hydraulic conductivity tensor and porosity, there are about 10 additional parameters to be estimated during calibration, adding to the uncertainty typically associated with single-phase saturated flow inverse modeling. Although not unique to multiphase and/or variably saturated inverse modeling, issues to be addressed include:

- Source definition
- Retention parameters
- Problem dimensionality
- History of water and/or tidal fluctuations
- Initial and boundary conditions
- Spatial heterogeneity
- Mesh refinement

Difficulties encountered in defining the source(s) of contamination include the composition of the organic compound, its spatial distribution, release

time, and the volume and flux rate. In some cases the release time and volume can be estimated from inventory records, purchase orders, or invoices. Additional historical information that may be useful for identifying the origin and volume of the release include:

- The facility locations where organic compounds were used, stored, and/or disposed
- Records of leaks/spills that provide information for determining the time, amount, and rate of such releases
- Inventory record reconciliation compared to the volume actually used
- The composition of the organic compound

Regardless of the availability and sufficiency of the historical records to identify the source and/or timing of the release, the spatial distribution of the contaminant in the surface remains a serious challenge for modeling. Most of uncertainty is due to an inaccurate description of soil heterogeneity and anisotropy, retention parameters, fluid density, and interfacial surface tension values. Procedures to age date the release using a direct estimate and nonlinear parameter estimation approach have been reported that may have merit in the absence of this type of historical or site specific information (Mull, 1970; Dracos, 1987).

LNAPL models normally assume that the phase-separate liquid is continuous and/or connected. If this assumption is incorrect, the LNAPL volume in the subsurface can be overestimated. If groundwater monitoring wells are separated by hundreds of feet and the LNAPL thickness is thin (less than a foot), the possibility that the LNAPL is not continuous is highly probable. Other challenges include:

- The accuracy of LNAPL or DNAPL fluid viscosity values
- The selected interfacial tension values for the different liquids
- The values used to represent the soil texture
- The accuracy of any phase-separate measurements

Fluid viscosity and density are rarely measured and are more commonly obtained from the literature. Published values can deviate significantly from actual field conditions due to factors such as weathering and/or co-mingling of the LNAPL with other compounds.

During a typical subsurface oil release and migration, pockets of residual NAPL may form due to fluctuations in the water table. When these fluctuations occur in low permeability strata, the dissolution and volatilization of trapped oil is extremely slow. Pockets of trapped NAPL in the unsaturated

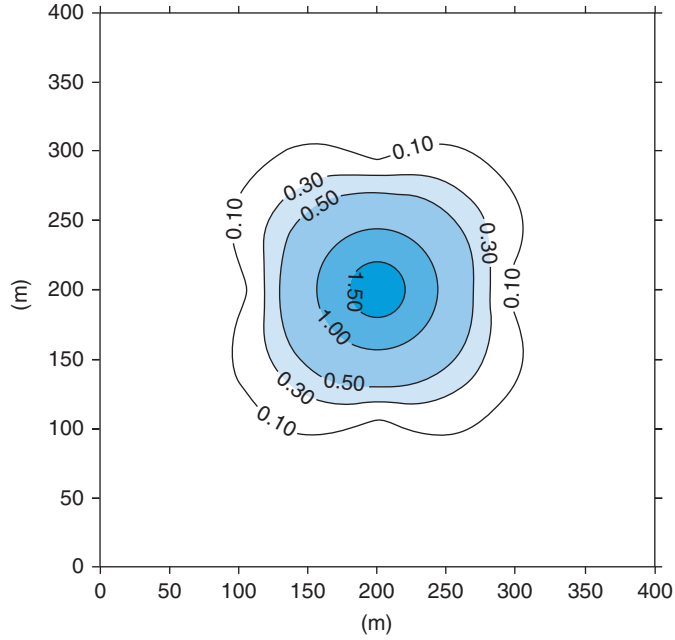


Figure 11.1
Initial apparent product thickness (m) distribution.

zone that are not in direct contact with groundwater may become potential sources of contamination when the water table rises. Similarly, when the groundwater water table decreases, pockets of trapped oil in the groundwater may form a pool of phase-separate hydrocarbon (free oil) or become trapped in the unsaturated zone. Therefore, as water table fluctuates, the mass loading to groundwater can change significantly, adding to the complexities of multiphase inverse modeling. Recharge to groundwater through the unsaturated zone subsequently dissolves some portion of the oil trapped above the water table.

A hypothetical scenario illustrates the phenomena and impact of trapped oil in a fate and transport model. The hypothetical is simulated using BIOSLURP (Resources and Systems International, 2000a), which assumes the initial LNAPL plume (apparent product thickness) distribution shown in Figure 11.1. The water table was allowed to rise 0.5 m over 185 days followed by a drop of 0.5 m over the next six months. Figures 11.2 and 11.3 depict the apparent product thickness at 185 and 365 days, respectively. As the water table decreases, the oil trapped in the saturated zone is liberated, although some remains trapped in the unsaturated zone (generally much smaller than the oil trapped in the saturated zone that is now free and co-mingled with the main pool of LNAPL).

Prior to the occurrence of oil at a given location, the system is treated as a two-phase air-water system described by the van Genuchten (1980) function:

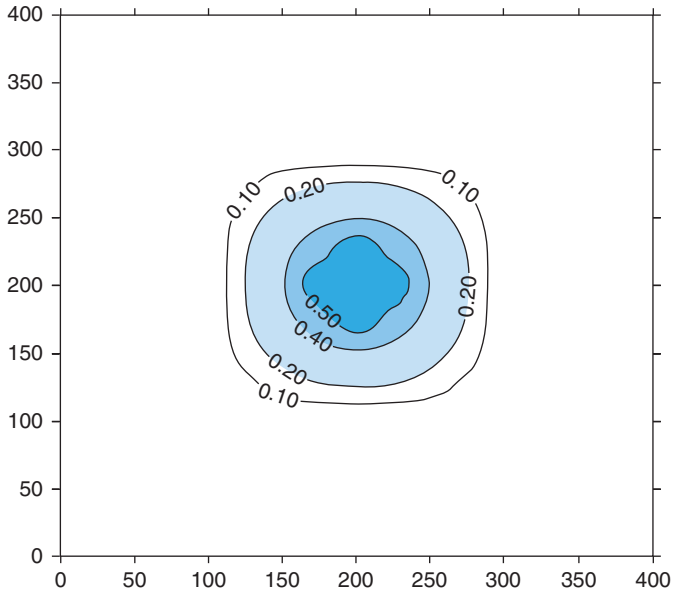


Figure 11.2
Apparent product thickness (m)—185 days.

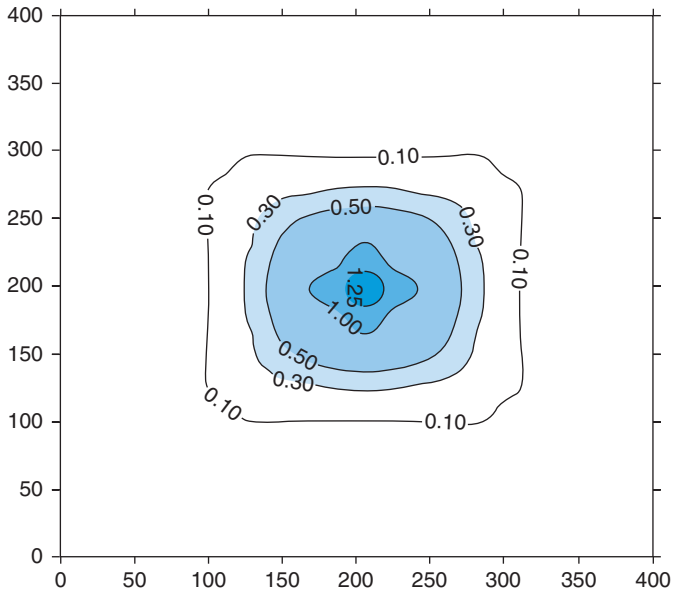


Figure 11.3
Apparent product thickness (m)—365 days.

$$\overline{S}_w = \left[1 + (\alpha \psi_{aw})^n\right]^{-m} \tag{11.58}$$

where $\overline{S}_w = (S_w - S_m)/(1 - S_m)$ is the "effective" water saturation, S_m is the "irreducible" water saturation, $\alpha[L^{-1}]$ and $n[-]$ are porous medium retention parameters, $m = 1 - \frac{1}{n}$, and $\psi_{aw} = h_a - \psi_w$ is the air-water capillary pressure.

Following the occurrence of oil at a location, the system is described by the three-phase relations

$$\overline{\overline{S}}_w = \left[1 + (\alpha\beta_{ow}\psi_{ow})^n\right]^{-m} \tag{11.59}$$

$$\overline{S}_t = \left[1 + (\alpha\beta_{ao}\psi_{ao})^n\right]^{-m} \tag{11.60}$$

where $\overline{\overline{S}}_w$ is the apparent water saturation, $\overline{S}_t$ is the effective total liquid saturation, $\psi_{ow} = \psi_o - \psi_w$ is the oil-water capillary pressure, $\psi_{ao} = h_a - \psi_o$ is air-oil capillary pressure, β_{ao} is a scaling coefficient that may be approximated by the ratio of water surface tension to oil surface tension, α is the water retention factor, and β_{ow} is a scaling factor approximated by the ratio of water surface tension to oil-water interfacial tension. The apparent water saturation is given by

$$\overline{\overline{S}}_w = \overline{S}_w + \overline{S}_{ot} \tag{11.61}$$

where $\overline{S}_{ot}$ is the effective trapped oil saturation.

Soil samples are collected from a site and representative moisture characteristic curves are determined by an appropriate laboratory. Computer programs such as RETC (van Genuchten *et al.*, 1991) and SOILPARA (Resources and Systems International, 2000c) using nonlinear parameter estimation algorithm can be used to estimate soil retention parameters (S_m, α, and n) from the moisture characteristics (saturation vs. suction) data. A user can also decide to estimate the saturated water content (≈ water filled porosity) along with the soil retention parameters. Such solutions, however, are not unique. Generally the uncertainty and error are likely to increase when too many variables are being estimated. A user may estimate and fix values for some parameters (such as saturated water content, and S_m) and use a computer program such as RETC (van Genuchten *et al.*, 1991), and SOILPARA (Resources and Systems International, 2000c) to estimate α and n. When too many variables are estimated in nonlinear parameter estimation programs, the final solution may depend on the initial set of values (guessed) of these parameters.

The complexity of the direct and/or inverse problems increases with dimensionality. Generally three-dimensional multiphase solutions are impractical for large-scale field sites due to the unavailability of sufficient data, excessive computational burden and prolonged investigation, and the high cost of investigation. Generally one-dimensional models are sufficient when vertical migration is dominant. Two-dimensional vertical-slice models such as MOFAT (Katyal *et al.*, 1991) are available to investigate multiphase flow and multicom-

ponent transport. Vertically averaged models such as BIOSLURP (Resources and Systems International, 2000a) are suitable for investigations involving LNAPL migration—application of BIOSLURP to a refinery site with a large oil plume has been reported (Wickland *et al.*, 2000). Three-dimensional models may be considered for focused study on a small section of the site (if this level of investigation is warranted).

The history of spatial and temporal change in fluid tables (air-oil, and air-water) elevation since the site operation commenced is important. The oil is trapped in the unsaturated and saturated zones as the air-oil table falls, and the oil-water table rises, respectively. When this condition reverses, the trapped oil becomes available to mix with the existing free pool of oil. The apparent product thickness observed in wells is indicative of only the free oil. If a free pool of oil at a location is trapped due to rise and/or fall of the fluid tables, the apparent thickness in the well at this location will reduce to zero. Therefore, neglecting historic trapped oil could significantly underestimate the NAPL source at a location.

Historic fluid level data also helps to correctly define the initial and the boundary conditions. Regional groundwater elevations and historic fluid levels versus time for the site are important, if the oil migration is governed by the gradient in the air-oil table (groundwater flows in response to the air-water gradient, although air-water gradient influences the air-oil gradient).

Several independent sets of fluid levels data are desirable when historic data is absent to verify the accuracy of the model parameters. A numerical model often is constructed using first best estimate of various parameters (soil retention, fluid, and hysteresis), and several simulations are performed to refine the parameter values. It is recommended that the user should fix values for parameters that have low uncertainty, and refine parameters (during calibration) that have relatively high uncertainty.

The accurate representation of initial and boundary conditions, and source/sink distribution of a model is crucial. Sufficient numbers of observation wells are needed to ensure that representative fluid levels are available for the source area and over the intended modeling domain. The modeling domain should be large enough to ensure that the phase separate (free) oil plume does not reach the boundary during the simulation. Transient fluid levels data should be used to define boundary condition—specifying air-water, and air-oil tables' fluctuations versus time. Source definition will include initial spatial distribution of free and trapped NAPL, dissolved phase source (mass loading in groundwater recharge versus time for each species). A comprehensive model should update the spatial distribution and mole fractions of all species (in the source) with time. The significance is apparent when relatively more soluble components dissolve/volatilize resulting in their

reduced mole fraction and increased mole fractions of relatively low solubility components in the source. This situation can result in a significant increase in the concentration of low solubility species in recharge to groundwater (especially for historical releases).

Although most multiphase codes offer these capabilities, the computational burden for three-dimensional field scale simulations are significant. BIOFT3D (Resources and Systems International, 2000b) is one of only a few single-phase dual porosity models that has the source update capabilities and that is computationally efficient for large-scale field applications.

Spatial distribution and temporal variation at the internal recharge and/or extraction locations (wells) should be accurately represented. For example, water recharge close to an oil plume can increase the air-oil table elevation enhancing plume mobility, but this is accompanied by a rise in the oil-water table causing excessive trapping of oil in the saturated zone and freeing of the oil trapped in the unsaturated zone. Since the trapped NAPL in the saturated zone is generally higher (by a factor of 2 to 4) than the oil trapped in the unsaturated zones, the overall oil transmissivity will decrease, which decreases the velocity of the NAPL. When historic fluid levels and injection/recharges are unknown, inverse modeling becomes difficult.

Sufficient data should be collected to define heterogeneity in hydraulic conductivity accurately. A network of boring logs describing soil texture at various depths can be used to construct a physically representative conceptual model. Conceptual models should be simplified when sufficient and/or reliable data is unavailable. The dimensionality of the problem may also be reduced. In a complex setting where sufficient data is available, a comprehensive three-dimensional model such as TOUGH2 (Pruess *et al.*, 1999) may be used. It is important that all available information is used, data gaps are identified, and additional data needed is obtained.

An appreciation of the importance of understanding the significance of spatial heterogeneity is illustrated in a hypothetical scenario. Consider a scenario where contamination of a deep aquifer has or may occur in future from a gasoline source (trapped oil) located 20 feet below ground surface (5 to 10 feet thick). The subsurface hydrogeology consists of sand, silt, and clay layers. Several of the clay layers are present in an extensive unsaturated zone. The hypothetical scenario also assumes that reliable data is either available and/or will be collected. A variably saturated flow and transport model incorporating subsurface heterogeneities is created. Such a model will need to include clay layers to mimic the physical reality of the site in terms of retarding groundwater recharge and the vertical migration of dissolved and gaseous plumes (and oil plume in multiphase models), and accurately simulating the lateral migration and development of perched zone on top of low permeability

strata. The contribution of diffusive gas phase transport to groundwater should be investigated and incorporated if significant.

In order to design a model to reflect this hypothetical environment, a variably saturated flow and transport model incorporating subsurface heterogeneities is required. For the clay layers present in the soil column and aquifer, the model requires a description for each clay layer and its impact on liquid (dissolved and NAPL) velocity (laterally and vertically through the clay). The contribution of diffusive gas phase transport within the fluid migrating through the clay layer present in the saturated and unsaturated zone should also be investigated and incorporated, if significant. The same consideration, and mathematical descriptions, for each soil/geologic media through which the contaminant migrates is required to mirror the transport of the fluid through the subsurface.

Another element in creating a multiphase model is its sensitivity to mesh refinement. The results of such an analysis may suggest a sensitivity to mesh refinement near wells and other hydraulic features, source locations, and boundaries. A coarser mesh can introduce significant numerical dispersion although it may compensate for the physical dispersion. The spatial distribution of vertical and lateral aspect ratios should be investigated to avoid serious violations. Additional model parameters such as dispersivity, retardation, and biodegradation should be identified that are representative of the subsurface environment. Literature values for the initial values of longitudinal and transverse dispersivities may be assumed (see earlier discussion). Initial estimates for retardation and bio-decay may be based on literature values and/or laboratory results.

Regardless of the source of the various model parameters, once the model parameters describing the fluid and subsurface environment are identified, it is desirable to refine these estimates through model calibration (Mercer and Cohen, 1990). Parameter values from similar sites should be compared to ensure that values used are within an acceptable range. The mass transfer coefficient is an especially critical calibration parameter. A large mass transfer coefficient value will be equivalent to instant equilibrium between the NAPL and dissolved phase.

It is important to investigate the nature of flow (saturated or variably saturated, steady or transient) prevalent during the period of interest. It may be possible to represent a transient velocity field as a piece-wise steady flow for saturated conditions, but such assumption may cause excessive mass balance errors in the variably saturated simulations. Therefore, it may be desirable to perform coupled (lagged by one time step) flow and transport simulations under variably saturated conditions.

Identification of any mass balance errors should be computed to assess the quality of the model simulations. Significant mass balance errors generally

are reduced by refining time step and/or mesh discretization, ramping boundary conditions, and/or source or sink allocations.

A tracer test investigation may also be performed to assess the degree of anisotropy as well as the presence of preferential flow paths. A series of iterative flow and transport simulations may be helpful to assess the degree of anisotropy and refine estimate of effective porosity.

11.6 CONCLUSIONS

Contaminant transport models for age dating and source identification of a contaminant release must be carefully selected and evaluated to determine the probability that the resulting information will be useful to all parties involved in the engagement. In the context of environmental litigation, the simulation results should be coupled with other corroborative evidence. Model results should be able to withstand intense scientific scrutiny, relative to the purpose for which the model was used.

REFERENCES

Alapati, S. and Kabala, Z. (2000) Recovering the release history of a groundwater contaminant via the non-linear least-squares estimation. *Hydrological Processes* 14, 1003–1016.

Alifanov, O. and Artyukhin, E. (1976) Regularized numerical solution of nonlinear inverse heat-conduction problem. *Journal of Engineering Physics* 29, 934–948.

Ames, K. and Epperson, J. (1997) A kernel-based method for the approximate solution of backward parabolic problems. *Journal of Numerical Analysis* 34, 1351–1471.

Anderson, M. and Woessner, W. (1992) *Applied Groundwater Modeling Simulation of Flow and Advective Transport*. Academic Press, San Diego, CA.

Andreini, M. and Steenhuis, T. (1990) Preferential paths of flow under conventional and conservation tillage. *Geoderma* 46, 85–102.

Ang, A. and Tang, W. (1990) *Probability Concepts in Engineering Planning and Design, Vol. II: Decision, Risk and Reliability*. A. Ang, University of California at Irvine and W. Tang, University of Illinois at Urbana-Champaign.

Aral, M., Guan, J., and Maslia, M. (2001) Identification of contaminant source location and release history in aquifers. *Journal of Hydraulic Engineering* (ASCE) 6(3), 225–234.

Atmadja, J. and Bagtzoglou, A. (2001a) Pollution source identification in heterogeneous porous media. *Water Resources Research* 37, 2113–2125.

Atmadja, J. and Bagtzoglou, A. (2001b) The marching-jury backward beam equation method and its application to backtracking non-reactive plumes in ground-water. Ph.D. dissertation. School of Engineering. Columbia University, New York, NY, 121.

Atmadja, J. and Bagtzoglou, A. (2001c) State of the art report on mathematical methods for groundwater source pollution identification. *Environmental Forensics Journal* 2(3), 205–214.

Baca, E. (1999) On the misuse of the simplest transport model. *Groundwater* 37(4), 483.

Bagtzoglou, A. (1990) Particle-grid methods with application to reacting flows and reliable solute source identification. Ph.D. dissertation. University of Irvine, Irvine, CA, 246.

Bagtzoglou, A., Tompson, A., and Dougherty, D. (1991) Probabilistic simulation for reliable solute source identification in heterogeneous porous media. *Water Resources Engineering Risk Assessment*. NATO Advanced Study Institute Series. Kluwer Academic Publishing. Volume G29, 189–210.

Bagtzoglou, A., Doughery, D., and Thompson, A. (1992) Application of particle methods to reliable identification of groundwater pollution sources. *Water Resources Management* 6, 15–23.

Ball, W. and Roberts, P. (1991) Long term sorption of halogenated organic chemicals by aquifer material. Part 1 Equilibrium. *Environmental Science and Technology* 24, 1223–1236.

Ball, W., Liu, C., Xia, G., and Young, D. (1997) A diffusion-based interpretation of tetrachloroethene and trichloroethane concentration profiles in a groundwater aquitard. *Water Resources Research* 33, 2741–2757.

Balshaw-Biddle, K., Oubre, C., and Ward, C. (1999) Subsurface contamination monitoring using laser fluorescence. Lewis Publishers. Boca Raton, FL, 160.

Barber, L., Thurman, M., Schroder, M., and LeBlanc, D. (1988) Long-term fate of organic micropollutants in sewage contaminated groundwater. *Environmental Science and Technology* 22, 205–211.

Barcelona, M. and Morrison, R. (1988) Sample collection, handling and storage: Water, soils and aquifer solids. In *Proceedings of a National Workshop on Methods for Groundwater Quality Studies*, November 1–3, 1988. Arlington, VA, 49–62. Agricultural Research Division, University of Nebraska at Lincoln, Lincoln, NE.

Bates, D. and Watts, D. (1988) *Nonlinear Regression Analysis and Its Applications*. John Wiley & Sons, New York.

Bear, J. (1979) *Hydraulics of Groundwater*. McGraw-Hill, New York, 569.

Bear, J., Beljin, M., and Ross, R. (1992) Fundamentals of ground-water modeling. EPA Groundwater Issue. United States Protection Agency, Office of Research and Development, Office of Solid Waste and Emergency Response. EPA/540/S-92/005, April 1991.

Berthouex, P. and Brown, L. (1994) *Statistics for Environmental Engineers*. Lewis Publishers, Boca Raton, FL.

Birchwood, R. (1999) Identifying the location and release characteristics of a groundwater pollution source using spectral analysis. In *Proceedings of the 19th Annual American Geophysical Union Hydrology Days Conference*, Colorado State University, Fort Collins, Colorado, August 16–18, 1999, 37–50.

Bonazountas, M., Hetrick, D., Kostecki, P., and Calabrese, E. (1997) *SESOIL in Environmental Fate and Risk Modeling*. Amherst Scientific Publishing. Amherst, MA.

Bradley, P., Chapelle, F., and Landmeyer, J. (2001a) Methyl t-butyl ether mineralization in surface-water sediment microcosms under denitrifying conditions. *Applied and Environmental Microbiology* 67(4), 1975–1978.

Bradley, P., Chapelle, F., and Landmeyer, J. (2001b) Effect of redox conditions on MTBE biodegradation in surface water sediments. *Environmental Science and Technology* 35(23), 4643–4647.

Brewster, M., Annan, A., Greenhouse, J., Kueper, B., Olhoeft, G., Redman, J., and Sander, K. (1995) Observed migration of a controlled DNAPL release by geophysical methods. *Journal of Ground Water* 33(6), 977–987.

Brown, A. (2000) Treatment of drinking water impacted with MTBE. In *Mealey's MTBE Conference*. May 11–12, 2000. Marina del Rey, CA, 475–518. Mealey Publications. King of Prussia, PA.

Buddemeier, R. and Rego, J. (1986) Colloidal radionuclides in groundwater. *Annual Report*. Lawrence Livermore National Laboratory, Livermore, CA. Report No. UCAR 10062/85–1.

Butcher, J. and Gauthier, T. (1994) Estimation of residual dense NAPL mass by inverse modeling. *Groundwater* 32(1), 71–78.

Butler, E. and Hayes, K. (1999) Kinetics of the transformation of trichloroethylene and tetrachloroethylene by iron sulfide. *Environmental Science and Technology* 33(12), 2021–2027.

Carslaw, H. and Jaeger, J. (1959) *Conduction of Heat in Solids*. Clarendon Press, Oxford.

Choy, B. and Reible, D. (1999) *Diffusion Models of Environmental Transport*. Lewis Publishers, Boca Raton, FL, 183.

Christy, T. (1998) A permeable membrane sensor for the detection of volatile compounds in soil. In *Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems*. Chicago, IL.

Clark, G. and Oppenheimer, S. (1994) Quasi-reversibility methods for non-well-posed problems. *Journal of Differentiation Equations* 1–9. No. 9.

Cleary, R. (1995) Introduction to applied mathematical modeling in groundwater pollution and hydrology with IBM-PC applications. *NGWA IBM-PC Applications in Groundwater Pollution and Hydrology*, Chapter 4. August 13–18, 1995. San Francisco, CA.

Cohen, R., Mercer, J., and Matthews, J. (1993) *DNAPL Site Evaluation*, EPA/600/R-93/022. US Environmental Protection Agency, Ada, OK.

Crank, J. (1985) *The Mathematics of Diffusion*, 2e. Clarendon Press, Oxford, 414.

Daily, W., Ramirez, A., and Johnson, R. (1998) Electrical impedance tomography of a perchloroethylene release. *Journal of Environmental and Engineering Geophysics* 2(3), 189–201.

Davidson, J. (2000) Fate, transport and remediation of MTBE. In *Mealey's MTBE Conference*. May 11–12, 2000, Marina del Rey, CA, 445–474. Mealey Publications, King of Prussia, PA.

Deeds, N., McKinney, D., and Pope, G. (2000) Laboratory characterization of non-aqueous phase liquid/tracer interaction in support of a vadose zone partitioning interwell tracer text. *Journal of Contaminant Hydrology* 41, 193–204.

Dilling, W., Tefertiller, N., and Kallos, G. (1975) Evaporation rates and reactivities of methylene chloride, chloroform, 1,1,1, trichloroethane, trichloroethylene, tetrachloroethylene, and other chlorinated compounds in dilute aqueous solutions. *Environmental Science and Technology* 9(6), 833–838.

Dimov, I., Jaekel, U., and Vereecken, H. (1996) A numerical approach for determination of sources in transport equations. *Computers in Mathematical Applications* 32, 31–42.

Domenico, P. (1987) An analytical model for multidimensional transport of a decaying contaminant species. *Journal of Hydrology* 91, 49–58.

Domenico, P. and Schwartz, F. (1990) *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.

Donaldson, J. and Schnabel, R. (1987) Computation experience with confidence regions and confidence intervals for nonlinear least squares. *Technometrics* 29(1), 67–82.

Dracos, T. (1987) Immiscible transport of hydrocarbons infiltrating in unconfined aquifers. In *Proceedings of Symposium on Oil in Fresh Water*, Alberta, Canada, J. van der Molen and S. Hrudey, S. eds. Pergamon Press, Oxford, 159–272.

Drogos, D. (2000) MTBE v. other oxygenates. In *Mealey's MTBE Conferenc*e, May 11–12, 2000, Marina del Rey, CA, 299–334. Mealey Publications, King of Prussia, PA.

Duke, L., Rong, D., and Harmon, T. (1998) Parameter-induced uncertainty in modeling vadose zone transport of VOCs. *Journal of Environmental Engineering* 124(5), 441–448.

Elden, L. (1982) Time discretization in the backward solution of parabolic equations. II. *Mathematical Computations* 39(159), 69–84.

Erickson, R. and Morrison, R. (1995) *Environmental Reports and Remediation Plans: Forensic and Legal Review*. John Wiley & Sons, Environmental Law Library, New York.

Ewing, R. (1975) The approximation of certain parabolic equations backward in time by Sobolev equations. *Journal of Mathematical Analysis*. 2, 283–294.

Feenstra, S. and Cherry, J. (1998) Aqueous concentration ratios to estimate mass of multi-component NAPL residual. In *Nonaqueous Phase Liquids: Remediation of Recalcitrant Compounds*, G.B. Wickramanayake and R. Hinchee, eds. Battelle Press, Columbus, OH, 55–60.

Fetter, C. (1994) *Applied Hydrogeology*, 3e. Prentice Hall, Upper Saddle River, NJ.

Fick, A. (1855) Ueber diffusion. *Annalen der Physik (Leipzig)* 170, 59–86.

Fishman, M. (1998) DNAPL infiltration and distribution in subsurface: 2D experiment and modeling approach. In *Nonaqueous-Phase Liquids: Remediation of Chlorinated and Recalcitrant Compounds*, G. Wickramanayake and R. Hinchee, eds. Battelle Press, Columbus, OH, 37–42.

Gao, J., Skeen, R., and Hooker, B. (1995) Effect of temperature on perchloroethylene dechlorination by a methanogenic consortium. In *Bioremediation of Chlorinated Solvents*, R. Hinchee, A. Leeson, and L. Semprini, eds. Battelle Press, Columbus, OH, 53–59.

Gauthier, T.D. and Murphy, B.L. (2003) Age dating groundwater plumes based on the ratio of 1,1-dichloroethylene to 1,1,1-trichloroethane: An uncertainty analysis. *Environmental Forensics* 4(3), 205–214.

Gelhar, L. and Axness, C. (1981) *Stochastic Analysis of Macro-dispersion in Three Dimensionally Heterogeneous Aquifers*. Report No. H-8, Hydraulic Research Program. New Mexico Institute of Mining and Technology, Soccorro, New Mexico.

Gelhar, L. and Axness, C. (1983) Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resources Research* 19(1), 161–180.

Gelhar, L., Welty, C., and Rehfeldt, K. (1992) A critical review of data on field-scale dispersion in aquifers. *Water Resources Research* 28(17), 1955–1974.

Ghadiri, H. and Rose, C. (1992) *Modeling Chemical Transport in Soils Natural and Applied Contaminants*. Lewis Publishers, Chelsea, MI.

Gorelick, S. and Remson, I. (1982) Optimal dynamic management of groundwater pollutant sources. *Water Resources Research* 18(1), 71–76.

Gorelick, S., Evans, B., and Remson, I. (1983) Identifying sources of groundwater pollution: An optimization approach. *Water Resources Research* 19(3), 779–790.

Grant, S. (2000) Physical and chemical factors affecting contaminant hydrology in cold environments. United States Army Corps of Engineers, Engineer Research and Development Center. Cold Regions Research and Engineering Laboratory. Hanover, New Hampshire. ERDC/CREEL TR-00-21.

Ham, P. (2006) Dispersion controlled natural attenuation. The role of conservative plume characteristics in modeling reactive mixing processes. Ph.D. Dissertation Thesis. University of Utrecht. Netherlands.

Hetrick, D. and Pandey, A. (1999) A methodology for establishing cleanup objectives in the unsaturated soil zone using sensitivity and uncertainty analysis for chemical fate and transport. *Journal of Soil Contamination* 5(5), 559–576.

Hillel, D. (1980) *Introduction to Soil Physics*. Academic Press, Orlando, FL, 365.

Hine, J. and Mookerjee, P. (1975) The intrinsic hydrophilic character of organic compounds. Correlations in terms of structural contributions. *Journal of Organic Chemistry* 40(3), 292–298.

Hughes, B., Gillham, R., and Mendoza, C. (1992) Transport of trichloroethylene vapors in the unsaturated zone: A field experiment. *Proceedings of the Conference on Subsurface Contamination by Immiscible Liquids*, International Association of Hydrogeologists, Calgary, Alberta, Balkema, Rotterdam, 81–88.

Huling, S. and Weaver, J. (1991) Dense nonaqueous phase liquids. United States Environmental Protection Agency, Office of Solid Waste and Emergency Response. USEPA/540/4-91-002.

Hunkeler, D., Hoehn, E., Hohener, P., and Zeyer, J. (1997) ^{222}Rn as a partitioning tracer to detect diesel fuel contamination in aquifer: Laboratory study and field observations. *Environmental Science and Technology* 31(11), 3180–3187.

Huyakorn, P.S., Pandey, S., and Wu, Y. (1994) A three-dimensional multiphase flow model for assessing NAPL contamination in porous and fractured media, 1. Formulation. *Journal of Contaminant Hydrology* 16, 109–130.

Imhoff, P., Thyrum, G., and Miller, C. (1996) Dissolution fingering during the solubilization of nonaqueous phase liquids in saturated porous media. 2. Experimental observations. *Water Resources Research* 32(7), 1929–1942.

Jarvis, N. (1998) Modeling the impact of preferential flow on nonpoint source pollution. In *Physical Nonequilibrium in Soils: Modeling and Application*, H. Selim and L. Ma, eds. Ann Arbor Press, Chelsea, MI, 195–221.

Jin, M., Delshad, M., Dawarakanath, V., McKinney, D., Pope, G., Sepehrnooci, *et al.* (1995) Partitioning tracer test for detection, estimation and remediation performance assessment of subsurface nonaqueous phase liquids. *Water Resources Research* 31(5), 1201–1211.

Jin, M., Butler, M., Jackson, R., Mariner, P., Pickens, J., Pope, G. *et al.* (1997) Sensitivity model design and protocol for partitioning tracers in alluvial aquifers. *Groundwater* 35(6), 964–972.

Jury, W., Spencer, W., and Farmer, W. (1984) Behavior assessment model for trace organics in soil: III. Application of screening mode. *Journal of Environmental Quality* 13(4), 573–579.

Jury, W., Sposito, G., and White, R. (1986) A transfer function model of solute transport through soil. 1. Fundamental Concepts. *Water Resources Research* 22, 243–247.

Jury, W. and Roth, K. (1990) *Transfer Functions and Solute Movement through Soil: Theory and Applications*. Birkhauser Verlag, Basel.

Jury, W. and Fuller, H. (1992) Transport of chemicals through soil: Mechanisms, models and field applications. *Advances in Agronomy*. Academic Press, London, 141–201.

Kaluarachchi, J., Parker, J., and Lenhard, R. (1990) A numerical model for areal migration of water and light hydrocarbon in unconfined aquifers. *Advances in Water Resources* 13, 29–40.

Kane, S., Beller, H., Legler, T., Koster, C., Pinkart, H., Halden, R., and Happel, A. (2001) Aerobic biodegradation of methyl tert-butyl ether by aquifer bacteria from leaking underground storage tanks. *Applied and Environmental Microbiology* 67(12), 5824–5829.

Katyal, A., Kaluarachchi, J., and Parker, J. (1991) MOFAT—A two-dimensional finite element program for multiphase flow and multi-component transport. R. S. Kerr Environmental Research Laboratory, USEPA, ADA, OK.

Katyal, A. and Parker, J. (1992) An adaptive solution domain algorithm for efficient solution of multiphase problems. *Computers and Geosciences* 18, 1–9.

Keller, C., Rossabi, J., Riha, B., Walker, J., Nashold, B., and Werenko, P. (2000) Utility of flexible liners installed via push rods. In *Proceedings of the Third International Symposium on Integrated Approaches to Site Characterization*. Chicago, IL. September 12–14, 2000.

Kezsbom, A. and Goldman, A. (1991) The boundaries of groundwater modeling under the law: Standards for excluding speculative expert testimony. *Environmental Claims Journal* 4(1), 5–30.

Kool, J., Parker, J., and van Genuchten, M. (1987) Parameter estimation for unsaturated flow and transport models: A review. *Journal of Hydrology* 91, 255–293.

Kornfeld, I. (1992) Comment to the boundaries of groundwater modeling under the law: Standards for excluding speculative expert testimony. *Tort and Insurance Law Journal* 28(1), 59–68.

Kosmatka, S. and Panarese, W. (1988) *Design and Control of Concrete Mixtures*, 13e. Portland Cement Association, Skokie, IL.

Kram, M. (1998) Use of SCAPS petroleum hydrocarbon sensor technology for real-time indirect DNAPL detection. *Journal of Soil Contamination* 17(1), 73–86.

Kung, J. (1990a) Preferential flow in a sandy vadose zone. 1. Field observation. *Geoderma* 45, 51–58.

Kung, J. (1990b) Preferential flow in a sandy vadose zone. 2. Mechanism and implications. *Geoderma* 46, 59–71.

Lieberman, S. and Knowles, D. (1998) Cone penetrometer deployable in situ video microscope for characterizing sub-surface soil properties. *Field Analytical and Technology* 2(2), 127–132.

Lincoff, A. and Gossett, J. (1984) The determination of Henry's law constant for volatile organics by equilibrium partitioning in closed systems. In *Gas Transfer at*

Water Surfaces, W. Brutsaert and G. Jirka, eds. Reidel Publishing, Germany, 17–25.

Liu, C. and Ball, W. (1999) Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware. *Water Resources Research* 35(7), 1975–1985.

Liu, J. (1995) Travel time and location probabilities for groundwater contaminant sources. Master's Thesis. Department of Earth and Environmental Science, New Mexico Institute of Mining and Technology. Socorro, New Mexico.

Lyman, W., Reehl, W., and Rosenblatt, D. (1982) *Handbook of Chemical Estimation Methods*. McGraw-Hill, New York, NY.

Mabey, W. and Mill, T. (1978) Critical review of hydrolysis of organic compounds in water under environmental conditions. *Journal of Physical and Chemical Reference Data* 7, 383–415.

MacDonald, J. (1995) Solution of an "impossible" diffusion-inversion problem. *Computers in Physics* 9, 546–553.

Maha, P. and Datta, B. (1997) Optimal monitoring network and ground-water pollution source identification. *Journal of Water Resource Planning and Management* 123, 199–207.

Maha, P. and Datta, B. (2000) Identification of pollution sources in transient groundwater systems. *Water Resources Management* 14(3), 209–227.

Maha, P. and Datta, B. (2001) Optimal identification of groundwater pollution sources and parameter estimation. *Journal of Water Resources Planning and Management* (ASCE) 27(1), 20–29.

Mahinthakumar, K. and Sayeed, M. (2005) Hybrid genetic algorithm-local search methods for solving groundwater source identification inverse problems. *Journal of Water Resources Planning and Management* (ASCE) 131(1), 45–57.

Mahinthakumar, K. and Sayeed, M. (2006) Reconstructing groundwater source release histories using hybrid optimization approaches. *Environmental Forensics* 7(1), 45–54.

Martin-Hayden, J. and Robbins, G. (1997) Plume distortion and apparent attenuation due to concentration averaging in monitoring wells. *Groundwater* 35(2), 339–347.

McCarty, P. (1993) In situ bioremediation of chlorinated solvents. *Biotechnology* 4, 323–330.

McCarty, P. (1994) An overview of anaerobic transformation of chlorinated solvents. *Symposium on Natural Attenuation of Groundwater*. USEPA/600/R-94/162. Office of Research and Development, Denver CO. August 30–September 1, 1994, 104–108.

McCoy, B. and Rolston, D. (1992) Convective transport of gases in moist porous media: Effect of absorption, adsorption, and diffusion in soil aggregates. *Environmental Science and Technology* 26(12), 2468–2476.

McNab, W., Rice, D., Bear, J., Ragaini, R., Tuckfield, C., and Oldenburg, C. (1999) Historical case analysis of chlorinated volatile organic compound plumes. Lawrence Livermore National Laboratory, University of California. Environmental Protection Department. UCRL-AR-133361. 31 (http://www-erd.llnl.gov/library/AR-133361.pdf).

Mercer, J. and Cohen, R. (1990) A review of immiscible fluids in the subsurface: Properties, models, characterization and remediation. *Journal of Contaminant Hydrology* 6, 107–163.

Mercer, J. (1991) Common mistakes in model applications. In *Proceedings of the American Society of Civil Engineers (ASCE) Symposium on Groundwater*, Nashville, TN. July 29–August 1, 1991. American Society of Civil Engineers, Washington, DC.

Miller, C., Gleyzer, S., and Imhoff, P. (1998) Numerical modeling of NAPL dissolution fingering in porous media. In *Physical Nonequilibrium in Soils Modeling and Application*, H. Selim and L. Ma, eds. Ann Arbor Press, Chelsea, MI, 389–415.

Miller, R., Biggar, J., and Nielsen, D. (1965) Chloride displacement in Panoche clay loam in relation to water movement and distribution. *Water Resources Research* 1, 63–73.

Millington, R. and Quirk, J. (1959) Permeability of porous media. *Nature (London)* 183, 387–388.

Montgomery, J. (1991) *Groundwater Chemicals Field Guide*. Lewis Publishers, Chelsea, MI.

Morrison, R. and Lowry, B. (1990) Sampling radius of a porous cup sampler: Experimental results. *Groundwater* 28(2), 262–267.

Morrison, R. (1998) Estimating the timing of a contaminant release via transport modeling. *Environmental Claims Journal* 10(2), 75–90.

Morrison, R., Hartman, B., Jones, J., and Erickson, R. (1998) Chlorinated Solvent Contamination: Technical and Legal Issues, Argent Communications Group, Foresthill, CA.

Morrison, R. and Newell, E. (1999) The cosolvation transport of DDT and xylene at a pesticide formulation facility. *Journal of Soil Contamination* 8(1), 63–80.

Morrison, R. (1999a) *Environmental Forensics. Principles and Applications*. CRC Press, Boca Raton, FL, 351.

Morrison, R. (1999b) Reverse and confirmation models: Applications and challenges. *Environmental Claims Journal* 12(1), 103–117.

Morrison, R. (2000a) Application of forensic techniques for age dating and source identification in environmental litigation. *Environmental Forensics* 1(3), 131–153.

Morrison, R. (2000b) Critical review of environmental forensic techniques. Part II. *Environmental Forensics* 1(4), 175–195.

Morrison, R. (2000c) Critical review of environmental forensic techniques. Part I. *Environmental Forensics* 1(4), 157–173.

Morrison, R. and Murphy, B. (2005) *Environmental Forensics: A Contaminant Specific Guide*. Elsevier Publishing. London, UK.

Mueller, J., Chapman, P., and Prichard, P. (1989) Action of a fluoranthene-utilizing community of polycyclic aromatic hydrocarbon components of creosote. *Applied and Environmental Microbiology* 55(12), 3085–3090.

Mull, R. (1970) The migration of oil products with regard to groundwater pollution by oil. In *Proceedings of the 5th International Conference on Advances in Water Pollution Control Research*, Moscow, S. Jenkins, ed. Pergamon Press, Oxford, 1–8.

Nelson, N., Oostrom, M., Wietsma, T., and Brusseau, M. (1999) Partitioning tracer method for the in situ measurements of DNAPL saturation. Influence of heterogeneity and sampling method. *Environmental Science and Technology* 33(22), 4046–4053.

Neuman, S. (1990) Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resources Research* 26(8), 1749–1758.

Neupauer, R. and Wilson, J. (1999) Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. *Water Resources Research* 35(11), 3389–3392.

Neupauer, R., Borhcers, B., and Wilson, J. (2000) Comparison of inverse methods for reconstructing the release history of a groundwater contamination source. *Water Resources Research* 36(9), 2469–2475.

Neupauer, R. and Wilson, J. (2004) Forward and backward location probabilities for sorbing solutes in groundwater. *Advances in Water Resources* 27(7), 2261–2269.

Newell, C. and Ross, R. (1991) Estimating potential occurrence of DNAPL at Superfund Sites. Quick Reference Guide Sheet. United States Environmental Protection Agency, Publication No. 9355.4-07FS. 3.

Odencrantz, J., Farr, J., and Robinson, D. (1992) Transport model parameter sensitivity for soil cleanup level determinations using SESOIL and AT123D in the context of the California leaking underground fuel tank field manual. *Journal of Soil Contamination* 1(2), 159–182.

Odermatt, J., Johnson, T., and Hummeldorf, R. (1993) Distribution of DDT residues (DDT, DDD and DDE) in California sediments. *Journal of Soil Contamination* 2(4), 315–329.

Olson, R. and Davis, A. (1990) Predicting the fate and transport of organic compounds in groundwater. *Hazardous Material Control* 3(3), 39–64.

Opdyke, D. and Loehr, R. (1999) Statistical analysis of chemical release rates from soils. *Journal of Soil Contamination* 6(5), 541–558.

Pandey, S., Forsyth, P., Falta, R., Wu, Y., and Huyakorn, P. (1995) Considerations for robust compositional simulations of subsurface non-aqueous phase liquid contamination and remediation. *Water Resources Research* 31, 1273–1289.

Pankow, J. and Rosen, M. (1988) Determination of volatile compounds in water by purging directly to a capillary column with whole column cryotrapping. *Environmental Science and Technology* 22(4), 398–405.

Pankow, J., Feenstra, S., Cherry, J., and Ryan, M. (1996) Dense chlorinated solvents in groundwater: Background and history of the problem. In *Dense Chlorinated Solvents and Other DNAPLs in Groundwater*, J. Pankow and J. Cherry, eds. Waterloo Press, Portland, OR, 1–46.

Parker, B., Gillham, R., and Cherry, J. (1994) Diffusive disappearance of dense, immiscible phase organic liquids in fractured geologic media. *Groundwater* 32(5), 805–820.

Parker, J. and Lenhard, R. (1990) Vertical integration of three phase flow equations for analysis of light hydrocarbon plume movement. *Transport in Porous Media* 5, 618–624.

Parker, J. (2000) Inverse plume modeling to identify release locations or timing. International Business Communications 3rd Annual Executive Forum on Environmental Forensics. *Applying Effective Scientific Methods to Decrease Cost and Liability*, June 26–28, 2000, Washington, DC.

Patzek, T. (2001) The forensic application of contaminant transport models; Chapter 6. In *Practical Environmental Forensics Process and Case Histories*, P. Sullivan, F. Agardy, and R. Traub, eds. John Wiley & Sons, Inc. Toronto, ON, 123–150.

Pfannkuch, H. (1984) Determination of the contaminant source strength from mass exchange processes at the petroleum groundwater interface in shallow aquifer systems, *Int. Conf. on Petroleum Hydrocarbons and Organic Chemicals*, NWWA, Houston, TX.

Philip, J. (1975) The growth of disturbances in unstable infiltration flows. *Soil Science. Society of America Proceedings* 39, 1049–1053.

Pohlmann, K., Icopini, G., McArthur, R., and Rosal, C. (1994) Evaluation of sampling and field-filtration methods for the analysis of trace metals in groundwater. United States Environmental Protection Agency. Environmental Monitoring Systems Laboratory Project Summary. EPA/600/SR-94/119.

Pruess, K., Oldenburg, C., and Moridis, G. (1999) TOUGH2 *User's Guide, Version 2.0,* Lawrence Berkeley National Laboratory Report LBNL-43134, Berkeley, CA.

Puls, R., Eychaner, J., and Powell, R. (1990) Colloidal-facilitated transport of inorganic contaminants in groundwater: Part I. Sampling considerations. United States Environmental Protection Agency. Environmental Research Brief. EPA/600/M-90/023.

Quisenberry, V. and Phillips, R. (1976) Percolation of surface-applied water in the field. *Soil Science Society of America Journal* 40, 484–489.

Ramamoorthy, S. and Ramamoorthy, S. (1998) *Chlorinated Organic Compounds in the Environment. Regulatory and Monitoring Assessment*. Lewis Publishers, Boca Raton, FL.

Resources and Systems International, Inc. (2000a) BIOSLURP—A code to simulate vacuum enhanced recovery of LNAPLs, *Technical Documentation & User Guide*. Blacksburg, VA.

Resources and Systems International, Inc. (2000b) BIOF&T—A code to simulate flow and transport in saturated and unsaturated zones, *Technical Documentation & User Guide*. Blacksburg, VA.

Resources and Systems International, Inc. (2000c) SOILPARA—A code to estimate retention parameters, *Technical Documentation & User Guide*. Blacksburg, VA.

Richards, L. (1931) Capillary conduction of liquids in porous media. *Physics* 1, 318–333.

Riha, B., Rosabi, J., Eddy-Dilek, C., and Jackson, D. (2000) DNAPL characterization using the ribbon DNAPL sampler: Methods and Results. In *Proceedings of the Second International Conference on Remediation of Chlorinated and Recalcitrant Compounds*. May 22–25, 2000. Monterey, CA. Battelle Press. Columbus, OH.

Ritsema, C. and Dekker, L. (1995) Distribution flow, a general process in the top layer of water repellent soils. *Water Resources Research* 31, 1187–1200.

Robbins, G. (1989) Influence of using purged and partially penetrating monitoring wells on contaminant detection, mapping and modeling. *Groundwater* 27(1), 155–162.

Roberts, P. and Dandliker, P. (1983) Mass transfer of volatile organic contaminants from aqueous solution to the atmosphere during surface aeration. *Environmental Science and Technology* 17(8), 484–489.

Robertson, W., Barker, J., LeBeau, Y., and Marcoux, S. (1984) Contamination of an unconfined sand aquifer by waste pulp liquor. A case study. *Groundwater* 22(2), 191–197.

Rong, Y. (1995) Uncertainty analyses of a vadose zone transport model for soil remediation criteria of volatile organic compounds. PhD dissertation. University of California at Los Angeles, School of Public Health, Los Angeles, CA.

Rong, Y. (1999) A study of vadose zone transport model VLEACH. *Journal of Soil Contamination* 8(2), 217–229.

Rong, Y., Wang, R., and Chou, R. (1998) Monte Carlo simulation for a groundwater mixing model in soil remediation of tetrachloroethylene. *Journal of Soil Contamination* 7(1), 87–102.

Rong, Y. and Wang, R. (2000) Monte Carlo vadose zone model for soil remedial criteria. *Soil and Sediment Contamination* 9(6), 593–610.

Rosenbloom, J., Mock, P., Lawson, P., Brown, J., and Turin, H. (1993) Application of VLEACH to vadose zone transport of VOCs at an Arizona Superfund site. *Groundwater Monitoring and Remediation* 13(3), 159–163.

Rossabi, J. and Nave, S. (1998) Characterization of DNAPL using fluorescence techniques. (U), WSRC-TR-98–00125. Westinghouse Savannah River Company. Aiken, SC.

Russo, D. and Dagan, G. (1991) On solute transport in a heterogeneous porous formation under saturated and unsaturated water flows. *Water Resources Research* 27, 285–292.

Salhotra, A., Mineart, P., Hansen, S., and Allison, T. (1993) Multimed, the Multimedia Exposure Assessment Model for Evaluating the Land Disposal of Wastes—Model Theory. EPA 600/R-93/081, May 1993.

Sanders, P. (1997) Calculation of soil cleanup criteria for volatile organic compounds as controlled by the soil-to-groundwater pathway: Comparison of four unsaturated soil zone leaching models. In *Sesoil in Environmental Fate and Risk Modeling*, M. Bonazountas, D. Hetrick, P. Kostecki, and E. Calabrese, eds. Amherst Scientific Publishing. Amherst, MA.

Scheidegger, A. (1960) The physics of flow through porous media. University of Toronto Press, Toronto, ON.

Schulin, R., Van Genuchten, M., Fluhler, H., and Ferlin P. (1987) An experimental study of solute transport in a stony field soil. *Water Resources Research* 23, 1785–1794.

Schwarzenbach, R., Giger, P., Hoehn, W., and Schneider, J. (1983) Behavior of organic compounds during infiltration of river water to groundwater. Field Studies. *Environmental Science and Technology* 17, 472–479.

Schwille, F. (1988) *Dense Chlorinated Solvents in Porous and Fractured Media—Model Experiments.* Translated by J. Pankow. Lewis Publishers, Boca Raton, FL.

Selim, H. and Ma, L. (1998) *Physical Nonequilibrium in Soils Modeling and Application*. Ann Arbor Press, MI.

Semprini, L., Hopkins, O., and Tasker, B. (2000) Laboratory, field and modeling studies of Radon-222 as a natural tracer for monitoring NAPL contamination. *Transport in Porous Media* 28, 223–240.

Sheppard, J., Campbell, M., and Kittrick, J. (1979) Retention of Neptunium, Americium and Curium by diffusible soil particles. *Environmental Science and Technology* 13(6), 680–684.

Sheppard, J., Campbell, M., Cheng, T., and Kittrick, J. (1980) Retention of radionuclides by mobile humic compounds. *Environmental Science and Technology* 14(11), 1349–1353.

Sidauruk, P., Cheng, A., and Ouazar, D. (1998) Groundwater contaminant source and transport parameter identification by correlation coefficient optimization. *Groundwater* 36(1), 208–214.

Skaggs, T. and Kabala, Z. (1994) Recovering the release history of a groundwater contaminant. *Water Resources Research* 30, 71–79.

Skaggs, T. and Kabala, Z. (1995) Recovering the history of a groundwater contaminant plume: Method of quasi-reversibility. *Water Resources Research* 31, 2669–2673.

Skaggs, T. and Kabala, Z. (1998) Limitations in recovering the history of a groundwater contaminant plume. *Journal of Contaminant Hydrology* 33, 347–359.

Snodgrass, M. and Kitanidis, P. (1997) A geostatistical approach to contaminant source identification. *Water Resources Research* 33, 537–546.

Sonnenborg, T., Engesgaard, O., and Rosbjerg, P. (1996) Contaminant transport at a waste residue deposit. I. Inverse flow and nonreactive transport modeling. *Water Resources Research* 32, 925–938.

Sposito, G., Gupta, V., and Bhattacharya, R. (1979) Foundational theories of solute transport in porous media: A critical review. *Advances in Water Resources* 2, 59–69.

Sun, N-Z. (1994) Inverse problems in groundwater modeling; Chapter 6. In *Theory and Applications of Transport in Porous Media*, J. Bear, ed. Kluwer Academic Publishers, The Netherlands.

Trevizo, C., Daniel, D., and Nirmalakhandan, N. (2000) Screening alternative degreasing solvents using multivariate analysis. *Environmental Science and Technology* 34(12), 2587–2595.

Tsien, H., Brusseau, G., Hanson, R., and Wackett, L. (1989) Biodegradation of trichloroethylene by methylosinus trichosporium OB3b. *Applied Environmental Microbiology* 55(12), 3155–3161.

United States Environmental Protection Agency. (1980) *Treatability Manual—Volume I: Treatability Data*. Office of Research and Development. USEPA Report 600/8-80-042a.

United States Environmental Protection Agency. (1992) Estimating potential for occurrence of DNAPL at Superfund Sites. Office of Solid Waste and Emergency Response. Publication No. 9355.4-07FS.

United States Environmental Protection Agency. (1993) Evaluation of the likelihood of DNAPL presence at NPL site. Office of Emergency and Remedial Response. Hazardous Site Control Division. USEPA/540/R-93/073. Washington, DC.

United States Environmental Protection Agency. (1996) Region 9 Preliminary Remediation Goals (PRGs) for 1996. USEPA, Region 9, San Francisco, CA. August 1, 1996.

United States Environmental Protection Agency. (1998) Geophysical techniques to locate DNAPLs: Profiles of federally funded projects. US USEPA, Office of Solid Waste and Emergency Response. Technology Innovation Office, Washington, DC. USEPA 542-R-98-020. December 1998.

Vand, V. (1948) Viscosity of solutions and suspensions, I. *Journal of Physical Chemistry* 52, 242–246.

van Genuchten, M. and Wierenga, P. (1976) Mass transfer studies in sorbing media, 1. Analytical solutions. *Soil Science Society of America Journal* 40, 473–480.

van Genuchten, M. (1980) A closed-form solution for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* 44, 892–898.

van Genuchten, M., Leij, F., and Yates, S. (1991) The RETC code for quantifying the hydraulic functions of unsaturated soils, EPA/600/2-91/065, USEPA, R.S. Kerr Environmental Research Laboratory, Ada, OK.

Villholth, K. (1999) Colloid characterization and colloidal phase partitioning of polycyclic aromatic hydrocarbons in two creosote-contaminated aquifers in Denmark. *Environmental Science and Technology* V33(5), 691–699.

Vogel, T. and McCarty, P. (1987a) Abiotic and biotic transformations of 1,1,1-trichloroethane under methanogenic conditions. *Environmental Science and Technology* 21, 1208–1213.

Vogel, T., Criddle, C., and McCarty, P. (1987b) Transformation of halogenated aliphatic compounds. *Environmental Science and Technology* 21(8), 722–736.

Waddell, M., Domoracki, W., and Temples, T. (1996) Use of seismic reflection amplitude versus offset (AVO) techniques to image dense nonaqueous phase liquids (DNAPL) at the M-area seepage basin. Savannah River Site, South Carolina. University of South Carolina, Earth Sciences and Resources Institute and the Board of Trustees.

Wagner, B. (1992) Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling. *Journal of Hydrology* 135, 275–300.

Warner, H., Cohen, J., and Ireland, J. (1987) Determination of Henry's Law Constants of Selected Priority Pollutants. Office of Science and Development. USEPA Report 600/D-87/229.

Washington State Department of Ecology. (2000) Occurrence of methyl tertiary butyl ether (MTBE) in groundwater at leaking underground storage tank sites in Washington. Publication No. 00-09-054.

White, R. (1985) The influence of macropores on the transport of dissolved and suspended matter through soil. *Advances in Soil Science*, Vol. 3. Springer Verlag, New York, 95–113.

Wickland, D., Mailloux, M., and Katyal, A. (2000) Application of field and laboratory analysis to modeling light non-aqueous phase liquid volume and recoverability at a Texas refinery. *Conference on The Petroleum Hydrocarbons and Organic Chemicals in Ground Water—Prevention, Detection and Remediation*, November 15–17, 2000. Anaheim, CA.

Wilson, J., Enfield, T., Dunlop, W., Cosby, R., Foster, D., and Baskin, L. (1981) Transport and fate of selected organic pollutants in a sandy soil. *Journal of Environmental Quality* 10, 501–506.

Wilson, J. and Linderfelt, W. (1991) Groundwater quality in pumping wells located near surface water bodies. Tech. Comp. Rpt. 261. New Mexico Water Resources Institute. Las Cruses, NM.

Wilson, J. and Liu, J. (1995) Backward tracking to find the source of pollution. In *Waste-management: From Risk to Reduction*, R. Bahda, *et al.*, eds. EDM Press, Albuquerque, NM, 181–199.

Wilson, J., Cho, J., Wilson, B., and Vardy, J. (2000) US Coast Guard fuel farm Elizabeth City, North Carolina. *Soil Sediment and Groundwater*, MTBE Special Issue, 54–56.

Wilson, J., Kolhatkar, R., Kuder, T., and Daugherty, S. (2005) Stable isotope analysis of MTBE to evaluate the source of TBA in ground water. *Ground Water Monitoring & Remediation* 25(4), 108–116.

Woodbury, A. and Ulrych, T. (1996) Minimum relative entropy inversion: Theory and application to recovering the release history of a groundwater contaminant. *Water Resources Research* 32(9), 2671–2681.

Woodbury, A., Sudicky, E., Ulrych, T., and Ludwig, R. (1998) Three-dimensional plume source reconstruction using minimum relative entropy inversion. *Journal of Contaminant Hydrology* 32(1–2), 131–158.

Wright, D., Abraham, J., Ellefsen, K., and Rossabi, J. (1998) Bore hole radar tomography and dielectric logging at the Savannah River site. *Proceedings of the Seventh International Conference on Ground Penetrating Radar*. May 27–30, 1998. University of Kansas, Lawrence, KS.

Zeng, C. and Bennett, G. (1995) *Applied Contaminant Transport Modeling. Theory and Practice*. Van Nostrand Reinhold, New York, NY.

CHAPTER 12

FORENSIC AIR DISPERSION MODELING AND ANALYSIS

Bruce A. Egan and Brian L. Murphy

12.1 Introduction 578
12.2 Deposition Processes 580
12.3 Estimating Soil Concentrations 581
12.3.1 Source Emission Parameters 581
12.3.2 The Influence of Meteorology and Terrain 582
12.3.2.1 Sources of Meteorological Data 586
12.3.2.2 The Influence of Nearby Terrain and Water Bodies 587
12.3.2.3 Sources of Terrain Elevation and Surface Characteristics Data 588
12.3.3 Plume Rise 588
12.3.4 Atmospheric Dispersion Modeling 589
12.3.4.1 The Gaussian Plume Equation 590
12.3.4.2 Gaussian Plume-based Models 594
12.3.4.3 Numerical Simulation Models 596
12.3.5 Determining Soil Concentrations 597
12.3.6 Soil Contamination Patterns 597
12.3.7 Concentration Ratios 598
12.3.8 Case Study 599
12.4 Redistribution Subsequent to Deposition 602
12.4.1 Effect of Soil Permeability 602
12.4.2 Drip Line Enhancement 602
12.4.3 Sediment Enrichment 603
12.4.4 Resuspension 604
12.5 Accidental Releases of Hazardous Substances 604
12.5.1 Heavy or Dense Gases 605
12.5.2 Estimation of Release Rates 606
12.6 Summary 607
Acknowledgment 608
References 608

12.1 INTRODUCTION

Releases of chemical contaminants to the atmosphere from even a small source can disperse over a wide area and potentially result in toxic exposures for a large number of people. For this reason toxic torts involving large numbers of plaintiffs often are based on the putative harm done by an air emission source. In principle, exposures can occur via direct inhalation of ambient air, contact with soils affected by the deposition of contaminants, consumption of water from water supplies contaminated by deposition processes, or indirectly, for example, via ingestion of vegetables grown in contaminated soils. Inhalation exposures generally occur during the time that the source is operating; however, exposures to soils, including inhalation exposures to fugitive dusts, can potentially occur long after a source has ceased operation. Accidental releases of substantial quantities of toxic chemicals to the atmosphere can result in large short-term exposures. Releases that go onto the soil, into containment structures such as dikes or into water bodies can also result in rereleases into the atmosphere.

In this chapter we will first address issues associated with routine emissions from conventional sources such as power plants, smelters, refineries, chemical plants, and other industries that would use chimneys or vents for the discharge of air emissions. The amount of the substances released into the atmosphere for these routine releases is generally not sufficient to alter the fluid dynamics or thermodynamics of the air flow field. Later, we will discuss the special issues associated with short-term accidental releases of significant quantities of toxic substances that may behave as heavier than air gases and consequently may alter the air flow fields that disperse the substances.

For air quality sources, forensic lines of inquiry generally involve a combination of environmental measurements and air dispersion modeling.

Direct measurements of ambient air quality concentrations are often available from state, federal, or local observational networks. However they are limited by the time periods and duration of the measurement programs and are limited spatially by the finite number of sampling locations. Interpreting air monitoring data in terms of currently operating sources is the subject of Chapter 8.

Measurements of the concentrations of contaminants in the soil can provide information of the cumulative effects of atmospheric deposition of emissions from historical sources of many compounds. For soil contamination, the source in question may not be responsible for the entirety of the observed measurements of contamination. Natural background, as well as emissions from other nearby or regional sources, may contribute to observed soil concentrations. When contaminants in soil are alleged to originate from a specific

air emission source, it is often possible to determine the relative contribution of this source by comparing observed soil concentrations to concentrations estimated using dispersion modeling.

Air dispersion modeling uses meteorological records and emissions rates as input information to mathematical models that calculate or simulate the effects of atmospheric motions in transporting, diluting, and dispersing pollutants. Concentrations can be predicted at as many receptor locations as entered into the model. The US EPA and state agencies heavily rely on dispersion modeling to quantify expected ambient air concentrations for purposes of permitting new or modified industrial emission sources. As a result, over the past three decades, there have been major research efforts focused on improving the predictive capabilities of atmospheric dispersion modeling methods.

For a source that is no longer operating or even no longer in existence, soil contamination by long-lived or recalcitrant compounds may be the only physical evidence of the source's impact while it was operating. The next sections of this chapter describe how soil concentrations and spatial distribution patterns are used to determine contributions from air emission sources. To the extent that soil sampling and data analyses provide evidence regarding the emission characteristics of a past source, these analyses can be used to estimate inhalation exposures that occurred when the source was operating.

In Section 12.2, wet and dry deposition processes are described. In Section 12.3, information required to estimate soil concentrations by air dispersion modeling is discussed. Often, historical source emission parameters needed for dispersion modeling are unknown or unavailable and must be estimated. Techniques for estimating source emission parameters are included in Section 12.3.1. The influence of meteorology and terrain is discussed in Section 12.3.2. Sources of meteorological and terrain data are described in Sections 12.3.2.1 and 12.3.2.2. Plume rise is discussed in Section 12.3.3.

Screening and refined atmospheric dispersion modeling techniques that estimate downwind concentrations based on source and meteorological variables are described in Section 12.3.4. The models range in complexity from the simple Gaussian plume equation to present state of the art methods. Both absolute soil concentrations and soil concentration patterns can be used to test whether or not a particular source is responsible. The two techniques are discussed in Sections 12.3.5 and 12.3.6. In Section 12.4 redistribution processes that can alter soil contamination patterns subsequent to deposition are described. Section 12.5 contains a discussion of using ratios of soil concentrations for different species to identify sources. A discussion of the special modeling needed to address the consequences of accidental chemical spills is included in Section 12.6.

Section 12.3.8 presents a case study illustrating the use of the various techniques described earlier.

12.2 DEPOSITION PROCESSES

Contaminants emitted to the atmosphere from a stack or other sources are eventually removed from the atmosphere via deposition or by chemical reaction. Over a sufficient period of time, emissions of contaminants that are long-lived in the environment can result in measurable soil concentrations above background. Long-lived contaminants include metals, many radio nuclides, dioxins, and other persistent organic pollutants. In addition to being emitted as particles, these substances may be emitted as vapors, attach themselves to coemitted particles or preexisting particles in the atmosphere, or coalesce to form new particles in the plume.

The air dispersion and deposition of these contaminants is determined by the behavior of the particles to which they are attached. Most combustion-generated particles from sources with particulate control equipment are less than a few microns (μm) in diameter. In this size range, gravitational settling is unimportant and the particles are affected mainly by winds and air turbulence. Some models use a sophisticated algorithm, incorporating surface roughness, to estimate a deposition rate. With this algorithm, particles less than about 0.2 mm have increasing deposition velocities because surface effects predominate over gravity in this size range.

However, there may be some soot and fly ash particles from uncontrolled sources that are as large as 50 mm or more in diameter (EPA, 1995). These larger particles will be affected by gravity and will be deposited closer to the source than the small particles. Typical deposition velocities are, as given by Schroeder *et al.* (1987), about 1 cm/s for particles between 1 and 10 mm in diameter and between 0.2 and 1.0 cm/s for smaller particles between 0.1 and 1.0 mm in diameter. Both gases and particles are deposited on the ground due to wet or dry deposition. In wet deposition contaminants are scavenged during precipitation events. In dry deposition small particles, less than about 10 mm in diameter, behave like gases and are brought to ground primarily by atmospheric turbulence. For larger particles gravitational settling can be a factor. Wet deposition can be significant in terms of total removal of contaminants over long distances; however, dry deposition is more important over a many year period in determining the deposition pattern near a source. If an air emission source is the primary cause of soil contamination, then the spatial distribution of contaminants in the soil should be determined, among other things, by the dry deposition pattern of the air emission source.

12.3 ESTIMATING SOIL CONCENTRATIONS

This section begins by discussing how to estimate source operating characteristics when, as for a historical source, these characteristics may be poorly known. These source characteristics are necessary together with meteorological data for dispersion modeling. As discussed later, dispersion modeling can be used to estimate ambient air concentrations of pollutants and the deposition of pollutants to ground surfaces for purposes of predicting soil concentration values and patterns.

12.3.1 SOURCE EMISSION PARAMETERS

In order to predict soil concentrations for comparison with observations, source emission characteristics must be known or be able to be estimated. Source emission characteristics include not only the emission rate for the chemical of concern, but for stack releases, also parameters such as exit velocity, temperature of the exhaust gases, and physical dimensions of the stack (i.e., inside stack diameter and stack height). This information is necessary to estimate the initial trajectory of the exhaust gases after exiting the stack, which depends upon the buoyancy and momentum of the exhaust gases. This information is often available from state environmental agency emissions inventories. When the source no longer exists and the emission parameters are poorly known it still may be possible to estimate the necessary parameters. For example, for combustion sources, fuel purchase or operating records should exist. For such a source, if complete combustion is assumed, stoichiometry dictates that each carbon atom in the fuel is used to generate one molecule of carbon dioxide (CO_2) and every two hydrogen atoms in the fuel are used to produce one molecule of water (H_2O). The amount of air required for complete combustion can then be derived from these oxygen (O_2) requirements knowing that oxygen comprises about 21% of air by volume. Calculations usually are performed assuming that an excess amount of air, typically about 40%, is actually used for combustion. The sum of air used in combustion plus excess air provides an estimate of stack gas volume as a function of fuel use.

Once the amount of stack gas from combustion is known, the gas temperature can be estimated from the actual, or nominal, heating values for the fuel used. The heated stack gas volume can then be estimated using this temperature and the ideal gas law. If the stack diameter is known, the stack gas exit velocity can be calculated from the volume generation rate. In this way the parameters necessary to calculate plume rise for dispersion modeling can be calculated from the fuel use rate.

Hazardous wastes, such as spent solvents, may be burned with fossil fuels. If the chemical composition of the hazardous waste is known or estimated, the volume of stack gas generated from the combustion of the hazardous waste can be estimated. Emissions of metals can be estimated simply by knowing the composition of the waste. Concentrations of combustion products such as dioxins or products of incomplete combustion in the stack gases are estimated based on the volume of stack gas evolved from the waste and on stack gas concentration measurements in similar facilities burning similar wastes. However, deciding what is a similar facility can be quite involved.

12.3.2 THE INFLUENCE OF METEOROLOGY AND TERRAIN

The meteorological variables that affect the dispersion of pollutants from a source and therefore determine ambient air concentrations and deposition processes to ground surfaces are the wind speeds and directions that are associated with the downwind transport of the pollutants, the wind and temperature gradients in the vertical direction, and the atmospheric turbulence mixing rates. These meteorological parameters vary with time and location as well as with height above the ground surface. They are also significantly affected by local topographic features. The estimation of the values of these variables to calculate the dispersion of contaminants depends, in turn, on a number of physical and thermodynamic processes within the atmospheric boundary layer flow that affect the variation of wind and temperature with height above the surface. These include solar radiation heat flux to the surface during the day, upward heat loss by radiation from the surface at night (which suppresses the generation of thermal eddies), surface roughness and its effect on creating mechanically induced turbulence within the surface layer, and the release or absorption of latent heat associated with humidity and the surface moisture content. The variation of the ambient temperature, humidity, and wind speed and directions with height above the surface are described by associated profiles. The field of planetary boundary layer meteorology addresses how these processes are parameterized for dispersion modeling calculations. Hanna *et al.* (1982), Randerson (1984), and the documentation of the ISC3 and AERMOD models (EPA, 1995b; Cimorelli, 2002) are good references to this subject matter.

A discussion of the relationship between meteorological conditions and dispersion rates used in modeling follows. Emphasis is placed upon issues that are often in contention in legal disputes.

Meteorological measurements typically are taken at one or more levels from a meteorological tower located away from buildings and structures that might alter the air flow. The 2-meter and 10-meter heights are commonly used as

reference levels for determining gradients of the wind and temperature. Higher levels of 30 to more than 100 meters are used in some settings to better characterize pollution transport, especially in complex terrain settings. However, wind data from a 10-meter level at a nearby U.S. National Weather Service (NWS) airport station is commonly used for dispersion modeling for sources in simple terrain. Continuous measurements are sampled rapidly and the digitized data is processed to provide short-term wind fluctuation data used to estimate turbulent dispersion rates and also to provide one-hour average wind, temperature, and humidity measurements.

Wind data are conveniently summarized in the form of a wind rose, which according to convention shows the direction that the wind is coming from. For example, surface wind data collected at the Dallas/Fort Worth Airport between January 1986 and December 1990 are summarized in Figure 12.1. Note that the "petals" of the "rose" corresponding to the compass directions also indicate wind speed information. As indicated in Figure 12.1, at this station surface winds originate predominantly from the south (about 20% of the time), or south-southeast (12%), followed by winds from the south-southwest (9%) and north (9%), and infrequently from the east or west. Consequently, any air emission source should produce a pattern of annual average concentration and soil contamination reflecting this strong north/south signature. In particular, the highest soil concentrations should be observed north of the emission source. Wind speed has a primary influence in diluting the emissions from a stationary source. Concentrations are roughly inversely proportional to the value of the wind speed. Because wind speeds generally increase with height above the ground, it is important to take that factor into consideration in analyses.

Mean wind speeds in most United States locations are 6–11 miles (9.6–17.7 km) per hour (Bair, 1992). Average daytime wind speeds are higher than nighttime wind speeds.

Atmospheric turbulence is responsible for the mixing and dilution of contaminants that are carried by the wind. Turbulence mixing rates commonly are associated with atmospheric stability conditions that primarily depend upon how a parcel of air would behave in different vertical temperature profiles when vertiucally displaced. Stability conditions vary by location and rapidly by time of day and with height above the ground. The determination of stability is therefore often an important finding in forensic analyses, especially for short-term exposures. We will therefore review this issue in the next paragraphs.

The pressure of the atmosphere decreases with height above the surface. Temperature may or may not decrease with height above the ground, and this becomes a very important factor in the determination of dispersion rates.

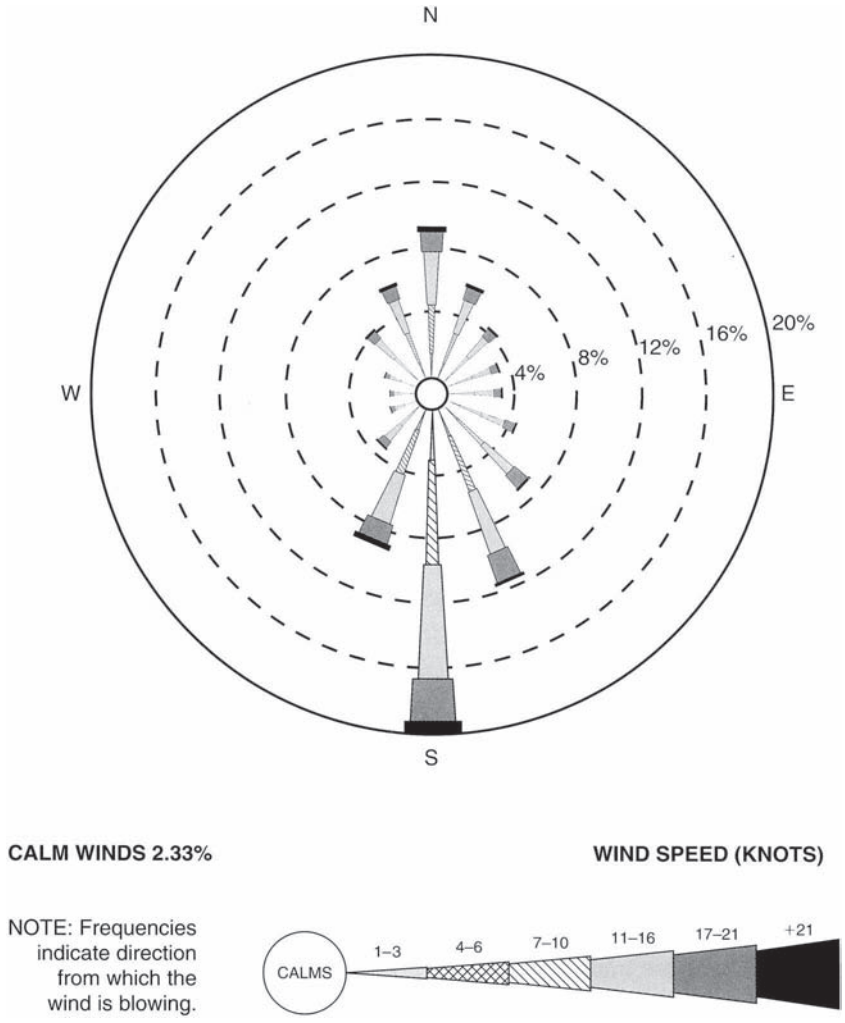


Figure 12.1
Dallas/Fort Worth meteorological data 1981–1984.

Consider a parcel of air that is moved vertically. The pressure surrounding the parcel will be less and the parcel will expand slightly, reducing its density. If no heat is added or subtracted by this motion (i.e., an adiabatic process), this expansion will also cause the parcel of air to cool, tending to increase the density. Whether the movement causes the parcel to become more or less dense than the surrounding air, and therefore whether the parcel wants to move further upward or downward depends upon the density gradient of the surrounding air. This density gradient depends upon the variation of temperature in the vertical direction.

The atmosphere is considered *stable* if the temperature increases with height and a parcel moved upward finds itself denser than its surroundings. The parcel would then respond by moving back downward toward its previous

position. The opposite would happen if the parcel were initially moved downward. Thus a stably stratified atmosphere will suppress vertical motions associated with turbulence. Stable conditions near the surface typically occur at night when radiational cooling of the ground results in the air temperature increasing with height above the ground. This is called a *temperature inversion*.

The atmosphere is considered *neutral* if the temperature gradient is such that the change of density associated with the temperature change with height is the same as the change of density associated with the change of pressure with height. For dry air, this equilibrium temperature gradient is about minus 1.0 degree Centigrade per 100 meters of height. This is also called the adiabatic lapse rate. A parcel moved upward or downward in a neutrally stable atmosphere would find itself surrounded by air of the same density and not experience any net upward or downward buoyancy force. Neutral stability conditions commonly occur with moderate and strong wind speeds and can occur at any time of day.

The atmosphere is considered *unstable* if the temperature gradient is such that a parcel of air moved upward finds itself less dense than its surroundings. The parcel will then tend to accelerate further upward. Again the opposite would happen if the parcel were initially moved downward. Thus an unstably stratified atmosphere enhances vertical turbulent motions. The most common cause of an unstable boundary layer is solar heating of the ground surface where buoyancy forces cause air warmed at the surface air to vigorously rise and mix with air at higher elevations.

The potential temperature gradient is used to define these conditions. It is given by

$$\partial\theta/\partial z = \partial T/\partial z + g/C_P \tag{12.1}$$

where $\partial T/\partial z$ is the vertical temperature gradient (k^0/m), g is the acceleration of gravity (m/s^2), and C_P is the specific heat of air at constant pressure (10^7ergs/gm – k^0). (Hanna *et al.,* 1982).

- If $\partial\theta/\partial z$ is >0, the atmospheric layer is stably stratified
- If $\partial\theta/\partial z$ is 0, the atmospheric layer is neutrally stratified
- If $\partial\theta/\partial z$ < 0, the atmospheric layer is unstable

There are a number of ways to refine the categorization of surface stability conditions. The most common, after Pasquill and Gifford and described in Hanna *et al.* (1982), defines six stability classifications: A (Extremely unstable) through F (Moderately stable). Class D is for neutral conditions.

The classifications are defined by ranges of observations of surface wind speed and solar insolation during daytime hours, or surface wind speed and cloud cover during nighttime hours. Other means used in advanced dispersion models involve ratios of the thermal heating effects and mechanically induced mixing. These involve direct measurements of additional factors of wind shear and solar and net radiation.

The vertical temperature structure is also important in determining turbulence rates well above the ground surface. For example, an elevated inversion can occur as a result of subsidence of air in a high pressure system. This inversion inhibits the vertical turbulent motions generated within the surface layer at the top of the surface layer defining what is known as the mixing depth or height.

We will return to describing the parameterization of winds and turbulence in Section 12.3.4 when we describe dispersion models more specifically.

12.3.2.1 Sources of Meteorological Data

Meteorological data can be obtained from a number of government and private resources. The National Climatic Data Center (NCDC) in Ashville, North Carolina, is a repository for meteorological data collected from a variety of sources throughout the world. Other sources in the United States include the National Weather Service (NWS), the Federal Aviation Administration (FAA), military bases, universities, agricultural sites, and some industrial sites. The USEPA also supplies a limited number of NWS meteorological observations suitable for use in air dispersion modeling through the Support Center for Regulatory Air Modeling (SCRAM). These can be accessed through the EPA web site. Meteorological data obtained directly from the NCDC or USEPA requires processing before they can be used in standard air models. The required meteorological processors generally are provided along with the actual air dispersion model. Private vendors can also provide meteorological data, typically retrieved from NCDC and processed for use in the air dispersion model of interest. When selecting meteorological data for use in air modeling, the quality assurance measures associated with the collection and processing of the raw data must be considered. Data collected by the NWS are considered to be of high quality.

The geographic locations of meteorological data should be selected on the basis of their ability to characterize the transport and dispersion conditions in the area of interest. Thus the selected data should be representative of the site being modeled in terms of spatial and temporal considerations. Spatial considerations include both horizontal and vertical directions. They would include a close proximity of the modeled area to the site where the meteorological data

were collected as well as similarities of such things as distance to large water bodies and wind-channeling-terrain features. From a temporal basis the selected meteorological data should be from a period of time as close to the period being modeled as possible. If a precise temporal match cannot be established between the available meteorological data and the period of interest for the modeling exercise, then the selected record should be of sufficient duration to capture all meteorological conditions expected at the site being modeled. In general, when modeling a long-term release, the USEPA (2005) considers a five-year period of meteorological data from an off-site location to be sufficient to reduce uncertainty in the model predictions to an acceptable level for regulatory purposes. Many sources elect to capture site-specific data to assure that the meteorological data would be considered representative of the site. In such cases, one year of data is generally considered adequate.

Upper air data is necessary to estimate the altitude of the mixing height below which vigorous dispersion can occur. Upper air data typically is measured twice a day but only at specific NWS stations. The data can be obtained from some of the same sources as identified earlier for surface data.

Precipitation data, for use in wet deposition modeling, must be obtained from NCDC or in some cases private vendors. Precipitation data is not on the EPA's SCRAM web site.

The processing necessary to use these data depends on the variable and the model. Much of the meteorological data can be processed using the EPA's PCRAMMET program. The program and user's guide can be obtained from the SCRAM web site. PCRAMMET will produce model ready meteorological input files including for dry and wet deposition modeling. Some models require the use of dedicated meteorological preprocessors.

12.3.2.2 The Influence of Nearby Terrain and Water Bodies

The presence of large terrain features and nearby water bodies is an additional factor in analyses of air flow patterns. In the presence of strong large-scale (synoptic) meteorological flows (such as associated with migratory high and low pressure areas), blockage and distortion effects of terrain features will alter the wind speed and direction fields and the thermal structure of the air flow within the planetary boundary layer. All these factors significantly affect dispersion rates. Under less strong large-scale meteorological flows, differential heating can cause drainage flows at night or alternatively upslope flow during the day. Detailed discussion of the topic of dispersion in complex terrain may be found in Egan and Snyder (2002).

Similar alterations of the flow and temperature field occur with the presence of nearby oceans or lakes. Here under light synoptic scale weather

conditions, local onshore flows often occur during the day and to a lesser degree, offshore winds at night. Some of the models discussed in Section 12.3.4 may be applied to assess pollutant transport in these complicated flow conditions. Models that address complicated flows generally require more detailed input information on topographic elevations and the characteristics of the surface in terms of ground cover (grass, forest, water, etc.) and their associated physical characteristics (albedo, surface roughness, moisture content).

12.3.2.3 Sources of Terrain Elevation and Surface Characteristics Data

Terrain elevation data are necessary when the typical plume height (stack height plus plume rise) is near or below the terrain height. Terrain data typically are obtained from U.S. Geological Survey (USGS), private vendors, and in some limited cases from state geographic information system (GIS) agencies. Data obtained from the USGS can be in electronic or paper (i.e., topographic maps) format. AERMOD has a special preprocessor, AERMAP.

Surface characteristics data can generally be obtained from state agency GIS databases. This information may need to be processed to produce the specific input information needed for running dispersion models.

12.3.3 PLUME RISE

For buoyant sources or for sources with high exit velocities, it is necessary to estimate the plume rise of the exhaust gases after exiting the stack. Dispersion models routinely used for regulatory permitting purposes include algorithms to estimate plume rise. The complexity of these algorithms varies by model depending primarily on whether the effects of nearby structures on the flow patterns near the stack are considered and upon how detailed the wind and temperature profiles are depicted in the model. These dispersion models will be discussed in more detail in Section 12.3.4. For purposes of providing estimates for emissions from a single stack not affected by nearby structures, the following formulae are commonly used (EPA, 1995b).

Under neutral or unstable conditions, plume rise Δh above the stack height due to buoyancy is given by:

$$\Delta h = 21.425F^{3/4}/u, \quad F < 55 \tag{12.2}$$

$$\Delta h = 38.71F^{3/5}/u, \quad F > 55 \tag{12.3}$$

The wind speed at the top of the stack is u and F is the buoyancy flux, given by:

$$F = g\left(1 - \frac{T_0}{T_P}\right) wR^2, \tag{12.4}$$

where g is the acceleration of gravity, T_P is the exhaust gas temperature (K^0), T_0 is the ambient temperature (K^0), w is the stack exit velocity, and R is the stack radius, all in MKS units.

For plume rise due to momentum under unstable or neutral conditions:

$$\Delta h = 6Rw/u. \tag{12.5}$$

The larger Δh from Equation 12.2 or 12.3 and Equation 12.5 should be used. For combustion sources Equation 12.2 or 12.3 will generally yield the larger Δh. Large plume heights are associated with tall stacks, high temperature exhaust gas, and large exhaust gas emission rates.

Note that in these equations the plume rise is inversely dependent on the value of the wind speed and is therefore very sensitive to the values assumed, especially when wind speeds are low. The formulae are generally not considered reliable for wind speeds less than about a meter per second.

For a buoyant plume during stable atmospheric conditions,

$$\Delta h = 2.6(F/us)^{1/3} \tag{12.6}$$

For a momentum dominated plume during stable atmospheric conditions,

$$\Delta h = 1.5(F_m/u\sqrt{s})^{1/3} \tag{12.7}$$

where F_m is the momentum flux defined by

$$F_m = w^2R^2T_0/T_P \tag{12.8}$$

where s is a stability parameter given by

$$s = g(\partial\theta/\partial z)/T_0 \tag{12.9}$$

and $(\partial\theta/\partial z)$ is the potential temperature gradient (degrees C/m) defined in Equation 12.1.

12.3.4 ATMOSPHERIC DISPERSION MODELING

Atmospheric dispersion modeling provides the means to relate emissions of pollutants to the atmosphere to expected ambient air concentrations of those

pollutants. The degree of detail of atmospheric motions and of chemical and physical transformation processes differs greatly among models depending upon the intended application. Also the methods for calculating the ambient air concentrations can be as simple as an empirical equation at one extreme to computational fluid dynamic (CFD) simulations that attempt to simulate the wide range of eddies and flow details associated with fluid mechanic and thermodynamic phenomena at the other extreme. The most commonly used models are between those extremes. Models also are characterized as screening or refined. Screening models are relatively simple to use and made purposely conservative in the sense that they overestimate concentration values. For some screening models the conservatism is associated with the fact that they rely upon preestablished sets or combinations of meteorological conditions. These models do not require any input of actual meteorological measurements. Other screening models require representative meteorological data and are conservative by virtue of the assumptions made about plume trajectories or behavior as a plume approaches a ground surface. This is especially the case in regions of complex terrain.

Refined models always require the use of representative meteorological data and generally require detailed information on terrain heights and features. Refined models are expected to predict more realistic and less conservative ambient air concentrations. There is therefore a tradeoff of conservatism and convenience versus accuracy and complexity that enters discussions of dispersion modeling techniques.

12.3.4.1 The Gaussian Plume Equation

We start by describing the simple Gaussian plume equation that has been a core component of dispersion modeling for decades. The model as commonly used assumes steady state meteorological conditions over a time period of an hour or sometimes less. In its simplest form, the meteorological information required is the wind speed and direction and the atmospheric turbulence or diffusion rates.

According to the Gaussian dispersion equation, downwind concentration c (x, y, and z) is given by:

$$c = \frac{Q}{2\pi u \sigma_y \sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2}\right)\left[\exp\left(-\frac{(h-z)^2}{2\sigma_z^2}\right)+\left[\exp\left(-\frac{(h+z)^2}{2\sigma_z^2}\right)\right]\right] \quad (12.10)$$

where Q is the mass per unit time emitted by the source, z is the height above the ground, y is the cross wind horizontal distance from the plume centerline, σ_y is the horizontal dispersion parameter, σ_z is the vertical dispersion parameter, h is the plume centerline height, and u is the wind speed at that height.

The term involving y represents a horizontal Gaussian profile centered at the plume centerline, $y = 0$. The two terms involving z arise because the plume is assumed to be reflected at the ground, $z = 0$. The first term involving z describes a Gaussian vertical profile with a peak at the plume centerline, $z = h$. The second term involving z describes a Gaussian vertical profile centered at $-h$. In other words the reflected portion of the plume is treated as if it originated from a source a distance $-h$ below the ground surface. This is known as an image source. The second term involving z often is preceded by a factor, f, so that if a fraction f of the plume striking the ground is reflected, a fraction $1 - f$ is deposited.

The dispersion parameters σ_y and σ_z represent the effects of atmospheric turbulence and need to be specified as a function of downwind distance, x. In general, σ_z is smaller than σ_y because the presence of the earth's surface suppresses turbulence in the vertical direction. The plume cross section in a plane perpendicular to the centerline thus appears as an ellipse.

Equation 12.10 conserves mass. That is, the mass per unit time emitted by the source is equal to the mass flux through a vertical plane downwind.

$$Q = u\int_0^\infty dz\int_{-\infty}^\infty dyc(y,z). \tag{12.11}$$

The highest concentrations occur along the plume centerline, $y = 0$. At ground level where deposition occurs, $z = 0$. In this case Equation 12.10 simplifies to:

$$c = \frac{1}{\pi\sigma_y\sigma_z u}\exp -\frac{h^2}{2\sigma_z^2} \tag{12.12}$$

Note that the wind speed, u, is assumed to be a constant. It should represent the value at the plume height.

Maximum Downwind Concentration

The simplicity of the Gaussian equation allows the derivation of simple relationships between the concentrations and the input information. To illustrate, assume that the variation of the dispersion parameters in the bent over plume is $\sigma_y = \alpha x^a$ and $\sigma_z = \beta x^b$, where x is the downwind distance. Differentiating the concentration, c, in Equation 12.12 with respect to x and setting the result equal to zero gives:

$$\sigma_z = h\sqrt{\frac{b}{a+b}} \tag{12.13}$$

showing that there is a single maximum value of c, which occurs at a distance where the vertical dispersion parameter is somewhat smaller than the plume height.

The maximum value of c at that location is given by

$$c = \frac{\alpha^{-1}\beta^{\frac{a}{b}}}{\pi u h^2}\left(h\sqrt{\frac{b}{a+b}}\right)^{\frac{-(a+b)}{b}} \exp[-\{(a+b)/2b\}] \tag{12.14}$$

showing that for $a = b = 1$, the maximum concentration is inversely proportional to the square of the plume height.

The vertical and horizontal dispersion parameters used in the Gaussian equation are empirically based upon field experiments performed in different locations. Dispersion models rely upon curve fitting techniques to develop algorithms to calculate the dispersion rates as a function of downwind distance. As plumes grow in size, the influence of larger eddies in the atmosphere in causing dispersion increases. For a plume with a small cross section, large eddies can be thought of as simply causing plume meander and not diluting the plume centerline concentrations. Thus the dispersion coefficients depend on the spectrum of turbulence in the atmosphere, which in turn depends on the contributions from mechanical mixing by friction from the wind blowing over obstructions to flow at the earth's surface or by the thermal effects associated buoyant eddies due to solar heating. The most commonly used dispersion coefficients to be used in the Gaussian equation are the Pasquill–Gifford values for rural conditions or those derived by Briggs for urban conditions. These can be found in graphical or equation form in dispersion modeling texts such as Turner (1994) or Hanna *et al.* (1982).

Figure 12.2 from Turner (1994) shows the distance to the downwind maximum along the abscissa or horizontal axis as a function of stability and plume height. The plume height or effective height of release in Figure 12.2 ranges from 2 meters to 500 meters.

Because different particle size ranges have different deposition velocities, the maximum values for particulate emissions will be somewhat closer to the source. For gaseous contaminants and uniform ground conditions the maximum deposition is expected to occur at the point of maximum ground level air concentration as shown in Figure 12.2.

Figure 12.2, which applies to flat terrain, does not imply that the same amount of deposition occurs at the maximum under different conditions. For example, all other things being equal, the amount of deposition will decrease at the maximum with increasing plume height. The air concentration at the maximum can be estimated from the ordinate or vertical scale of Figure 12.2.

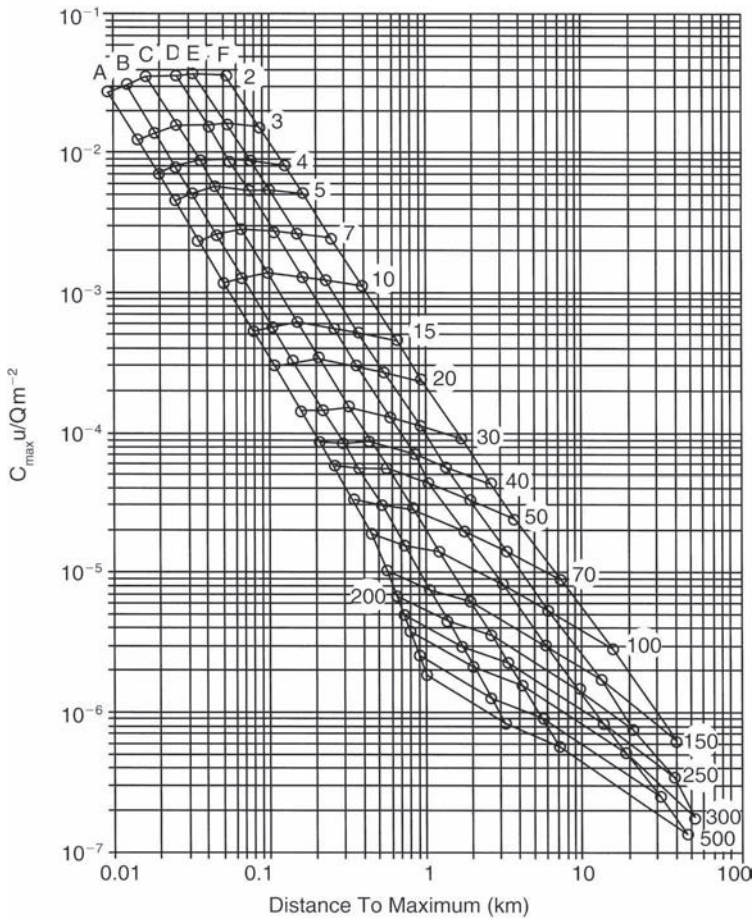


Figure 12.2
Relative maximum concentration normalized for wind speed and emissions vs. distance to maximum concentration as functions of Pasquill stability class and effective height of release. (Reprinted with permission from Workbook of Atmospheric Dispersion Estimates, 2e, *1994, D. Bruce Turner. Copyright CRC Press, Boca Raton, Florida.)*

The ordinate scale m^{-2} gives the air concentration (grams per cubic meter) at the maximum multiplied by the wind speed (meters per second) and divided by the source emission rate (grams per second).

There is considerable imprecision in estimating the distance to a downwind maximum using Figure 12.2. A better estimate is obtained by dispersion modeling using hourly meteorological data for the appropriate time period together with source operating parameters. However, if a soil deposition pattern is inconsistent with the overall pattern of (a) highest concentrations in the prevailing wind direction, and (b) a maximum occurring at approximately the distance indicated by Figure 12.2, then the source identification is either incorrect or concentration patterns have been greatly rearranged subsequent to deposition. The latter possibility is considered further in Section 12.4.

12.3.4.2 Gaussian Plume-based Models

The Gaussian plume equation, with significant elaborations in parameterizations, forms the basis for several dispersion models used in regulatory practice, which are described in this section.

SCREEN3

This model is available from the EPA's SCRAM web site. It is a screening model designed to be conservative by overestimating maximum predicted ambient air concentrations.

Using only stack exit parameters and nearby structure dimensions as input, the model calculates the one-hour average maximum ground level concentrations for combinations of preset wind speed categories and stability classifications. The conditions resulting in the largest predictions are displayed. Many of the algorithms are the same as those used in ISC3, described next. A simple scaling procedure is used to relate the predicted one-hour averages to longer averaging times. This scaling procedure is responsible for a large degree of the conservatism in this model. Deposition processes are not included. SCREEN3 also computes maximum values associated with shoreline fumigation phenomena for offshore sources located near coastlines.

The Industrial Source Complex Model (ISC3)

This model (EPA, 1995b) has been used extensively by EPA for regulatory applications without many changes for well over a decade. It is based upon the Gaussian plume equation with plume dispersion algorithms categorized by the stability classification schemes identified earlier for rural and urban settings. In normal operation, the model requires that one or more years of meteorological data be used. The model uses that information sequentially to calculate one-hour average concentrations for each hour of data. Those calculations are then post-processed to generate concentrations for longer averaging times such as for daily or annual averages. The model includes building downwash algorithms and deposition algorithms. It assumes steady state meteorology and straight line plume trajectories for each hour of calculation. At the time of writing ISC3 is being replaced by AERMOD (described next) for many regulatory applications. However, the deposition algorithms for AERMOD have not been finalized. Modelers have an option of using ISC3 for deposition calculations until AERMOD is approved for such use.

The American Meteorological Society and EPA Regulatory Model Improvement Committee Model (AERMOD)

AERMOD was in development for over 12 years and has undergone extensive statistical performance testing by the EPA. AERMOD is a steady-state,

straight-line trajectory model based upon a much more detailed understanding of the physics of air flow within the planetary boundary layer. It was promulgated as a guideline model to replace ISC3 as a refined multipurpose model in November 2005 (EPA, 2005). The model is applicable to sources in complex terrain settings with terrain height both above and below the sources' stack heights. It also includes new algorithms (PRIME) for downwash effects simulation that have been shown to improve the accuracy under these conditions. AERMOD includes new algorithms for dealing with buoyant plumes that may partially penetrate elevated inversions. It uses numerical iterative techniques to calculate plume rise in the stable boundary layer.

AERMOD was developed to be run with different levels of sophistication of meteorological data. For example, it can use wind and temperature data from multiple levels of meteorological towers in an area of complex terrain. Alternatively, it can be run with a minimum of near surface meteorological data coupled with atmospheric sounding data routinely collected by the National Weather Service's network of upper air stations (Cimorelli *et al.*, 1998). Studies have shown (Paine, 2001) that the accuracy of AERMOD improves when more meteorological information is used as inputs.

A key improvement to AERMOD is the incorporation and use of three physical parameters that describe the ground surface conditions important to the rate of dispersion in the vertical direction. These are surface roughness, albedo, and the Bowen ratio, a ratio of sensible to latent heat content in moist air. These parameters are used together with wind fluctuation statistics, solar radiation, and mixing depths, to determine the rate of mechanical and thermal mixing in the planetary boundary layer. In AERMOD, these parameters are used in the construction of the vertical profiles of temperature and wind speed and turbulent diffusion rates.

The California Puff Model (CALPUFF)

CALPUFF (Scire *et al.*, 2000) is a variable trajectory model suitable for long-range transport modeling where meteorological conditions vary spatially and may change during the time period of transport. It may also be applicable to cases when meteorological conditions are unsteady, such as for periods of wind shifts. The model uses Gaussian-like dispersion rates, but for individual plume segments. It includes algorithms for building downwash, dry and wet deposition, and quasi-linear chemical reactions. Puff models can also model short-term releases if meteorological data for periods less than an hour is available. The meteorological inputs used by CALPUFF need to be produced by its meteorological preprocessor, CALMET.

The Complex Terrain Dispersion Model Plus Algorithms for Unstable Situations (CTDMPLUS)

CTDMPLUS (Perry *et al.*, 1989) is a Gaussian-based refined model for use in mountainous terrain settings. It was developed on the basis of a number of field studies and laboratory experiments and has been extensively validated for use where the terrain height exceeds that of elevated point sources. Multiple level meteorological data are required to create the input files.

CTSCREEN

The algorithms of CTDMPLUS form the basis for CTSCREEN (Perry *et al.*, 1990), a screening model for complex terrain. CTSCREEN uses a preestablished set of meteorological conditions. It does require the use of detailed topographic information on a fine scale.

The Offshore and Coastal Dispersion Model (OCD)

OCD (DiCristofaro and Hanna, 1989) is a straight-line Gaussian-based model for use in modeling shoreline fumigation phenomena where a high concentration in a stable layer due to an elevated source is brought rapidly to ground by an inversion breakup. This situation can occur in coastal locations when cooler air over water is warmed as it passes over land.

12.3.4.3 Numerical Simulation Models

Advances made in numerical simulation techniques and in computer speeds have encouraged the development of models capable of simulating the detailed air flows of both large and small physical scales on finite difference type computational grids. Numerical techniques routinely used in weather forecasting can be scaled down to generate meteorological fields with resolutions of a few kilometers or less. The Meteorological Model Version 5 (MM5) is an example of a meteorological driver model than can interface with CALPUFF or other puff-type models. Seaman *et al.* (2002) discuss the application of MM5 as a driver to different dispersion models simulating the effects of sea breezes on Cape Cod.

Computational Fluid Dynamics (CFD) models can apply to even smaller scale flows and are used to simulate fluid flows around objects and obstructions in minute detail. Although much of the early development and application of CFD models has been in the field of aerodynamics, present CFD models are capable of simulating fluid flows at both much smaller (e.g., microscopic) and larger (e.g., weather forecasting) scales. Recent work has shown CFD models to perform well against wind tunnel studies for assessing air quality issues associated with downwash and other recalculating flows around buildings. We expect CDF models to have larger roles in future forensic air quality analyses (Patil, 2005).

12.3.5 DETERMINING SOIL CONCENTRATIONS

A rough estimate of the expected soil concentration can be obtained by assuming that the predicted deposition over the years of source operation is mixed into the top few centimeters of the soil column. The predicted soil concentration is then equal to the product of the ambient air concentration, deposition velocity, and time of operation, divided by the mixing depth and bulk density of the soil as given by Equation 12.15.

$$C_s = 31.6\frac{C_a V_d t}{\rho x} \tag{12.15}$$

where C_s is predicted soil concentration (mg/kg); C_a is air concentration ($\mu g/m^3$); V_d is deposition velocity (cm/s); t is time of operation (years); ρ is density of soil (g/cm^3); and x is depth of mixing (cm).

The constant in Equation 12.15 is based on the assumption that the source operates 24 hours a day, every day of the year. Therefore it should be reduced proportionally for sources that are intermittent in operation. The air concentration, C_a, required in Equation 12.15 is determined by dispersion modeling. Most dispersion models also estimate the deposition velocity V_d given a particle size description. For comparison purposes, the depth of mixing should be taken as the depth of soil collected in the soil-sampling program. Then if predicted soil concentrations are much less than measured concentrations the source is not a credible explanation for the observations.

The modeled soil concentrations can be compared to observed concentrations on an absolute or relative basis. Comparison of absolute concentrations need not require highly accurate knowledge of the source emission characteristics. As suggested in the preceding paragraph, it may be possible to demonstrate that no reasonable set of emission assumptions can produce soil concentrations comparable to those that are observed. The observations therefore are due to background or other sources. When absolute concentrations cannot be compared reliably, comparison of concentration patterns is still possible.

12.3.6 SOIL CONTAMINATION PATTERNS

The deposition pattern predicted from the wind rose shown in Figure 12.1 for a particular source is illustrated in Figure 12.3. Contours of constant soil deposition are shown. The source is a small short-stack industrial boiler located in the southwest corner of the shaded region corresponding to the facility. Consistent with the wind rose in Figure 12.1, contaminants are

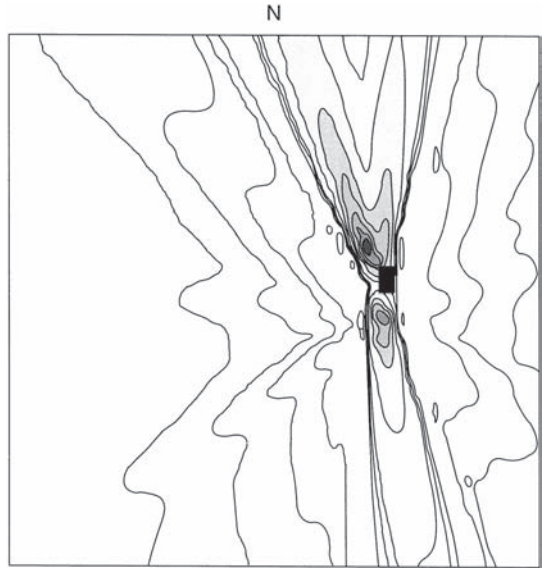


Figure 12.3
Deposition isopleths resulting from the Figure 12.1 meteorology.

deposited in a strong north–south pattern, with deposition occurring preferentially to the north. In this case the distance from the stack source to the maximum deposition location to the northwest of the facility is about 0.4 km. For a taller stack or a larger source with a more buoyant plume the distance to the point of maximum deposition would be greater, as discussed in Section 12.3.4.2.

In order to compare predicted and observed concentrations it is best to sample along a transect in the downwind prevailing wind direction in order to minimize interference due to background. For the same reason the soil sampling program should include the expected point of maximum deposition. Of course, if soil concentrations are indistinguishable from background this sets an upper limit on exposures that could previously have existed. The statistical methods described in Chapter 5 can be used to distinguish a source signature from background.

12.3.7 CONCENTRATION RATIOS

In some cases difficulties in determining sources due to redistribution processes can be obviated by examining ratios of different species. If a source is known to have emitted several different contaminants then, in principle, these can be used to find the source's signature in soils. For example, Kimbrough and Suffet (1995) used characteristic ratios of lead:antimony:arsenic in air, soil, and waste samples collected near a secondary lead smelter to identify the source of these elements in off-site soils. Characteristic ratios of these elements

were maintained even as absolute concentrations decreased with distance downwind of the plant. A similar analysis has been reported near a smelter in Denver (Christen, 2000). In this case cadmium concentrations in soil decreased with distance as expected for smelter emissions but arsenic concentrations did not. Instead arsenic concentrations varied with land use. These observations were partly responsible for the conclusion that arsenic concentrations were largely due to herbicide use.

Care is required in examining ratios of chemicals that degrade in the environment at different rates. For example, during gas phase transport, dioxins and furans are degraded by sunlight with dechlorination of more highly chlorinated isomers occurring more rapidly than that of less chlorinated isomers. This can result in similar dioxin and furan homologue ratios in soils even when emission profiles from different sources were originally distinct.

Care must also be taken in basing a ratio analysis on contaminants that occur naturally at high concentrations or on materials that are ubiquitous, such as gasoline products, phthalate plasticizers, or wood-treating products. Dioxins and furans are also widespread contaminants in the environment as discussed in the next section.

12.3.8 CASE STUDY

This case study is a toxic tort in which the plaintiffs claimed dioxin and furan exposure. Defendants were the operators of and suppliers to a furniture manufacturing operation not far from Dallas, Texas. Waste solvents, paints, and waxes were burned at the site both in a gas-fired boiler and in an open pit used for fire training exercises. Wastes were burned in the boiler on a weekly basis, whereas the burn pit was used only occasionally. Most of the plaintiffs were in an urbanized area to the west of the facility. The highest measured concentrations were in the sediments at one location in a creek to the southwest.

All dioxin measurements were converted to toxic equivalent quantities (TEQs), consistent with the EPA's, and all quantitative results reported below are in TEQ. At the creek hot spot, several samples were above 1000 parts per trillion (ppt). At other locations concentrations ranged from 2 to 50 ppt. However, only 2 out of 34 values were above 10 ppt. In the EPA's draft 2003 National Academy of Sciences, dioxin reassessment the North American rural and urban background concentrations were estimated as 2.6 and 8.8 ppt, respectively. Since the site is on the outskirts of a small city the question arises as to whether most of the values are simply representative of background.

The creek hot spot was not in the prevailing wind direction, which is from the south. In fact, according to Figure 12.1, winds blow from the northeast

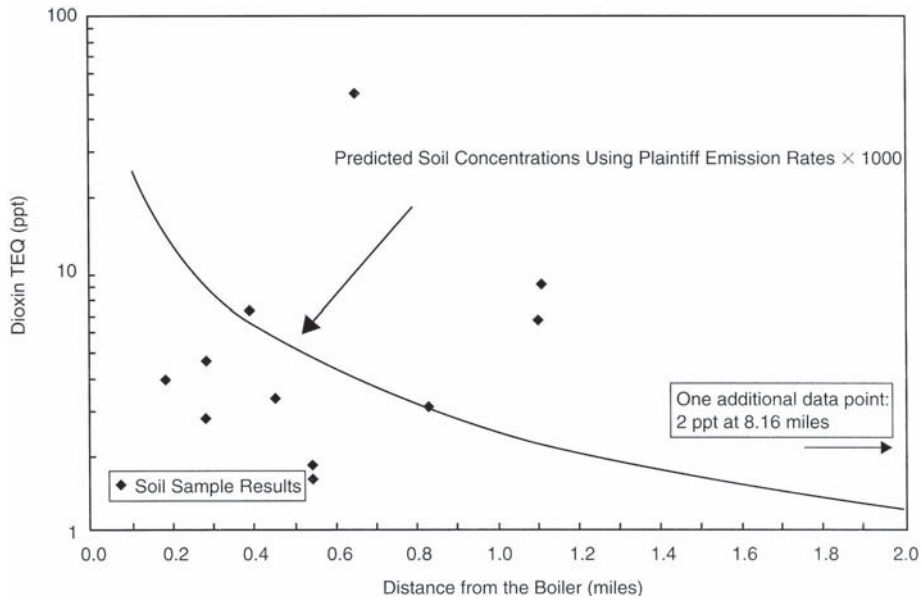


Figure 12.4
Dioxin TEQ (parts per trillion: ppt) vs. distance from the facility boiler for samples collected in the prevailing wind direction.

toward the hot spot only about 4% of the time compared with about 20% of the time from the south. The isopleths based on dispersion modeling shown in Figure 12.3 also indicate that the highest values would be found to the north of the facility. The hot spot location was wooded but permitted vehicle access. There was evidence it had been used for dumping. The location was also across the street from a hospital with a recently removed medical waste incinerator.

Since the furniture manufacturing facility boiler was small with a short stack, the peak air concentration for emissions from both the boiler and the burn pit were expected to occur quite close to the facility. A transect of soil measurements was taken going north of the plant in the prevailing wind direction. Results on a semilog plot are shown in Figure 12.4. Overall there is no indication that concentrations are falling off with distance from the putative source.

The highest concentration shown of nearly 50 ppt occurred at the base on the lee side of a hill where, unlike the other samples, there was a clay patch. The other measurement above 10 ppt, namely 24 ppt, occurred at a drainage ditch for runoff from the facility parking lot. Thus there is some evidence that at least the highest soil concentrations were determined largely by very local conditions.

Note that in Figure 12.4 the curve of predicted soil concentrations has been multiplied by 1000. We used the plaintiff expert's estimated emissions for the dispersion modeling but found that these had to be multiplied by about a

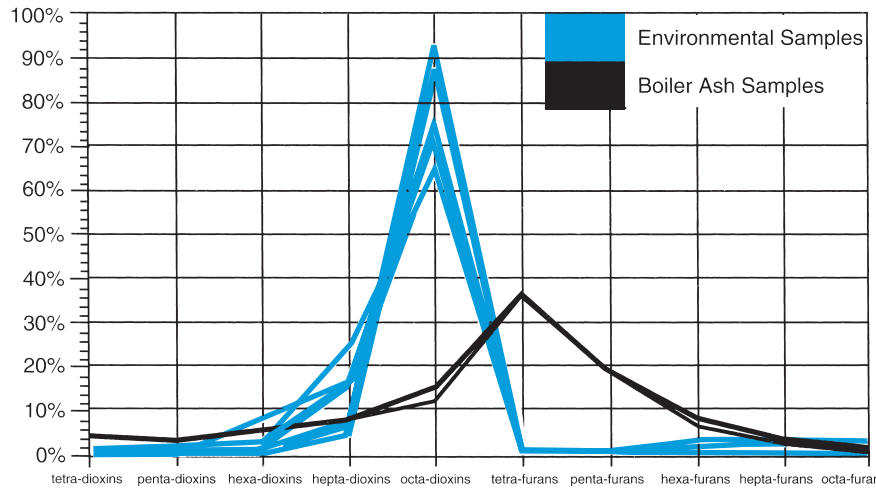


Figure 12.5
Dioxin/Furan homologue fingerprint. (There is no particular rationale for octa-dioxins to be connected by a line to tetra-furans but this is how plaintiff's expert chose to display the data.)

factor of 1000 in order to approximate the observed soil concentrations. For the deposition modeling it was assumed that all contaminants reaching the ground would be deposited in the top 2 inches (5 cm) from which the samples were drawn. The plaintiff expert's estimate of emissions was based on calculating the stack gas volume due to burning waste based on stoichiometry and then assuming that dioxins and furans would be present in this stack gas volume at the highest level found in any incinerator in an EPA survey (USEPA, 1994). The argument was that since emissions from the small industrial boiler were uncontrolled, dioxin levels should be at least this high. However, the reality is counterintuitive, in that the use of control equipment for particles or sulfur dioxide, as was the case for the incinerators surveyed by EPA, necessitates lowering gas temperatures into the dioxin formation range. In facilities without control equipment, stack gas temperature may be above the range of dioxin formation until the stack gases are emitted. In the case of the manufacturing facility boiler the stack gas temperature was at least 900˚F (482˚C) above the optimum dioxin formation range (250–450˚C or 482–842˚F). The rapid temperature drop and dilution of stack gases that occurs outside the stack would be effective in limiting dioxin formation.

The plaintiff's expert also noted that the homologue profile or fingerprint shown in Figure 12.5 was the same for all the environmental samples. According to this expert, that demonstrated that there was only a single source. However, when the congener profiles for bottom ash samples from the boiler were added, as shown in the figure, these had a distinctly different profile. In fact, the environmental samples probably resemble each other because of the preferential degradation of specific homologues based on the degree of chlorination. The bottom ash samples differ, not so much because they come from

a different source, although that is true, but because they have not been degraded, in particular photodegraded, in the same way.

As a result of these arguments we were able to conclude that, except for the creek hot spot, dioxins in soils represented background such as would be found in other similar communities. Although there were several possible explanations for the creek hot spot, the manufacturing facility operations were not one of the credible ones.

12.4 REDISTRIBUTION SUBSEQUENT TO DEPOSITION

The soil concentration patterns that have been discussed thus far assume that contaminants remain where they are deposited. Even for contaminants that are not degraded in the environment, redistribution occurs both vertically and horizontally as described next.

12.4.1 EFFECT OF SOIL PERMEABILITY

Although it does not necessarily affect the total contaminant mass in the soil column, soil concentrations can be affected by soil permeability. For example, one expects concentrations to be higher at the surface in locations where precipitation puddles and evaporates rather than percolating into the ground (e.g., in flat areas with clay soils). Vertical distribution effects can be mitigated by sampling to sufficient depth to capture the entire contaminant loading.

12.4.2 DRIP LINE ENHANCEMENT

Material deposited on roofs and walls of buildings can subsequently be washed down by rain and collect along building perimeters or "drip lines." A further enhancement may occur under drain spouts. This results in higher concentrations in samples collected near building foundations than in samples collected in open areas. Mielke (1994) found drip line enhancement of lead in soils around unpainted buildings in inner city locations compared to open yard samples. The source of lead was vehicles burning leaded gasoline. Of course, some care must be used in interpreting this effect for contaminants such as lead because drip line enhancement can also be caused by deterioration of exterior paint. Not surprisingly, Mielke found that painted buildings had a greater lead drip line enhancement than masonry buildings.

In general, with the exception of paint-related materials, a drip line enhancement of contaminants is a signature of an air source. Because of soil removal or addition, the effect may or may not be present at a specific

building. It may be necessary to sample a population of buildings in order to see a statistically significant effect. Although we are unaware of any studies that have investigated the following possibility, it seems reasonable to expect that the amount of the enhancement for a population of buildings will vary between the side of the building facing the source and other sides of the building.

12.4.3 SEDIMENT ENRICHMENT

Precipitation may redistribute materials deposited on impervious surfaces such as roads and parking lots. Finer particles, which typically contain higher concentrations of contaminants, as discussed next, are easily suspended in heavy rains, transported along drainage pathways, and collected in sedimentation areas of rivers and streams. For this reason, fine sediments collected from rivers, streams, and drainage ditches may contain higher concentrations of contaminants originating from combustion-related sources compared to open area soils. Sediment samples (and drip line samples) should be considered as populations distinct from open area soil samples.

As previously mentioned, particles emitted from controlled combustion sources are generally smaller than a few micrometers in diameter and behave essentially as a gas. Smaller particles typically have higher concentrations of condensed contaminants, because of their larger surface to volume ratio, and are more susceptible to redistribution processes than larger soil particles. Higher concentrations associated with smaller particles often are reported in terms of an enrichment ratio or enrichment factor. Schroeder *et al.* (1987) define an elemental enrichment factor (EF) based on a comparison of concentration ratios:

$$EF = \frac{C_i/C_n(ambient)}{C_i/C_n(background)} \tag{12.16}$$

where C_i is the concentration of the element of interest and C_n is the concentration of a ubiquitous background element such as iron or silicon. The ratio of C_i/C_n in ambient air particulate is compared to the ratio of C_i/C_n in background soils. The quotient is defined as the enrichment factor.

Similarly, Ghadri and Rose (1991) define an enrichment ratio for sorbed chemicals (nutrients and pesticides) as the "ratio of a chemical concentration in eroded sediment to that of the original soil from which the sediment originates." The authors hypothesize that finer particles, richer in sorbed contaminants, are also created by the action of raindrops stripping the concentrated outer layer of larger particles. In soil aggregate peeling experiments, the authors report enrichment ratios for pesticides in the removed outer layers of

natural soils of less than a factor of two. Sheppard and Evenden (1992) review enrichment ratios for metals and conclude that most natural redistribution processes result in enrichment ratios less than 10.

12.4.4 RESUSPENSION

Contaminants attached to particles may be resuspended by winds following deposition. When this occurs repeatedly it is referred to as *saltation*. Wind tunnel studies have shown that resuspension emissions rise by the cubic power of wind speed above some threshold velocity (Cowherd *et al.*, 1985). Thus the redistribution pattern over sufficient time will be determined by the high wind speed portion of the wind rose illustrated in Figure 12.1. This can be an important consideration in assessing remediation plans for hazardous waste sites where contaminated soils may be blown downwind into populated areas over a period of years. Solid fences will reduce the rate of migration during high wind conditions.

12.5 ACCIDENTAL RELEASES OF HAZARDOUS SUBSTANCES

Accidental releases of toxic substances from containers or pipelines require special attention to how the release may affect the transport of the chemical. For example, the release of a heavier-than-air gas may result in very high concentrations of a gas cloud at ground level that initially disperses at rates very different from those of a passive gas. A heavier-than-air gas will tend to flow down terrain gradients and will mix much more slowly with the ambient air. In the initial phases of release, the cloud can flow horizontally under the influence of gravity, even upwind in some circumstances. This flow behavior is of special concern to emergency planning and emergency response teams because the locations of high exposure risks may not be directly downwind of the release point and will be close to the ground. Forensic analyses of the consequences of accidental releases also need to address the release rates and durations of the releases themselves as these are very much situation-specific and, of course, have relevance to the potential exposure assessments. Some of the regulatory implications of accidental releases and emergency planning activities are discussed in Egan and Heinold (1997). Accidental releases generally involve a combination of mechanical, thermodynamic, and chemical processes. A catastrophic event often involves explosions together with the release of quantities of hazardous chemicals. In this section we will focus on the mechanisms of accidental releases that would affect populations downwind by

the transport of toxic gases. Those interested in the dynamics and consequences of explosions should review CCPS (1994) and the simple models for explosive impacts in EPA (2005).

We first describe what constitutes a heavy gas and then highlight some simple physical factors important to understanding how analyses of the consequences of accidental releases are approached.

12.5.1 HEAVY OR DENSE GASES

A dense gas is one having a specific gravity or density that is significantly greater than that of the surrounding atmosphere. A density difference greater than 0.1% to 1% commonly is associated with dense gas flow behavior. There are three release mechanisms that often result in dense gas behavior. They can occur individually or in combination.

(1) The gas may have a molecular weight greater than that of air (~29 g/mole). In this case if the released gas and the atmosphere are at the same temperature, a gas with a greater molecular weight than air, when released rapidly, will be heavy relative to air.

(2) A gas may be cooled by the expansion processes associated with the release. In this case the change of density will initially be approximately inversely proportional to the drop of temperature divided by the absolute temperature.

(3) If the gas is vaporizing from a liquid state within a containment vessel and discharging into the atmosphere, it may be entraining liquid droplets as it exits. The resulting mixture of liquid droplets and the gas will behave as a heavy gas.

In each of these cases, as ambient air is mixed into the plume, the density of the mixture will decrease. The rate of entrainment of ambient air into the plume is suppressed, however, by the difference of density at the interface. At some distance downwind, enough air will be entrained to cause the plume to be dispersed passively; that is, with the ambient air. As a cool plume travels downwind it may also gain heat from the ground, which will further decrease the plume's density. The analysis issue is often to determine how far downwind the plume will be highly concentrated and behave as a dense gas. The Center for Chemical Process Safety of the American Institute of Chemical Engineers (CCPS) has described some of the dispersion models developed to address these complicated dynamics (CCPS, 1996). Most models treat the plume as a heavy gas initially and then transition to equations such as the Gaussian equation to describe the passive dispersion further downwind.

12.5.2 ESTIMATION OF RELEASE RATES

In this section we briefly review some of the physical and thermodynamic factors that determine the release rate of an accidental release of a chemical. More information on these mechanisms and on quantitative methods can be found in the EPA's guidance documents for consequent analyses of accidental releases including EPA (1999), EPA (2004), and guidance from the CCPS (1996).

Gases under pressure at ambient temperature. If a pipe or storage tank under pressure ruptures, the resulting gas stream will tend to cool adiabatically as it expands to reach atmospheric pressure. As described earlier, the resulting gas may behave initially as a dense gas. The gas flow rate through the rupture can be estimated if one knows the shape and size of the rupture, the internal pressure, and the volume of the container. This will be a time varying rate. Also, the associated cooling may cause icing conditions at the rupture exit, further complicating the calculations.

Liquefied gases under pressure. If there are both gas and liquid phases of a compound in a container, then as the gas suddenly exits and the pressure inside drops, the liquid phase will boil and produce more gas. This is called *flash vaporization*. The boiling will continue until the temperature of the contents drops to the boiling point. One can calculate the amount of gas evaporated by flash vaporization by equating the heat of vaporization to the loss of heat from the liquid phase. The equation for the mass of the gas initially generated is

$$m_G = m_L(T_0 - T_{BP})C_P/L \tag{12.17}$$

where m_L is the initial mass of the compound in its liquid phase, T_0 is the initial temperature of the liquid phase, T_{BP} is the boiling point temperature of the compound, C_P is the specific heat of the liquid phase (cal/g/K), and L is the latent heat of evaporation for the compound (cal/g).

Evaporation of liquid spills. If a liquid spills into a containment area, or upon the ground, the release rate of the gas phase to the atmosphere depends upon the thermodynamic properties of the compound, the area of the spill, and upon the rate that the spill is heated by applicable heat transfer processes. These will generally include conduction of heat to the liquid from the ground or underlying surfaces, solar radiation, and heat from the air flowing over the spill.

Accidental spills occur from a wide variety of causes, and the physical, thermodynamic, and chemical processes that determine the spill rate and

duration can be quite complex. Much of the research work done to quantify the consequences of accidental spills has been performed in laboratories or in highly controlled field experiments under idealized circumstances. Real-world spills often have more than one contributing cause and the damage done by the accident can make causality difficult to determine. Forensic investigations therefore may require more fundamental approaches that go beyond the modeling capabilities of off-the-shelf models to explain the range of observed consequences.

12.6 SUMMARY

Forensic analyses of the contribution of specific sources to an environmental or human health outcome involves a combination of measurement and mathematical modeling techniques. For source identification based upon soil measurements there should be a rough correspondence between measured surface soil concentrations and the concentrations that result from dispersion and deposition modeling. Soil concentrations should be highest downwind in the prevailing wind direction and should be roughly proportional to wind direction frequency in other directions. In the absence of downwash and terrain a single maximum should occur downwind at a distance determined by the typical plume height, which in turn is determined by source operating parameters. Background effects will be minimized if predictions are compared to soil concentrations obtained near this maximum. Comparison will also be facilitated if soil sampling is done to sufficient depth to capture most of the soil loading, particularly in permeable soils.

Redistribution processes also affect soil concentration patterns. Soil concentrations due to airborne sources are enhanced along building perimeters. A concentration enhancement can also occur in sediments relative to soils. In some cases, ratios of different contaminant concentrations can be used to identify sources, even with redistribution, when background and degradation effects are properly accounted for.

For analyses that focus on short- or long-term exposures to contaminated ambient air, the choice and use made of atmospheric dispersion models to determine causality is important. There are a number of dispersion models available for different terrain settings and applications. They also differ by the level of technical sophistication. The performance of these dispersion models also depends critically upon the appropriateness, the quality, and the quantity of the meteorological input data.

Analyses of the consequences of accidental releases of chemicals need to carefully address the physical factors that determine the release rates to the environment and also how the subsequent dispersion of the releases might be

affected by the thermodynamic state of the release and the magnitude of the release.

ACKNOWLEDGMENT

Certain of the preceding material appeared in different form in the *Environmental Claims Journal*, Vol. 11, No. 2, Winter 1999.

REFERENCES

Bair, F.E. (1992) *Weather of US Cities*. Gale Research, Detroit, MI.

CCPS. (1994) *Guidelines for Evaluating the Characteristics of Vapor Cloud Explosions, Flask Fires, and BLEVEs.* Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York, NY.

CCPS. (1996) *Guidelines for Use of Vapor Cloud Dispersion Models, 2e*. Center for Chemical Process Safety of the American Institute of Chemical Engineers, New York, NY.

Christen, K. (2000) Arsenic in old herbicides comes back to haunt Denver. *Environmental Science and Technology* 34, 376A–377A.

Cimorelli, A., Wilson, R.B., Perry, S.P., Venkatram, A., Weil, J.C., Paine, R.J., *et al.* (1998) Minimum Meteorological Data Requirements for AERMOD—Study and Recommendations. Draft Document, EPA December 14, 1998.

Cimorelli, A. *et al.* (2005) AERMOD: A Dispersion model for industrial source applications. Part I: General model formulation and boundary layer characterizations. *Journal of Applied Meteorology* 44(5), 682–693.

Cowherd, C., Muleski, G., Engelhart, P., and Gillette, D. (1985) *Rapid Assessment of Exposure to Particulate Emissions from Surface Contamination*. Office of Health and Environmental Assessment, US Environmental Protection Agency, Washington, DC, EPA/600/8–85/002, NTIS PB85-192219.

DiCristofaro, D.C. and Hanna, S.R. (1989) *OCD; The Offshore and Coastal Dispersion Model, Version 4. User's Guide.* NTIS Nos. PB 93-144384 and PB93-144392.

Egan, B.A. and Heinold, D.W. (1997) *A Guide to EPA's Risk Management Program under Section 112(r) of the Clean Air Act: Strategies for Effective Compliance*. Air and Waste Management Association Publication OTHP-10. Pittsburgh. PA

Egan, B.A. and Snyder, W.E. (2002) *Atmospheric Dispersion: Complex Terrain.* Volume 1. Encyclopedia of Environmetrics. John Wiley & Sons, Ltd, Chichester, UK, 109–125.

Ghadri, H. and Rose, C.W. (1991) Sorbed chemical transport in overland flow: 1. Nutrient pesticide enrichment mechanism. *Journal of Environmental Quality* 20, 628–633.

Hanna, S.R., Briggs, G.A., and Hosker, R.P. (1982) *Handbook on Atmospheric Diffusion*. Technical Information Center, US Department of Energy.

Kimbrough, D.E. and Suffet, I.H. (1995) Off-site forensic determination of airborne elemental emissions by multi-media analysis: A case study at two secondary lead smelters. *Environmental Science and Technology* 29, 2217–2221.

Mielke, H.W. (1994) Lead in New Orleans soils: New images of an urban environment. *Environmental Geochemistry and Health* 16(3/4), 123–128.

Paine, R.J. (2001) Meteorological Input Data for AERMOD Applications. *AWMA Conference on Air Quality Models: A New Beginning.* Newport, Rhode Island. April 2001.

Patil, R.S. (2005) Performance evaluation of CFD Model PANAIR for air pollution dispersion of industrial stack emissions. *Proceedings of NOAA/EPA Golden Jubilee Symposium on Air Quality Modeling and Its Applications.* USEPA, RTP, NC.

Perry, S.J., Burns, D.J., Adams, L.H., Paine, R.J., Dennis, M.G., Mills, M.T., *et al.* (1989) *User's Guide to CTDMPLUS: Volume 1. Model description and User's Instructions*. EPA Pub No. EPA-600/8-89-041. US EPA, RTP, NC.

Perry, S.J., Burns, D.J., and Cimorelli, A.J. (1990) *User's Guide to CTDMPLUS: Volume 2. The Screening Mode (CTSCREEN)*. EPA Pub No. EPA-600/8-90-087. US EPA, RTP, NC.

Randerson, D. (1984) *Atmospheric Science and Power Production* U.S.DOE Doc No. DOE/TIC-27601. Available from NTIS. US Dept of Commerce. Springfield VA.

Schroeder, W.H., Dobson, M., Kane, D.M., and Johnson, N.D. (1987) Toxic trace elements associated with airborne particulate matter: A review. *Journal of the Air Pollution Control Association* 37(11), 1267–1285.

Scire, J.S., Strimaitis, D.J., and Yamartino, R.J. (2000) *A User's Guide for the CALPUFF Dispersion Model*. Earth Tech, Inc., Concord, MA.

Sheppard, S.C. and Evenden, W.G. (1992) Concentration enrichment of sparingly soluble contaminants (U, Th, and Pb) by erosion and by soil adhesion to plants and skin. *Environmental Geochemistry and Health* 14, 121–131.

Turner, D.B. (1994) *Workbook of Atmospheric Dispersion Estimates—An Introduction to Dispersion Modeling*. CRC Press, Chapel Hill, NC.

US Environmental Protection Agency. (1994) *Combustion Emissions Technical Resource Document* (CETRED), Draft. Office of Solid Waste and Emergency Response, Washington, DC. EPA 530-R-94-014.

US Environmental Protection Agency. (1995a) *Air Quality Criteria for Particulate Matter*, Vol. I, EPA/600/AP-95/001a. Office of Research and Development, Washington, DC.

US Environmental Protection Agency. (1995b) *User's Guide for the Industrial Source Complex Dispersion Models. Volume II.* Doc. EPA-454/B-95-003b.

US Environmental Protection Agency. (1999) *Risk Management Program Guidance for Offsite Consequence Analysis*. EPA Doc. 550-B-99-009 (www.epa.gov/ceppo/).

US Environmental Protection Agency. (2003) *Exposure and Human Health Reassessment of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and Related Compounds.* National Academy Sciences (NAS) review draft. U.S. Environmental Protection Agency.

US Environmental Protection Agency. (2004) *User's Manual for ALOHA*. Chemical Emergency Preparedness and Prevention Office. Washington, DC.

US Environmental Protection Agency. (2005) *Appendix W to Part 51—Guideline on Air Quality Models*, 40 CFR Ch. I (11-9-05 Edition).

CHAPTER 13

ENVIRONMENTAL FORENSIC MICROSCOPY

James R. Millette and Richard S. Brown

13.1 Introduction 611
13.2 Sampling and Analysis Equipment 612
13.2.1 Air Sampling 612
13.2.2 Surface Dust Sampling 613
13.2.2.1 Microscopy Equipment 614
13.3 Determining the Nature of Contaminants 615
13.3.1 Particle Analysis 615
13.3.2 Product Identification by Microscopy 616
13.3.2.1 Sampling 616
13.3.2.2 Analysis 617
13.4 Measuring the Extent of a Specific Contaminant 619
13.4.1 Asbestos 619
13.4.1.1 Vermiculite Analysis 619
13.4.2 Nonasbestos Fibers 620
13.4.2.1 ISO Method 14966 620
13.4.2.2 Glass Fibers 621
13.4.2.3 Ceramic Whiskers 621
13.4.3 Nonfibrous Particulate 622
13.5 Case Studies—Examples of Environmental Forensic Microscopy Investigations 624
13.5.1 Elevated Lead in a Child 624
13.5.2 A Spot Called Ralph 625
13.5.3 Automobiles with a Sooty Deposition—1 626
13.5.4 Automobiles with a Sooty Deposition—2 626
13.5.5 WTC Signature Search 628
References 630

13.1 INTRODUCTION

Environmental forensic microscopy is the application of microscopy to the identification, collection, and analysis of small particles and the interpretation of the analyses performed as they pertain to environmental investigations

and litigation. Environmental investigations are generally assessments of the nature and extent of a contaminant. Microscopy is most useful in examining, classifying, and identifying particulate contaminants. Determining the nature of a contaminant and measuring the extent of a contaminant are two fundamentally, although often overlapping, questions. Often the questions put to the environmental forensic microscopist are "What is this dust?" and "Where did it come from?" However, the environmental forensic microscopist may also be asked to quantify the amount of a specific contaminant in several samples in an effort to determine the extent of the problem and to establish a procedure for further sampling and analysis. Answering the two questions has given rise to different sets of methods. In answering the question about how much of a specific contaminant is present, that contaminant must be sorted and identified. Other particles comprising the matrix of the sample may be ignored. In answering the questions about the nature of a contaminant dust, a broader range of knowledge may be required to sort the various contaminants into appropriate classes or categories. Environmental forensic microscopy measurement methods tend to be fairly specific with a designated set of counting rules defined so that different analysts can get comparable numerical results. Environmental forensic microscopy methods dealing with the nature of a contaminant sample tend to be less formalized because they must take into account a vast variety of potential particle types. In answering the question "What is this dust?", the microscopist cannot ignore all the particle types but one. The microscopist must allow the constituents of the sample to dictate which basic standard microscopy procedures and other techniques he or she will use in solving the analytical problem. When dealing with an environmental particle contamination situation of unknown origin, the forensic investigation of the nature of the contaminant naturally takes place first. However, in some situations where the nature of the contaminant is known, measurement methods may be employed first.

13.2 SAMPLING AND ANALYSIS EQUIPMENT

The sampling and analysis methods used in environmental forensic investigations are drawn primarily from the criminal forensics, industrial hygiene, and environmental monitoring areas. Combining various aspects of these disciplines allows the investigator to generate a procedure that fits the varied and sometimes very complex environmental situation.

13.2.1 AIR SAMPLING

Collecting samples of airborne particles for environmental forensic investigations usually is done using membrane filter sampling cassettes. Various types of membrane filters are used in environmental and industrial hygiene

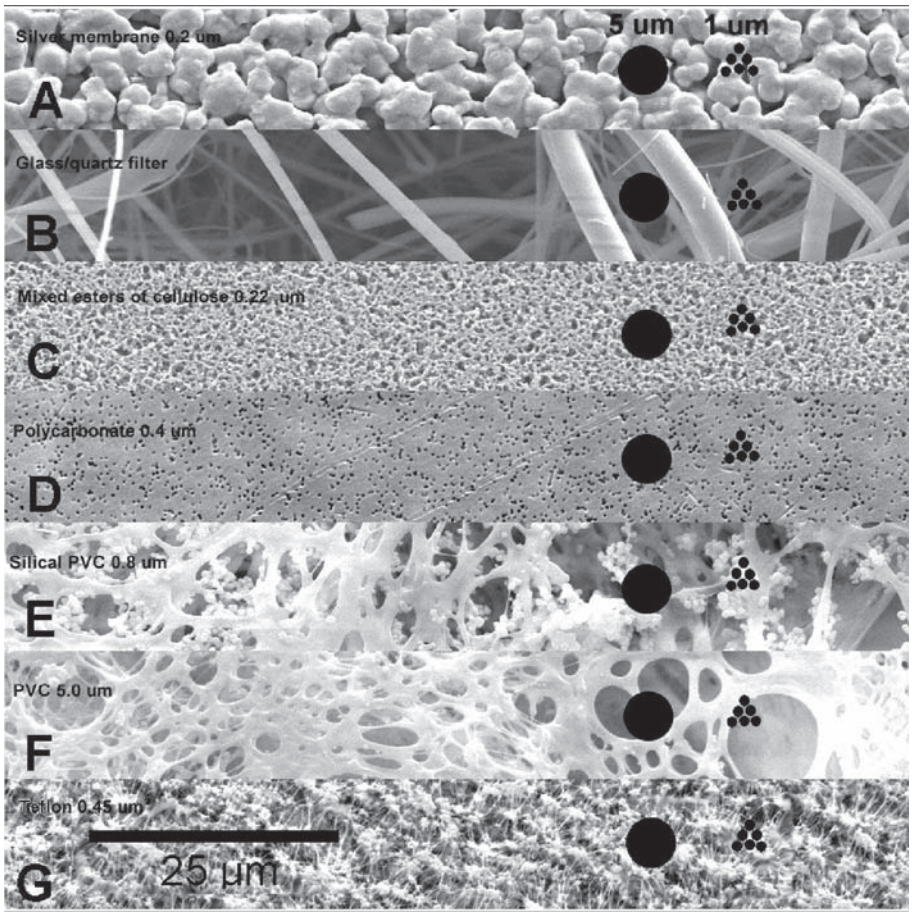


Figure 13.1
Scanning electron microscope images of filters used to collect airborne particles. Spheres of 5 micrometers and 1 micrometer are overlaid to show how small particles would appear on the filter surface.

sampling. These include filters made of mixed cellulose ester polymer, polycarbonate polymer, glass fiber filters, quartz fiber filters, silver membrane filters, and polyvinyl chloride (PVC) filters. Figure 13.1 shows a comparison of the surfaces of the various filter types. Spheres representing particles of 5 micrometer diameter and 1 micrometer diameter are overlaid on the images. With the glass fiber filters, particles can be lost in the filter thickness. Backwashing the glass filter to remove the particles produces a sample that contains glass fiber fragments. It is clear why microscopists prefer air samples collected on polycarbonate (PC) filters. The PC filters have a flat surface because they are a "straight through" filter rather than a tortuous path design that is found in the other filter types.

13.2.2 SURFACE DUST SAMPLING

Surface dust particle collection techniques, varying according to situation, are chosen to maximize probative value and when possible to preserve a

portion for future analysis. Common particle collection techniques for environmental forensic microscopy investigations of surface particulate include scraping, brushing, taping with cellophane tape on a slide or post-it notes, scoop and bag techniques, wipe sampling, and vacuuming with either a microvac (air cassette) or large-scale vacuuming (Ness, 1994; Millette and Few, 2001). Sample collection procedures with cotton or polyester balls, wipes, and tape slides for outdoor surfaces are described in the ASTM Standard Practice D6602 (ASTM, 2003). ASTM D6602 is a standard procedure that serves as an excellent basis for environmental forensic microscopy investigations where soot might be involved. This ASTM standard was designed primarily for the determination of carbon black among soot particles and other dark particles but provides the framework for the microscopy studies necessary to determine the identity of all particles and possible sources of surface contamination.

13.2.2.1 Microscopy Equipment

Microscopic analyses in environmental forensic investigations are accomplished utilizing a combination of visible light, infrared light, and electron microscopy. The light microscopy usually is performed with polarized light microscopy (PLM), but may involve phase contrast (PCM), darkfield, or fluorescence microscopy. Infrared microscopy is done using Fourier transform infrared microspectroscopy (micro-FTIR). FTIR is very useful when identifying organic molecules such as plastics and polymers. Scanning electron microscopy (SEM) allows the analyst to see particles that are smaller than can be seen with light microscopy, and when equipped with an x-ray analysis unit, allows the analyst to determine the elemental composition of the particles. Transmission electron microscopy (TEM) also allows the analyst to see particles that are smaller than can be seen with light microscopy, and when equipped with electron diffraction capabilities and an x-ray analysis unit, allows the analyst to determine the crystal structure and elemental composition of the particles. ASTM Practice D6602 lists different types of microscopes that can be used to investigate environmental particles: a stereobinocular microscope, capable of 1 to 60× magnification; a polarized light microscope (PLM), equipped with objectives in the 4 to 100× range of magnification (for a total magnification between 40 and 1000×); a transmission electron microscope (TEM), equipped with a suitable camera; and a scanning electron microscope (SEM), equipped with energy or wavelength dispersive analysis equipment (EDS or WDS). A TEM equipped with selected area electron diffraction (SAED) and EDS is referred to as an analytical electron microscope (AEM).

13.3 DETERMINING THE NATURE OF CONTAMINANTS

For questions involving the nature of particulate contaminants, the general procedure in environmental forensic investigations is the same as the general forensic science approach to trace analysis; that is, to collect samples of the particles (dust, dirt, or suspensions in liquid) in question, evaluate the samples to classify and identify the particles, identify possible sources, and then to compare the sample particle types to the suspect source(s) particle types (see Table 13.1).

13.3.1 PARTICLE ANALYSIS

The McCrone Particle Atlas is the standard reference for environmental forensic microscopical particle analysis (McCrone, 1973). Published first in 1972, this six-volume set (now available only on CD-ROM) contains thousands of images of different types of particles and information about their microscopical characteristics observable by light and electron microscopy. There are also clear discussions of various topics important to the proper use of PLM in identifying particles including birefringence, refractive index, and dispersion staining. Another useful reference atlas for the forensic microscopist was published by Petraco and Kubic (2003). Illustrations and information about the morphological characteristics of pollen grains can be found in Faegri, Iversen, and Waterbolk (1964). Images and information about the appearances of fungal spores can be found in Hawksworth *et al.* (1995). Images of soil minerals can be found in Graves (1979), Krumbein and Pettijohn (1938), Palenik (1979), and Hopen (2004), along with useful information about the identification of soil minerals as individual (detrital) grains. Images of soot have been published by Medalia and Rivin (1982), Huffman *et al.* (2000), and Clague *et al.* (1999). Images of aciniform flare combustion can be seen in Kostiuk, Johnson, and Thomas (2004).

Although environmental forensic microscopical particle analysis of dusts may seem daunting due to the many types of particles possible, it has been found that for most common situations such as normal residential and

Collect
Analyze
Identify particle type
Identify possible sources
Compare sample particles to suspect source(s)
Report
Interpret findings as they pertain to law/science

Table 13.1

General forensic science approach to trace evidence analysis.

building dust, the particles can be classified using approximately 20 categories (Millette, Lioy *et al.*, 2003; Turner, Millette *et al.*, 2005).

13.3.2 PRODUCT IDENTIFICATION BY MICROSCOPY

Product identification is a special case of environmental forensic microscopy because it involves, in addition to particle classifications, the combining of the results of several analytical techniques to determine the composition of a product and then compare the overall results to product formula information. The same process may be used to compare dust particles with building products or other man-made compounds. Sections 13.3.2.1 and 13.3.2.2 describe the general procedures used for product identification when matching asbestos-containing products with manufacturers' formulae.

13.3.2.1 Sampling

Sampling involves removing the material in a manner that will maintain the layer structure and integrity of layered samples. This can be accomplished with a polycarbonate tube used as a cork borer or by carefully cutting through the product down to the substrate, wrapping in paper and cushioning the sample so as the product cannot be crushed or crumbled. Wet samples should be thoroughly air dried prior to packaging to discourage fungal and mold growth.

Sampling a variety of products having different physical dimensions and cohesion requires the person doing the sampling to be creative and cognizant of the type of analysis that will be performed. Chain-of-custody forms must accompany the samples to the laboratory. Once the sample reaches the laboratory the packing is carefully removed and the sample's condition "as received" can be recorded, in some cases, with photographs. Examination with reflected light under low magnification is performed to determine if the sample represents a single product or multiple products displayed as multiple layers. If multiple layers are present, the color and layer sequence can be documented with a photograph. The layers are examined as individual products as the analysis continues. Each sample is split into four vials. One vial is sent to each analysis station for PLM, SEM-EDS, TEM, or acid-soluble weight percent. Quality control for each analysis is checked by comparing the results of each examination for a particular sample against each other. For instance, a sample that contained chrysotile as identified by PLM should also have chrysotile characterized by SEM and identified by TEM. Should an ingredient that was determined to be a major component of a sample not be observed by each of the microscopical techniques, a review of the analysis would be undertaken to determine why. Some particles, such as montmorillonite clay, may be of a

Acoustical plaster
Ceiling tiles
Fireproofing
Pipe insulation

Table 13.2
Common asbestos-containing product types that have been tested for product ID.

Diatoms	Starch
Kaolinite	Perlite
Montmorillonite	Pumice
Fiberglass®	Hair
Quartz and soil minerals	Lithopone
Iron chromite	Titanium dioxide
Chrysotile	Mineral wool (soluble in some procedures)
Amosite	Cotton/paper cellulose fibers
Crocidolite	Vermiculite
Tremolite, actinolite, anthophyllite	Mica
Fly-ash	

Table 13.3
Acid insoluble particle types.

Mineral wool (soluble in some procedures)	Calcium and magnesium silicates
Calcium carbonate	Gypsum
Portland cement	Sodium silicates

Table 13.4
Common acid soluble particles.

low percentage that when combined with other ingredients, such as Portland cement, may be difficult to find. Here looking at the acid insoluble material collected on the filter may be necessary to confirm the presence of the clay.

13.3.2.2 Analysis

A combination of microscopy techniques is used to analyze asbestos-containing building materials. Polarized light microscopy (PLM) is used in combination with stereobinocular microscopy to estimate the volume percent of the constituents. Table 13.2 lists some common product types that are the subjects of product identification. Tables 13.3 and 13.4 list the common constituents that are readily identified using polarized light microscopy. A description of the sample includes color and number of layers present. A portion of the sample is placed into a refractive index liquid of known value. The sample is observed using the PLM to note relative refractive index, color, pleochroism, birefringence, morphology, and relative abundance. Particles such as paint, calcium carbonate, Portland cement, synthetic calcium silicate, synthetic magnesium silicate, asbestos, paper fiber, synthetic fiber, glass fibers, mineral wool fibers (slag and rock wools), starch, and clays can be characterized and their volume percentage estimated.

Scanning electron microscopy coupled with energy-dispersive x-ray spectrometry (SEM-EDS) is used to analyze a portion of the sample to complement

the PLM analysis. The sample is prepared by adding a few milligrams to a clean glass vial (disposable plastic pressure fit cap half-dram vials work well). Add one drop of collodian and 15 to 20 drops of amyl acetate. Using clean forceps, crush and stir the sample in the vial. Using a clean glass Pasteur pipette, mix gently and draw a few drops of the mixture into the pipette. Deposit onto a 12 millimeter diameter carbon planchet controlling the loading of solids by adding and removing the mixture until a light loading of the sample particulate has been deposited on the planchet. Allow to dry three to four hours and carbon coat. Analysis is carried out at 500× magnification. Find a heavily loaded area and collect an image and an EDS spectrum at 100× magnification to get a representation of the elements present. Search the majority of the planchet with a combination of secondary electron and backscattered electron imaging. Collect EDS spectra and images of particles that represent each type of constituent present. The SEM-EDS is especially useful when looking at glass fibers, Portland cement, gypsum, mineral wools, heavy metals, paint, perlite, and pumice. Clays such as montmorillonite and kaolinite can be differentiated if there are a minimum of interfering materials present. Diatoms, synthetic magnesium silicates, and calcium silicates can also be characterized by morphology and elemental composition.

Transmission electron microscopy coupled with an energy dispersive x-ray spectrometer and capable of electron diffraction (AEM) is useful to characterize and identify mineral fibers (such as asbestos), clays, pigments, and other very fine, thin particles that may present themselves. The sample is prepared as a drop mount directly on a carbon film copper grid support and placed directly into the AEM. Collect EDS spectra and images of particles that represent each type of constituent present.

To determine the acid-soluble weight percent, approximately 250 milligrams of the sample is weighed and dissolved in warm 10% hydrochloric acid. The undissolved particles are recovered by filtering through a preweighed polycarbonate filter (0.4 micrometer pore size). The insoluble fraction will typically contain clays, glass fibers, perlite, pumice, polymers, paint fragments, and other acid insoluble materials (see Table 13.2).

Occasionally, difficult or complex samples will require the analysis of the acid insoluble particles collected on the polycarbonate filter by PLM, SEM, and AEM. Occasionally multilayered samples will require all the forementioned analyses performed on each layer. Some particles will actually be aggregates such as concrete or paint flakes and will require further testing.

The final report involves matching results of particle classification to formulae that have been obtained from manufacturers and entered into a database. The formula information from various manufacturers varies considerably. So complex were the terms used to describe processes and so varied between manufacturers were the terms and proprietary names for ingredients that in

one case hundreds of hours were needed to analyze documents associated with the formulas and a dictionary of terms was created to decipher the formulae (Hopen, Millette *et al.*, 1999).

13.4 MEASURING THE EXTENT OF A SPECIFIC CONTAMINANT

13.4.1 ASBESTOS

There are quite a few standard microscopical methods for determining quantities of one environmental contaminant: asbestos. The asbestos methods use one of four types of microscopy: polarized light microscopy (PLM), phase contrast microscopy (PCM), transmission electron microscopy (TEM), or scanning electron microscopy (SEM). The TEM analysis includes characterization of fiber crystal structure by selected area electron diffraction (SAED). Methods for both types of electron microscopy (SEM and TEM) utilize an energy dispersive x-ray spectrometer (EDS) to analyze the fibers for their elemental composition. Many methods have been developed for determining asbestos levels in bulk building materials or concentrations in airborne samples, and some have been developed for asbestos in surface dust samples (Millette and Hays, 1994; ASTM, 2002). Because a chapter was devoted to asbestos in Volume 1 of *Environmental Forensics: Contaminant Specific Guide* (Van Orden, 2006), and additional information on available asbestos analysis methods can be found in Millette and Bandli (2005) and Millette (2006), only one asbestos-related environmental forensic microscopy method will be considered here.

13.4.1.1 Vermiculite Analysis

Environmental forensic microscopy can be used to determine if a sample of vermiculite can be traced to a specific mine in Libby, Montana. Vermiculite products originating from the mine near Libby, Montana are currently in millions of homes and businesses across the nation. Associated with the Libby vermiculite are fibrous amphiboles including tremolite-asbestos. Many of these amphiboles occur in the asbestiform habit. Recent studies of the health of former mine workers and residents of Libby, Montana have shown an increased incidence of asbestos-related lung disease (ATSDR, 2001). The former W.R. Grace vermiculite mine and milling operations and numerous residences and commercial properties in Libby have been added to the U.S. Environmental Protection Agency's (EPA) National Priorities List for immediate environmental remediation.

Because other vermiculite products are not contaminated with fibrous amphiboles, it is an important environmental forensic question to determine

whether a particular vermiculite material is from Libby or not. The USEPA Research Method for Sampling and Analysis of Fibrous Amphibole in Vermiculite Attic Insulation (VAI) (USEPA, 2004) can be used to analyze a sample of vermiculite for fibrous amphiboles that are characteristic of Libby vermiculite. In the EPA VAI Method (sometimes called the Cincinnati Method), samples are processed/separated using a water density gradient into three fractions: floats, sinks, and suspended particles. The EPA VAI Method contains the procedures for making two measurements of amphibole in VAI. Light microscopy is used to analyze the sinks fraction because denser particles that sink to the bottom may contain large amphibole fiber bundles. Electron microscopy is used to analyze the suspended particle fraction because some fine amphibole fibers may be suspended in the water. Using the cone and quartering technique replicate subsamples of approximately 10 grams each are produced from the original ziplock bag of material (usually a one-gallon container). These subsamples are dried for two hours at 100 degrees and weighed using an analytical balance. Following the methods described in EPA/600/R-04/004, the denser particles (sinks) are separated from the less dense particles (floats). The sinks are dried overnight on a hot plate at 60 degrees and weighed. The sinks are examined at low magnification utilizing a stereobinocular microscope. Fiber bundles are picked with tweezers. The total weight of fiber bundles is determined for each subsample. Representative fiber bundles are then analyzed by polarized light microscopy (PLM). Representative fiber bundles from one of the subsamples may also be examined and analyzed by scanning electron microscopy (SEM) coupled with an energy dispersive x-ray spectrometry (EDS) system.

If no fibers are found by stereobinocular microscopy, the suspended particle fractions are prepared from the subsamples using the liquid suspensions after the floats and sinks have been removed. The suspensions are brought up to a volume of one liter and sonicated for two minutes. The suspensions are then agitated by bubbling filtered oxygen for one minute through the liquid using a 10 ml glass pipette at a flow rate of approximately 4 liters per minute. One- and 10-milliliter aliquots of each suspension are removed and filtered through 0.2 μm pore size polycarbonate filters. The filters are prepared and analyzed following the standard procedures as described in ISO 13794. Analyses are performed on a transmission electron microscope with an x-ray analysis system.

13.4.2 NONASBESTOS FIBERS

13.4.2.1 ISO Method 14966

The International Standards Organization (ISO, 2002) Method 14966 uses an SEM-EDS approach for determination of the concentration of inorganic

fibrous particles in the air. It was designed to consider various types of fibers longer than 5 micrometers in length. With the ISO method, an analyst may count fibers on air filters and determine concentrations of different fiber types based on their elemental composition. The method specifies the use of gold-coated, capillary-pore, track-etched membrane filters, through which a known volume of air has been drawn. Using energy-dispersive x-ray analysis, the method can discriminate between fibers with compositions consistent with those of the asbestos varieties (e.g., serpentine and amphibole), gypsum, and other inorganic fibers. Annex C provides a summary of fiber types that can be measured. Information is included about fibers of mullite ($Al_6Si_2O_{13}$), willemite (Zn_2SiO_4), sillimanite (Al_2SiO_5), dumortierite ($Al_7O_3(BO_3)(SiO_4)$), wollastonite ($CaSiO_3$), and several zeolites; industrial crystalline fibers such as silicon carbide (SiC), zirconia (ZrO_2), and tungsten (W); inorganic amorphous fibers such as glass wool, rock wool, and slag wool, ceramic, and quartz.

ISO 14966:2002 is applicable to the measurement of the concentrations of inorganic fibrous particles in ambient air. The method is also applicable for determining the numerical concentrations of inorganic fibrous particles in the interior atmospheres of buildings. It is important to note that the ability of the method to detect and classify fibers with widths lower than 0.2 micrometers is limited. If airborne fibers in the atmosphere being sampled are predominantly less than 0.2 micrometers in width, a transmission electron microscopy method is recommended.

13.4.2.2 Glass Fibers

Glass fibers found in airborne samples collected on membrane filters can be analyzed by phase contrast microscopy using the "B" rules of the National Institute of Occupational Safety and Health (NIOSH) 7400 method (NIOSH, 1994a). Although the phase contrast method does not identify the type of fiber counted, it does provide a useful system for monitoring airborne concentrations in areas where glass fibers are manufactured or products are fabricated with the fibers. In situations where mixtures of fibrous materials are anticipated, the PCM method should be augmented with additional microscopical procedures such as polarized light microscopy (PLM) or electron microscopy. Although not a standard method, surface dust samples were analyzed for glass fibers by Schneider *et al.* (1990). They used gelatinous foils to collect the samples and then analyzed them by light microscopy.

13.4.2.3 Ceramic Whiskers

There are four standard microscopy measurement methods related to airborne ceramic whiskers including silicon carbide and silicon nitride fibers (ASTM, 2001a,b,c,d). The scanning electron microscopy method for single-

crystal ceramic whiskers (SCCW) is a method that has the ability to distinguish among many different types of fibers based on energy dispersive x-ray spectrometry (EDS) analysis. This test method may be appropriate for other man-made mineral fibers (MMMF). This test method is applicable to the quantification of fibers on a collection filter that are greater than 5 micrometers (μm) in length, less than 3 μm in width, and have an aspect ratio equal to or greater than 5:1. The data are directly convertible to a statement of concentration per unit volume of air sampled. This test method is limited by the diameter of the fibers visible by SEM (typically greater than 0.10 to 0.25 μm in width) and the amount of coincident interference particles. The transmission electron method is applicable to the quantification of fibers on a collection filter that are greater than 0.5 μm in length and have an aspect ratio equal to or greater than 5:1.

13.4.3 NONFIBROUS PARTICULATE

There is a standard procedure for determining particle size information on samples of environmental particulate emissions. Administered by the Louisiana Department of Environmental Quality (LDEQ), the Louisiana Environmental Laboratory Accreditation Program (LELAP) is recognized by the National Environmental Laboratory Accreditation Program (NELAP). Under the LELAP program, samples of particulate from industrial emissions that have been collected on filters (PC filters are preferred although they do have a temperature limit) are analyzed for total particle sizing by automated scanning electron microscopy. The analyses are performed using a scanning electron microscope operating in automated mode under the control of an x-ray analysis system utilizing a standard operating procedure. Figure 13.2 shows a dispersion of lead-rich particles in the backscattered electron mode and Figure 13.3 shows how the automated sizing procedure assigns dimensional values to all the particles. Approximately 1000 to 2000 particles from each sample are individually sized and classified according to five client-requested size categories, typically, 0.5–2.5, 2.5–5.0, 5.0–7.5, 7.5–10.0 and >10 μm. The particle size data are presented in terms of particle number and in terms of estimated mass. The assumption is made that the particles are all of similar density and therefore the particle volume distribution is equivalent to the particle mass distribution.

Automated Particle Analysis is a useful environmental forensic microscopy tool. Digital imaging under computer control with the scanning electron microscope makes possible the morphological and chemical analysis of hundreds or thousands of particles without intervention by the microscopist. With computer control of the electron beam and the sample position, digital images

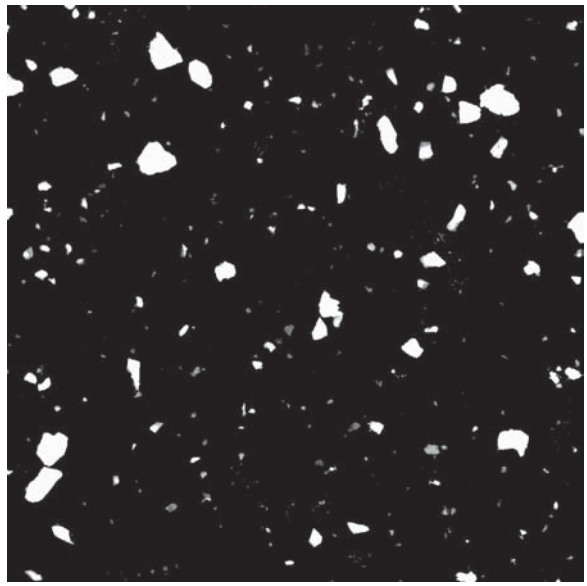


Figure 13.2

Backscattered electron image of lead ore dust sized by automated scanning electron microscopy.

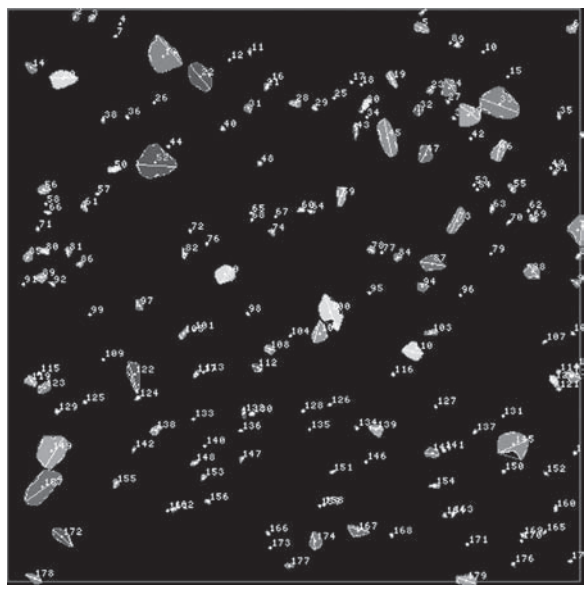


Figure 13.3

Same image field as Figure 13.2 after the microscope computer has applied conditions of particle definitions for automated particle sizing. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

can be acquired into computer memory and the image searched for particles based on computer programs that recognize image contrast. Size and shape calculations can be made and recorded for each particle found and the electron beam can be driven back to the particle for chemical analysis by energy dispersive x-ray spectrometry. Automated SEM particle analysis has been used extensively in the investigation of leaded particles in soil (Cotter-Howells and

Thornton, 1991; Hunt *et al.*, 1991, 1998; Rybicka, Wilson, and McHardy, 1994; Thornton *et al.*, 1994; Franck and Herbarth, 2002; Pirrie, 2003) and in other types of environmental forensic studies including the size of urban air particles (Jambers, De Bock, and Van Grieken, 1995) and those involving cement particles and grout (Millette and Brown, 2002).

13.5 CASE STUDIES—EXAMPLES OF ENVIRONMENTAL FORENSIC MICROSCOPY INVESTIGATIONS

13.5.1 ELEVATED LEAD IN A CHILD

It is well known that young children can be exposed to lead by ingesting lead paint dust that is generated by the deterioration, sanding, or scraping of lead-based paints in the home. Other sources of lead present in homes include deposits of particles from the use of leaded gasoline and from industrial sources like smelting. In one case, finding a child with elevated blood lead raised the concern that the child was ingesting lead-based paint dust. However, x-ray fluorescence testing of the paint in the home detected no lead. Samples of the dust were collected on adhesive notes and sent to the laboratory for analysis by SEM-EDS. The results showed the source of the lead in dust was lead-containing fly ash (Millette, Brown *et al.*, 1991). A search of the neighborhood resulted in sampling piles of material at a nearby industrial site (see Figure 13.4). The SEM-EDS showed lead-containing fly ash consistent with that found within the home (see Figure 13.5). The So-Green company apparently had been paid to receive tons of electric arc fly ash from a company in another state and was turning it into a product to be sold as a fertilizer. Unfortunately the fertilizer business was not keeping up with the shipments of fly ash and large piles were accumulating around the facility. The site was



Figure 13.4

Piles of electric-arc fly ash at the So-Green site. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

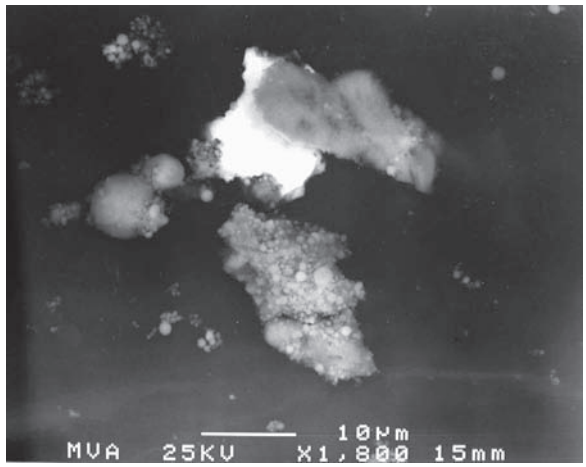


Figure 13.5
Backscattered electron image of fly ash and lead-rich particle (white) typical of particles found in homes surrounding the So-Green facility.

ultimately declared a Superfund site. In other cases involving concern over lead particulate, samples were collected from homes and analyzed by SEM-EDS. The leaded particles were classified as lead paint (Pb-Cr, Pb-Zn), auto exhaust (Pb-Br), industrial lead (Pb+metals), soil lead (Pb-Si, Al and/or Ca), and undetermined (Pb-only) based on the other elements associated with the lead in the particle. In one project, lead-arsenate particles were found in household dust (Millette *et al.*, 1995). In another project, lead-chlorine particles from PVC electric cable covering were found in an office building (Ghazi and Millette, 2005).

13.5.2 A SPOT CALLED RALPH

In a courthouse in South Carolina a mysterious stain appeared in the new carpet. Employees even gave the spot a name—"Ralph." At first it was the size of a half dollar but it grew after cleaning to about 2 square feet. Environmental mold specialists initially tested the stain in the carpet and determined that it was not caused by mold. A section of the stain was cut from the carpet and delivered to the environmental forensic microscopy laboratory for inspection. Analysis by light and scanning electron microscopy showed that the carpet contained a variety of particles typical of the particles often found in office dusts. A sticky substance was also found on the carpet fibers. FTIR analysis of the sticky material showed that it was consistent with corn syrup. It is apparent that someone spilled a soft drink on the carpet and the stain was caused by office dust particles adhering to the sticky drink residue. Efforts to clean the stain removed the dark office dust particles but did not completely remove the sticky drink residue. In fact, the cleaning efforts spread the sticky residue, which collected more office dirt over time and therefore appeared to grow in size. Not surprisingly, the newspaper reporter who interviewed the laboratory



Figure 13.6
Dark "sooty" looking deposit on an automobile trunk surface.

after the findings were made public was disappointed that the stain called "Ralph" was not something exotic, just something caused by a spilled soft drink.

13.5.3 AUTOMOBILES WITH A SOOTY DEPOSITION—1

The ASTM standard, D4610, states "Microbial growth is a major cause of discoloration and deterioration of paint films." This is true not only of painted house surfaces but also of automobile finishes. Environmental forensic microscopy investigations of what is causing the dark "sooty" layer on light colored automobiles such as the one shown in Figure 13.6 were performed using the ASTM D6602 standard (ASTM, 2003a). A clear cellophane tape is applied to the surface of the auto, removed, and placed on a microscope slide. The use of a PLM microscope at magnifications from 50 to 400× showed that the darkening agent on the automobile shown in Figure 13.6 was a biological growth (see Figure 13.7). A cotton ball is used to collect a sample for analysis by transmission electron microscopy of the darkening agent if it is suspected of being soot. The TEM analysis is the only way to differentiate various types of aciniform soots including carbon black that have particle sizes in the nanometer range. In a similar case, a residential mailbox was reportedly covered with soot from a nearby power plant. The environmental forensic investigation showed that the darkening agent on the mailbox was not soot but a sooty mold.

13.5.4 AUTOMOBILES WITH A SOOTY DEPOSITION—2

Cars parked in an industrial area were becoming coated with "black, sticky debris." The debris was described as "sticky and difficult to wash off" and

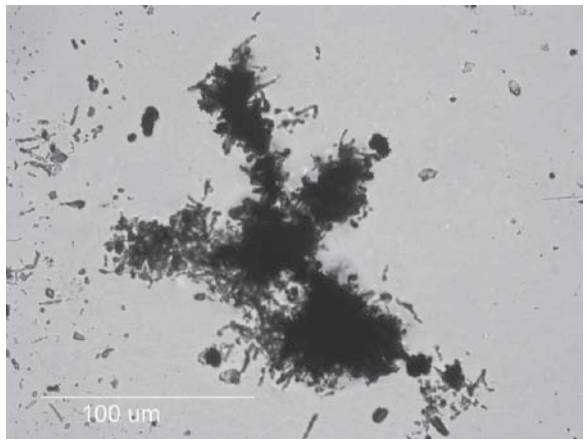


Figure 13.7

PLM image of the mold/biofilm material causing the darkening shown in Figure 13.6.

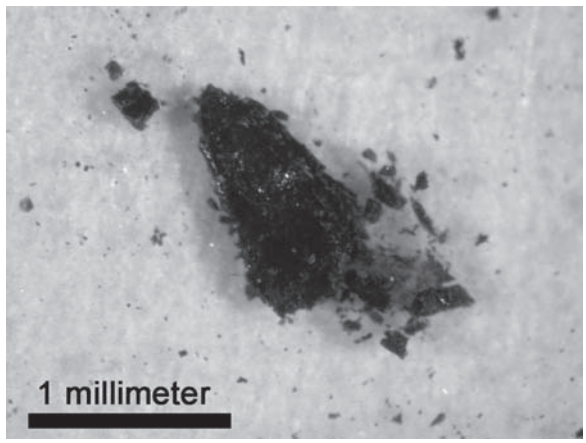


Figure 13.8

Brightfield reflected light image of black debris from parked automobile (Sooty Deposition 2). A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

scratched the automobile painted finish if rubbed. Samples were collected with tape and sent for microscopical examination. The black, sticky debris was examined using a combination of polarized light microscopy (PLM), scanning electron microscopy-energy dispersive x-ray spectrometry (SEM-EDS), and Fourier transform microspectroscopy (FTIR), and was found to consist primarily of metal particles, quartz sand, and a resinous material (see Figures 13.8, 13.9, and 13.10). Aggregates were extracted with drops of acetone on a glass slide and the acetone soluble fraction was redeposited on a reflective e-glass slide for FTIR analysis. Samples from exhaust vents of a nearby metal casting industry were collected. The residue from the exhaust vents was found to be similar in physical and elemental composition to the black debris appearing on the parked automobiles.

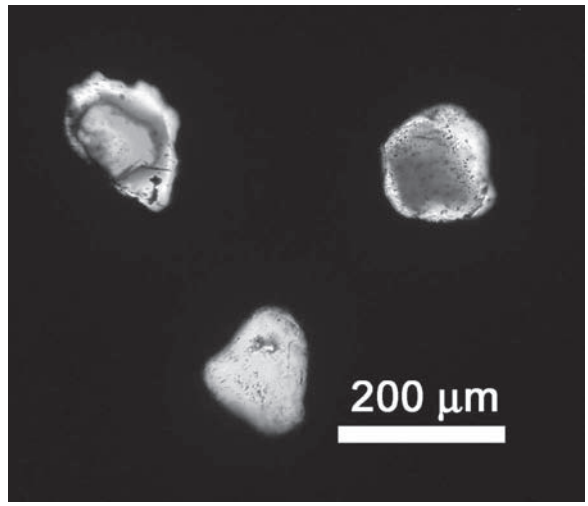


Figure 13.9
Transmitted light image using crossed polarizers of quartz particles in black debris.

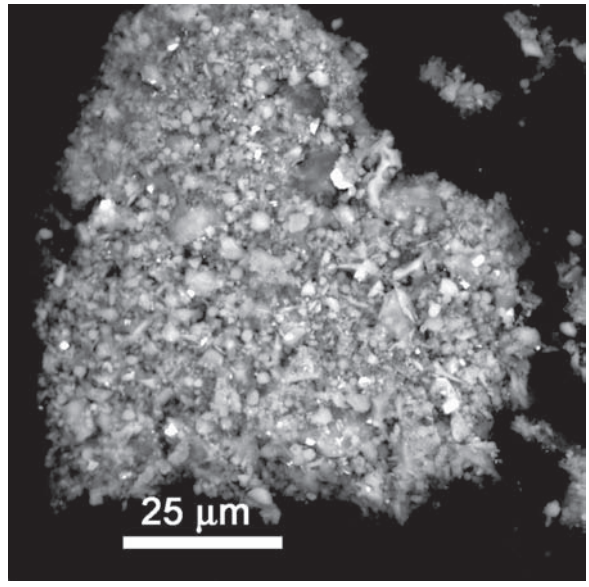


Figure 13.10
Backscattered electron image of inorganic particles comprising black debris.

13.5.5 WTC SIGNATURE SEARCH

The dust cloud generated by the collapse of the World Trade Center (WTC) towers on September 11, 2001 was a complex mixture of building material particles (Millette *et al.*, 2001; Lioy *et al.*, 2002a; Yin *et al.*, 2004). In April 2005, there remained questions about where the dust particles had settled and how to determine which residences in New York City still required cleaning of the WTC dust. EPA conducted a study to determine if a WTC Dust Signature Protocol could be used to determine where residue of the WTC dust remained.

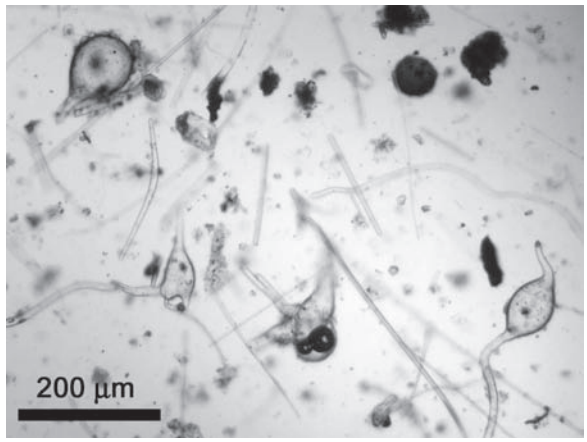


Figure 13.11
Brightfield transmitted light image of WTC dust showing glass fibers, cement particles, and gypsum.

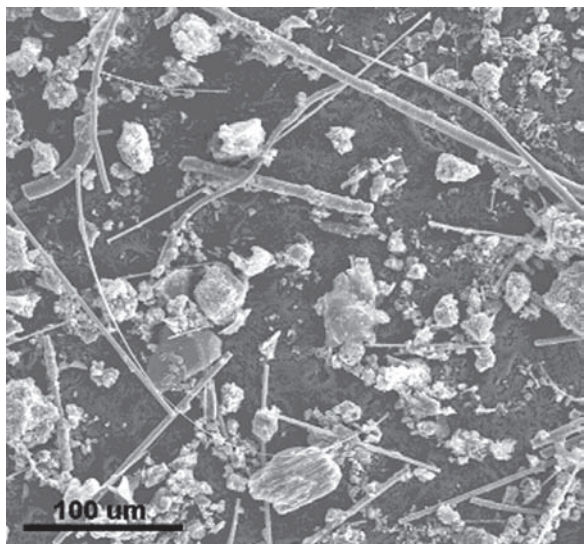


Figure 13.12
Secondary electron image of WTC dust showing glass fibers, cement particles, and gypsum.

They asked a number of independent laboratories to test a proposed protocol based on work done by the U.S. Geological Survey (Lowers *et al.*, 2005; Meeker *et al.*, 2005). Based on the fact that glass fibers, cement particles, and gypsum were the major components of the WTC dust (see Figures 13.11 and 13.12), the USGS protocol involved elemental mapping of calcium and sulfur by computer-controlled scanning electron microscopy—x-ray analysis (SEM-EDS) and an independent analysis for slag wool fibers. In the summer of 2005, based on the work of the various laboratories, the EPA decided that the Computer Controlled SEM/EDS analysis was not reproducible between labs and proposed a method based only on Slag Wool Analysis by SEM-EDS (USEPA, 2005). The proposed protocol was sent out to a Peer Review Panel. In October

2005, the peer review panel (Crane *et al.*, 2005) concluded, "EPA has not made the case that its proposed analytical method can reliably discriminate background dust from dust contaminated with WTC residue." The USEPA is proceeding with a $7 million Test and Clean Program that covers the area south of Canal Street and west of Pike-Allen Streets in New York City. Instead of a decision based on a WTC dust signature, a cleaning decision will be based on the finding of elevated levels of one of four contaminants of concern. Dust samples will be analyzed for asbestos, man-made vitreous fiber (MMVF), lead, and polycyclic aromatic hydrocarbons (PAH). The proposed EPA Criteria for determining whether an accessible area is to be cleaned are:

- Asbestos—5000 structures/cm^2
- MMVF—5000 fibers/cm^2
- Lead—40 $\mu g/ft^2$
- PAHs—150 $\mu g/M^2$

Standard methods exist for the analysis, in surface dust, of asbestos (ASTM, 2003b), lead (NIOSH, 1994b; USEPA, 1996; OSHA, 2002), and PAHs (USEPA, 1986), but no standard method currently exists for the analysis of MMVF in surface dust. It is anticipated that a method for MMVF in surface dust will be developed by combining the ASTM microvac TEM method for asbestos in dust (ASTM, 2003) with analytical procedures developed by the U.S. Geological Survey (Lowers and Meeker, 2005) using the scanning electron microscope to classify the glass fibers. A method using PLM identification of the glass fibers is also possible. There is no advantage to using the TEM for glass fiber analysis because although asbestos and other naturally occurring mineral fibers are characterized by their crystalline structure, which can be examined by electron diffraction, man-made vitreous fibers are amorphous; that is, they are glassy, with no discernible crystalline array (US Navy, 1997).

REFERENCES

ASTM. (2001a) Standard Practice for Determining Concentration of Airborne Single-Crystal Ceramic Whiskers in the Workplace Environment. *American Society for Testing and Materials, Method D6058-96.* Conshohocken, PA.

ASTM. (2001b) Standard Test Method for Determining Concentration of Airborne Single-Crystal Ceramic Whiskers in the Workplace Environment by Transmission Electron Microscopy. *American Society for Testing and Materials, Method D6056-96.* Conshohocken, PA.

ASTM. (2001c) Standard Test Method for Determining Concentration of Airborne Single-Crystal Ceramic Whiskers in the Workplace Environment by Scanning Electron Microscopy. *American Society for Testing and Materials, Method D6059-96.* Conshohocken, PA.

ASTM. (2001d) Standard Test Method for Determining Concentration of Airborne Single-Crystal Ceramic Whiskers in the Workplace Environment by Phase Contrast Microscopy. *American Society for Testing and Materials, Method D6057-96.* Conshohocken, PA.

ASTM. (2003a) Sampling and Testing of Possible Carbon Black Fugitive Emissions or Other Environmental Particulate, or Both. *American Society for Testing and Materials, Method D6602.* Conshohocken, PA.

ASTM. (2003b) Standard Test Method for Microvacuum Sampling and Indirect Analysis of Dust by Transmission Electron Microscopy for Asbestos Structure Number Surface Loading. *American Society for Testing and Materials, Method D5755-03.* Conshohocken, PA.

ATSDR. (2001) Libby Environmental Health Update. *Agency for Toxic Substances and Disease Registry*, January 12, 2001.

Clague, A.D.H., Donnet, J.B., Wang, T.K., and Peng, J.C.M. (1999) A Comparison of Diesel Engine Soot with Carbon Black. *Carbon* 37, 1553–1565.

Cotter-Howells, J. and Thornton, I. (1991) Sources and pathways of environmental lead to children in a Derbyshire mining village. *Environmental Geochemistry and Health* 13(2), 127–135. ISSN: 0269-4042 (Paper) 1573–2983 (Online). DOI: 10.1007/BF01734304.

Crane, D., Ehrenfeld, F., Gunter, M., McConnell, E., Su, S., Webber, J., and Wilhelmi, J. (2005) Peer Review of the U.S. Environmental Protection Agency's Final Report on the World Trade Center (WTC) Dust Screening Study. Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC., EPA Contract No. 68-C-02-060, Task Order 107.

Faegri, K., Iversen, J., and Waterbolk, H.T. (1964) *Textbook of Pollen Analysis.* Hafner Publishing Company, New York, 237.

Franck, U. and Herbarth, O. (2002) Using scanning electron microscopy for statistical characterization of the diameter and shape of airborne particles at an urban location. *Environmental Toxicology* 17, 98–104. Published online in Wiley InterScience (www.interscience.wiley.com); DOI 10.1002/tox.10037.

Ghazi, A.M. and Millette, J.R. (2005) Environmental forensic application of lead isotope ratio determination: A case study using laser ablation, Sector ICP-MS. *Environmental Forensics* 5(2), 97–108.

Graves, W.J. (1979) A mineralogical soil classification technique for the forensic scientist. *JFSCA* 24, 323–338.

Hawksworth, D.L., Kirk, P.M., Sutton, B.C., and Pegler, D.N. (1995) *Ainsworth and Bisby's Dictionary of the Fungi.* CABI International, Oxon, UK, 616.

Hopen, T.J., Millette, J.R., Boltin, W.R., and Brown, R.S. (1999) A dictionary of terms related to additives found in asbestos building products. *Microscope* 47(3), 163–171.

Hopen, T.J. (2004) The value of soil evidence. In *Trace Evidence Analysis*, M.M. Houck, ed. Elsevier Academic Press, Boston, 105–122.

Huffman, G.P., Huggins, F.N., Shah, N., Huggins, R., Linak, W.P., Miller, C., Pugmire, R.J., Meuzelaar, H.L.C., Seehra, M.S., and Manivannan, A. (2000) Characterization of fine particulate matter produced by combustion of residual fuel oil. *Journal of the Air & Waste Management Association* 50, 1106–1114.

Hunt, A., Johnson, D.L., and Thornton, I. (1991) Descriptive apportionment of lead in house dust by automated SEM. *Water, Air and Soil Pollution*, 57–77.

Hunt, A., Hawkins, J., Gilligan, E., and Bhatiab, S. (1998) A comparison of the lead particle content of indoor dust before and after a lead paint abatement: A new source of lead recontamination. *Indoor and Built Environment* 7, 32–46 (DOI: 10.1159/000024558).

ISO. (1999) Ambient air—determination of asbestos fibres—indirect-transfer Transmission electron microscopy method. International Standards Organization, Method 13794.

ISO. (2002) Ambient air-determination of numerical concentration of inorganic fibrous particles—Scanning electron microscopy method. International Standards Organization, Method 14966.

Jambers, W., De Bock, L., and Van Grieken, R. (1995) Recent advances in the analysis of individual environmental particles. A review. *The Analyst* 120, 681–692. DOI: 10.1039/AN9952000681.

Kostiuk, L.W., Johnson, M.R., and Thomas, G.P. (2004) University of Alberta flare research project final report November 1996—September 2004. University of Alberta, Department of Mechanical Engineering.

Krumbein, W.C. and Pettijohn, F.J. (1938) *Manual of Sedimentary Petrography.* Appleton-Century-Crofts, New York, 439–441.

Lioy, P.J., Weisel, C.P., Millette, J.R., Eisenreich, S., Vallero, D., Offenberg, J. *et al.* (2002a) Characterization of the dust/smoke aerosol that settled east of the

World Trade Center (WTC) in Lower Manhattan after the collapse of the WTC 11 September 2001. *Environmental Health Perspectives* 110(7), July 2002, 703–714.

Lowers, H.A., and Meeker, G.P. (2005) Particle atlas of World Trade Center dust. US Geological Survey, Open File Report 2005-1165.

Lowers, H.A., Meeker, G.P., and Brownfield, I.K. (2005) Analysis of background residential dust for World Trade Center signature components using scanning electron microscopy and x-ray microanalysis. US Geological Survey, Open File Report 2005-1073.

McCrone, W.C., Delly, J.G., and Palenik, S.J. (1973) The Particle Atlas, 2nd Ed. Ann Arbor Science Publishers, Inc., Ann Arbor, Michigan, ISBN: 0250400081.

Medalia, A.I. and Rivin, D. (1982) Particulate carbon and other components of soot and carbon black. *Carbon* 20, 481–492.

Meeker, G.P., Bern, A., Lowers, H., and Brownfield, I.K. (2005) Determination of a diagnostic signature for World Trade Center dust using scanning electron microscopy point counting techniques. US Geological Survey, Open File Report 2005-1031.

Millette, J.R., Brown, R.S., Barnett, J., and Mount, M.D. (1991) Scanning electron microscopy of post-it notes used for environmental sampling. *NAC Journal,* 32–35.

Millette, J.R. and Hays, S.M. (1994) *Settled Asbestos Dust: Sampling and Analysis*. Lewis Publishers, Boca Raton.

Millette, J.R., Brown, R.S., and Mount, M.D. (1995) Lead arsenate. *Microscope* 43(4), 187–191.

Millette, J.R. and Few, P. (2001) Sample collection procedures for microscopical examination of particulate surface contaminants. *Microscope* 49(1), 21–27.

Millette, J.R., Boltin, R., Few, P., and Turner, W. Jr. (2002) Microscopical studies of World Trade Center disaster dust particles. *Microscope* 50(1), 29–35.

Millette, J.R. and Brown, R.S. (2002) Examining the fine particle component of grout dust by automated SEM. *Microscope* 50(1), 1–3.

Millette, J.R., Lioy, P.J., Wietfeldt, J., Hopen, T.J., Gipp, M., Padden, T. *et al.* (2003) A microscopical study of the general composition of household dirt. *Microscope* 51(4), 201–207.

Millette, J.R. and Bandli, B.R. (2005) Asbestos identification using available standard methods. *Microscope* 53(4), 179–185.

Millette, J.R. (2006) Chapter 2: Asbestos analysis methods. In *Asbestos: Risk Assessment, Epidemiology, and Health Effects*, R.F. Dodson and S.P. Hammar, eds. CRC, Taylor & Francis, Boca Raton, 9–38.

Ness, S.A. (1994) *Surface and Dermal Monitoring for Toxic Exposures*. Van Nostrand Reinhold, New York, 561.

NIOSH. (1994a) Asbestos and other fibers by phase contrast microscopy (PCM). *NIOSH Manual of Analytical Methods*, Method 7400, 4e, 94–113.

NIOSH. (1994b) Lead surface wipe samples. *NIOSH Manual of Analytical Methods*, Method 9100, 4e.

OSHA. (2002) Metal and metalloid particles in workplace atmospheres (ICP Analysis). Occupational Safety and Health Administration, Method ID-125G.

Palenik, S.J. (1979) The determination of geographical origin of dust samples. In *The Particle Atlas*, 2e, Volume 5, W.C. McCrone, J.G. Delly, and S.J. Palenick, eds. Ann Arbor Science Publishers Inc., Ann Arbor, 1347–1361.

Petraco, N. and Kubic, T. (2003) *Color Atlas and Manual of Microscopy for Criminalists, Chemists, and Conservators.* CRC Press, New York, 328.

Pirrie, D., Power, M.R., Rollinson, G., Camm, G.S., Hughes, S.H., Butcher, A.R., and Hughes, P. (2003) The spatial distribution and source of arsenic, copper, tin and zinc within the surface sediments of the Fal Estuary, Cornwall, UK. *Sedimentology* 50(3), 579–595.

Rybicka, E.H., Wilson, M.J., and McHardy, W.J. (1994) Chemical and mineralogical forms and mobilization of copper and lead in soils from a Cu-smelting area in Poland. *Journal of Environmental Science and Health, Part A.* A29, 3, 531–546.

Schneider, T., Nielsen, O., Bredsdorff, P., and Linde, P. (1990) Dust in buildings with man-made mineral fiber ceiling boards. *Scandinavian Journal of Work Environment and Health* 16, 434–439.

Thornton, I., Watt, J.M., Davies, D.J.A., Hunt, A., Cotter-Howells, J., and Johnson, D.L. (1994) Lead contamination of UK dusts and soils and implications for childhood exposure: An overview of the work of the Environmental Geochemistry Research Group, Imperial College, London, England 1981–1992. *Environmental Geochemistry and Health* 16, 113–122; 10.1007/BF01747907.

Turner, W.L., Millette, J.R., Boltin, W.R., and Hopen, T.J. (2005) A standard approach to the characterization of common indoor dust constituents. *Microscope* 53(4), 169–177.

USEPA. (1986) Polynuclear Aromatic Hydrocarbons. U.S. Environmental Protection Agency, Method 8310.

USEPA. (2004) Research Method for Sampling and Analysis of Fibrous Amphibole in Vermiculite Attic Insulation. U.S. Environmental Protection Agency, EPA/600/R-04/004.

USEPA. (2005) Final Report on the World Trade Center (WTC) Dust Screening Method Study. U.S. Environmental Protection Agency, August 17, 2005.

US Navy Environmental Health Center. (1997) Man-made vitreous fibers. Technical Manual, NEHIC-TM6290.91-REVA.

Van Orden, D. (2006) Asbestos. In *Environmental Forensics: Contaminant Specific Guide*, R.D. Morrison and B.L. Murphy, eds. Elsevier, Amsterdam, 19–31.

Yiin, L.M., Millette, J.R., Vette, A., Ilacqua, V., Quan, C., Gorczynski, J. *et al.* (2004) Comparisons of the dust/smoke particulate that settled inside the surrounding buildings and outside on the streets of southern New York City after the collapse of the World Trade Center, September 11, 2001. *Journal of the Air & Waste Management Association* 54, 515–528.

CHAPTER 14

APPLICATIONS OF LASER ABLATION INDUCTIVELY COUPLED PLASMA MASS SPECTROMETRY (LA-ICP-MS) IN ENVIRONMENTAL FORENSIC STUDIES

A. Mohamad Ghazi

14.1	Introduction	637
14.2	Magnetic Sector ICP-MS	639
14.3	Laser Ablation	641
14.3.1	Operation of Laser Ablation	641
14.3.2	Sampling Strategy	643
14.3.3	Rastering Analysis	646
14.3.4	Calibration Strategies for Quantitative Analysis	649
14.4	Applications of Laser Ablation ICPMS in Environmental Studies	650
14.4.1	Dendrochemical Studies	651
14.4.2	Fingerprinting the Source of Lead in the Environment	652
14.4.3	Base Paint	653
14.4.4	Lead in Contaminated Soil	655
14.5	Environmental Impacts to Metal Exposure in Tissue	657
14.6	Environmental Formation of Diamonds	660
14.7	Summary	662
References		662

14.1 INTRODUCTION

Analytical instrumentation and the tools available to environmental scientists (e.g., geochemist, forensic chemists) have become increasingly sophisticated, due in part to the large volume of data sets evaluated in environmental investigations. As a result of the most recent development in micro-analytical techniques coupled with applications of Inductively Coupled Plasma Mass Spectrometry (ICP-MS) during the past 20 years, forensic scientists have been able to extract an overwhelming wealth of data from single samples. These techniques provide routine bulk analysis as well as detailed major- and trace-element determinations, as well as isotopic analysis of individual

components from a wide range of material. Although ICP-MS was developed initially by geochemists, other types of materials have been successfully analyzed by laser ablation ICP-MS.

Inductively coupled plasma (ICP) is generated by applying a radio frequency around a copper coil to turn a stream of an inert gas such as argon into high-energy plasma. The generation of plasma is accomplished by subjecting a few argon electrons to intense oscillations by the radio-frequency magnetic fields. Collisions that occur between the free electrons and atoms of argon gas cause further ionization. Each particle formed in the argon gas is subjected to a magnetic field that causes it to travel in annular paths within the torch, where the temperature could rise up to 10,000 K (e.g., Lahaye, 2002).

This technique was commercially introduced in 1983 and since its introduction, among the high-end analytical instrumentation, inductively coupled plasma instruments (ICPs) arguably have seen the greatest attention in instrument development and application. This is particularly true since ICP has been coupled with mass spectrometry, and now ICP-MS has matured into one of the most sophisticated and successful methods in atomic spectrometry such as atomic absorption and optical emission spectrometry, including ICP Atomic Emission Spectroscopy (ICP-AES) (e.g., Houk *et al.*, 1980; Date and Gray, 1981, 1983). This is due to a number of advantages this technique offers over all other atomic spectrometry methods, such as a superior sensitivity, lower detection limits, and the ability to make multielement and isotopic ratio measurements with sufficient precisions that are used for age determination of various rocks and minerals (Date and Gray, 1989; Hinners *et al.*, 1989; Russ, 1989; Ketterer *et al.*, 1991; Jarvis and Jarvis, 1992; Jarvis *et al.*, 1992; Koeppenaal, 1992; Rienhard and Ghazi, 1992; Ghazi, 1994; Monna *et al.*, 1997; Chaudhary-Webb *et al.*, 1999, 2003; Jackson *et al.*, 2001; Machado and Simonetti, 2001).

An important advantage of ICP-MS techniques is the versatility in sample introduction. Regardless of the type of introduction method (i.e., solution, solid, slurry), all samples are introduced at atmospheric condition. This provides ICP-MS the potential of being a true application-specific instrument with perhaps the widest range of inorganic analytical applications; there is a vast literature on various aspect of this technique (e.g., Gray and Jarvis, 1989; Gray *et al.*, 1992; Hieftje, 1992; Montaser and Golightly, 1992; Ghazi *et al.*, 1993, 1996; Longerich *et al.*, 1993; Sargent and Webb, 1993; Newman, 1996; Montaser, 1998). Other advantages of ICP-MS over comparable chemical analytical instruments such as atomic absorption, graphite furnace AA, and optical emission spectrometry, such as ICP Atomic Emission Spectroscopy (ICP-AES) include detection limits, which for most elements are equal to or better than those obtained by Graphite Furnace Atomic Absorption Spectros-

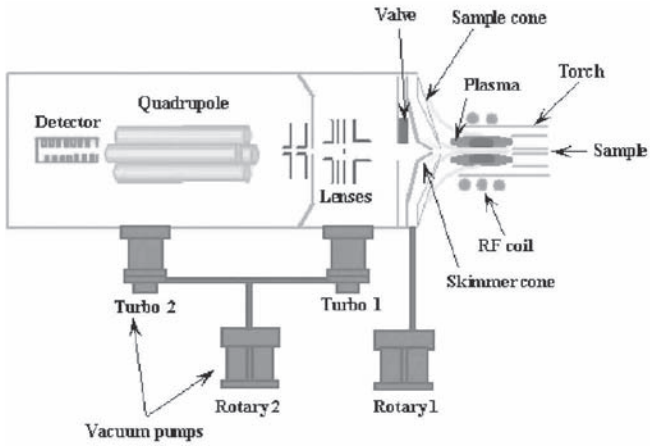


Figure 14.1
Schematic representation of a quadrupole ICP-MS (PlasmaQuad 3). Source: http://www.nature.cuni.cz/igmnz/icplab0.html

copy (GFAAS); higher throughput than GFAAS; and the ability to handle both simple and complex matrices with a minimum of matrix interferences due to the high-temperature of the ICP source.

Traditional ICP-MS instruments are based on using quadrupole mass spectrometer (Q-ICP-MS), which provides a resolution of one atomic mass unit. A quadrupole mass filter consists of four parallel metal rods. By varying the electric signals to a quadrupole it is possible to transmit only ions within a very small m/z ratio, where m is the atomic mass and z is the number of charge (other ions are neutralized and carried away as uncharged particles) (see Figure 14.1).

14.2 MAGNETIC SECTOR ICP-MS

Magnetic sector ICP-MS (also known as high resolution ICP-MS), and single and multicollector ICP-MS are the most recent instrumental developments in analytical inorganic chemistry for elemental and isotope ratio analysis. Both of these instruments are double focusing sector magnetic ICP-MSs in reverse Nier-Johnson geometry (e.g., Walder *et al.*, 1993a, 1993b; Walder and Furuta, 1993; Shuttleworth and Kremser, 1996; Becker and Dietze, 1997; Jakubowski *et al.*, 1998; Becker, 2002) (see Figure 14.2). The major difference between these two instruments is that one is equipped (or has the option) with a multicollector detector system for simultaneous detection of isotopes for more precise than isotope ratio analysis.

The operational philosophy of sector instruments is significantly different to that of quadrupole instruments. Ions are resolved by double focusing with the magnetic sector and the electrostatic analyzer (ESA). When the slit apertures at the entrance and exit of the mass spectrometer are small, then the

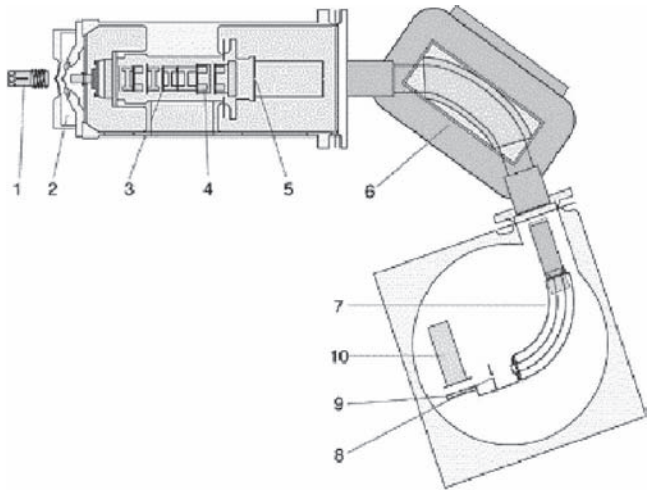


Figure 14.2
Sketch of a double focusing single collector High Resolution ICPMS (HR-ICP-MS). The components are as follows: (1) ICP ion source; (2) interface region, including sampler and skimmer cones; (3) transfer and focusing optics; (4) acceleration and beam focusing; (5) entrance slit; (6) electromagnet; (7) electric sensor; (8) exit slit; (9) conversion dynode; (10) electron multiplier (Courtesy of ThermoFinnigan).

double focusing of ions results in high resolution. Three fixed resolutions—300, 3000, and 10,000; low, medium, and high, respectively (dependent on the slit aperture widths)—are possible with the instrument by ThermoQuest (Finnigan MAT). The multicollector ICPMS instruments such as Neptune (by ThermoFinnigan) and Nu Plasma™ (by Nu instrument) offer even greater analytical capability by allowing TIMS-level precision on simultaneous isotopic measurements of up to eight isotopes.

These instruments scan over different masses by changing the field strength of the magnetic sector or by changing the applied accelerating voltage. Three models of data acquisition are available: (1) magnet scanning, (2) acceleration voltage scanning, and (3) a combination of the two methods. Magnet scanning is a relatively slow way of scanning the mass range because each magnet scan requires a settling time (1–300 ms) to reach stability with the magnetic field strength of the magnet. This mode of operation is also of limited use with laser ablation (LA) acquisitions because temporal differences in the laser signal can result in poor accuracy and precision when settling times are long. In contrast, acceleration voltage scanning is a relatively fast way of scanning the mass range. Jumps from peak to peak can be achieved in less than one millisecond but unfortunately, this method is limited to only 30% of the mass range relative to the magnet park mass. Therefore, this method of acquisition is especially suited to analysis of laser signals, but is limited to the collection of subsets of elements. When high and low mass elements are to be analyzed by laser ablation in one acquisition, then a combination of magnet and acceleration voltage scanning is applied. Wide analyzer slit apertures produce peaks that have large flat tops enabling scanning of only a small portion of the peak and thus reducing the magnet settling time (1–10 ms) without compromising analytical performance. In this way the instrument can quickly

peak hop across the mass range with a series of magnet and acceleration voltage jumps and acceleration voltage scans.

Nevertheless, in routine elemental and isotopic analysis, the Q-ICP-MS instruments are by far the most commonly used and the work horse for multielemental analysis in most of the commercial analytical laboratories.

14.3 LASER ABLATION

Another breakthrough associated with ICP mass spectrometry is the development and the use of lasers as micro-sampling mechanisms for high sensitivity ICP-MSs (Jeffries *et al.*, 1989; Pearce *et al.*, 1992; Neal *et al.*, 1995; Perkins and Pearce, 1995; Shuttleworth, 1996; Shuttleworth and Kemser, 1997; Günther *et al.*, 1998) (see Figure 14.3). Among the spectrum of *in situ* micro-analytical methods, laser ablation ICP-MS analyses is increasingly recognized for its strengths and versatility, including major to trace element analyses, isotopic analyses, application to a wide variety of matrices (e.g., metals, oxides, organics, and fluids), short analysis time, and low costs. The use of LA-ICP-MS method has increased as unique applications are developed, particularly for UV laser ablation. Sylvester (2001) reviewed applications of LA-ICP-MS in the earth sciences with a chapter on the application of LA-ICP-MS in biologically related environmental studies as well as a number of chapters on other technical aspects of this technique.

14.3.1 OPERATION OF LASER ABLATION

Laser ablation technique uses high-energy photons generated in a laser (Light Amplification by Stimulated Emission of Radiation) to liberate fine particles

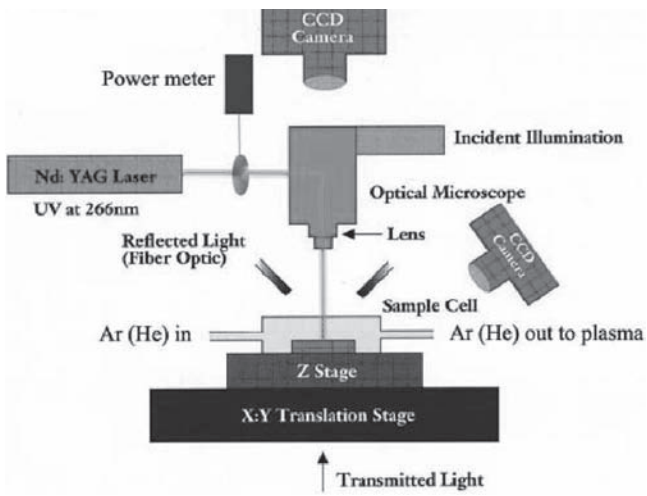


Figure 14.3
General schematic of a Nd-YAG-UV laser ablation system (modified VG Microprobe I). Source: http://www.nature.cuni.cz/igmnz/icplab0.html

(aerosol) from the solid sample. There are two main types of ultraviolet lasers widely used in earth, environmental, and forensic sciences. In solid lasers, wavelength (λ) of the laser light emitted using a Nd : YAG laser is generally at 1064 nanometer (nm) in the infrared range. This wavelength couples easily with samples containing significant quantities of the transition elements. Longerich *et al.* (1993) incorporated a harmonic generator into the laser apparatus, which allowed shorter wavelength (532 nm, 266 nm, and 213 nm) laser radiation to be generated. Most commercially available lasers are 266 nm and 213 nm instruments. The second type of lasers are gas-based, and include the excimer lasers (e.g., Kr-F, ArF, and F2, ($\lambda = 193$ nm)). In excimer gas lasers, light is emitted by a short-lived molecule made up of one rare gas atom (Bea *et al.*, 1996; Ducreux-Zappa and Mermet, 1996; Balhaus *et al.*, 1997; Gahderi *et al.,* 1997; Günther *et al.*, 1997; Kinsley *et al.*, 1997; Sinclair and McCulloch, 1997; Sylvester, 1997; Sylvester and Ghaderi, 1997; Eggins *et al.*, 1998). Its operation depends on electronic transitions in molecules.

In this technique the energy of photons is converted into thermal energy, which is responsible for the ablation and vaporization of the exposed sample surface. Laser ablation results in damage of the sample surface and in the formation of a laser pit that is up to a few hundred microns in diameter and ca. 100–200 microns deep. The ablated material is carried by the carrier gas (argon or helium) to the inductively coupled plasma source of the mass spectrometer in the form of microscopic solid particulates or the fine aerosol from the solution (e.g., fluid inclusions) that enters the RF-generated argon plasma (temperatures up to 10,000 K) where virtually all atoms and molecules are converted to ions before they enter the vacuum stage of the mass spectrometer through a small entrance orifice in the interface (see Figures 14.1 to 14.3).

The laser is operated in two modes: (a) Q-Switched, where a short laser pulse (10 ns) contains practically all of the energy; and (b) Fixed-Q or Free-Running, where the laser pulse is much longer (120–150 sec) and the power delivered is considerably less (Denoyer *et al.*, 1991). The resulting ablation characteristics are very different and produce very different ablation pits, thus affecting the size of the sample analyzed. In Q-switched mode, the laser energy is higher, and much of the ablation occurs through total vaporization and mechanical ablation. In Fixed-Q or Free-Running mode, the power of the laser is lower; the laser interacts with the sample for a longer period of time and produces a deeper ablation crater of smaller diameter relative to Q-Switched mode. Generally, most lasers are operated in Q-Switched mode, and the beam profile is typically Gaussian. The laser beam is focused onto the surface of the sample, which is generally housed inside an ablation cell. The aerosols (solid, gas, or even liquid for some samples) generated during the ablation process are swept into the argon plasma of the ICP via an argon gas stream.

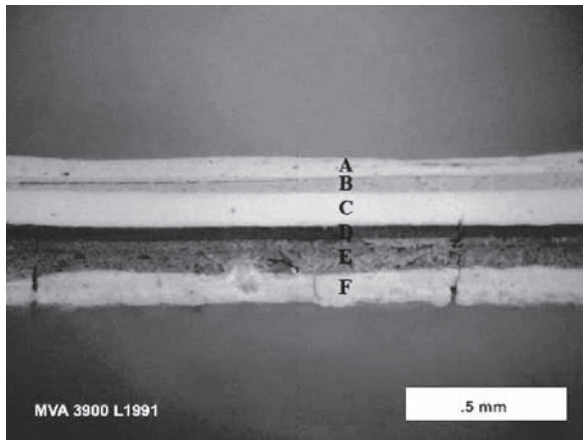


Figure 14.4
Light microscope photomicrograph of cross-section of layered paint sample, standing up, showing six visible layers of different paint. Alphabets are used to identify each layer and correspond with text. The sample was prepared for analysis by scanning electron microscopy and laser ablation ICP-MS, by cutting a section of paint chip approximately 10 mm wide with a scalpel (Ghazi and Millette, 2004). A full color version of this figure is available at books.elsevier.com/companions/9780123695222.

14.3.2 SAMPLING STRATEGY

Two approaches are available for sampling materials by laser ablation ICPMS. The first is performed at a fixed point by applying a series of laser pulses at a specified energy level. With this method, each successive pulse removes a layer of material and a crater is drilled into the sample with an opening at the surface of the sample, which remains almost constant in size or cross section during the ablation process. In this type of sampling the laser ablation stage is fixed at one spot, thus this type of sampling strategy is called *spot sampling*. This method of sampling provides elemental variation within the thickness of the sample. Figure 14.4 depicts a layered lead-bearing paint sample, which was the subject of a paper by Ghazi and Millette (2004). Figure 14.5a is a scanning electron microscope (SEM) photo of the same paint sample after spot analysis using a Nd:YAG 213 nm laser ablation system. Figure 14.5b shows six separate peaks, each representing all four lead isotopes, which were obtained from each hole. Of note is that the height of each curve corresponds to relative natural abundance of each lead isotope (i.e., ^{204}Pb (1.4%) < ^{207}Pb (22.1%) < ^{206}Pb (24.1%) < ^{208}Pb (52.4%). Figure 14.6 shows the cross-sectional cuts of one of the ablation holes in this lead-bearing paint sample (Ghazi and Millette, 2004, 2006). An examination of the ablation profiles showed that the holes were slightly curved; such minute deformation is in part due to the interaction of laser beam and the paint, which is relatively soft. All analyses were performed under identical laser conditions (e.g., beam diameter, laser intensity, energy flux, ablation time, and distance between ablation pits).

Another example of spot analysis with laser ablation is in the area of inclusion studies (i.e., melt or fluid). Fluid inclusions are basically microscopic entrapped fluids within host minerals (e.g., quartz, halite, other types of geological samples, as well as synthetic materials) that carry the physiochemical

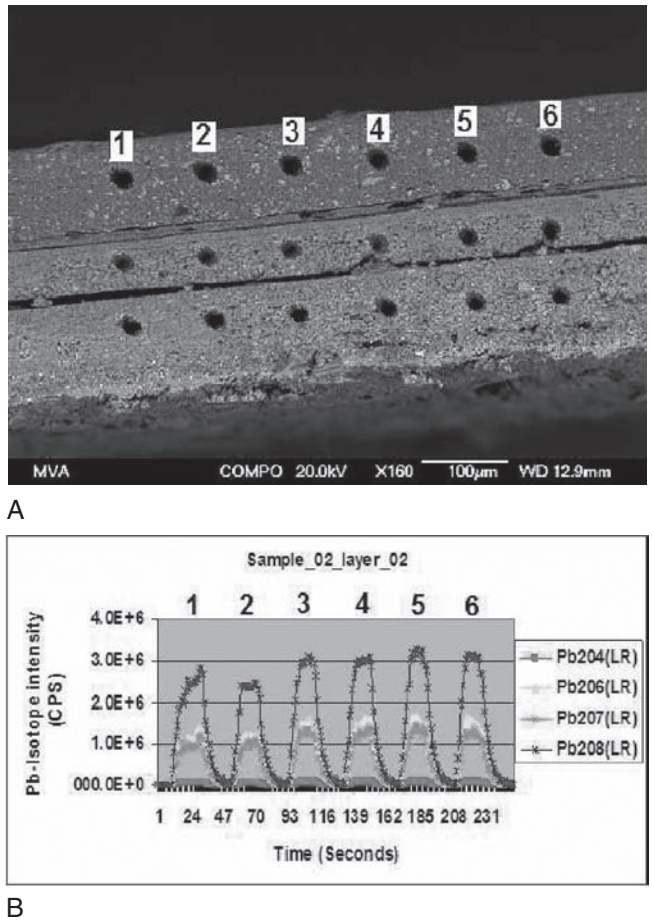


Figure 14.5

***A**. Scanning electron microscope (SEM) photomicrograph of several equally spaced laser ablation pits along straight lines in the center of each layer of paint. Ablation holes are 20 µm in diameter and the distance between holes is 100 µm. **B**. Time-resolved Pb isotopic profile of distribution of 204Pb, 206Pb, 207Pb, 208Pb in layer (1) of the paint sample. From left to right each peak corresponds to holes 1–6. Isotopes in order of intensity from largest to smallest are Pb208, Pb206, Pb207, and Pb204. A full color version of this figure is available at books.elsevier.com/companions/9780123695222.*

signature of the environment at the time of formation of the substance. Figure 14.7a is an example of a fluid inclusion-bearing synthetic quartz. In this type of analysis, the objective is to use the laser beam to drill through the host mineral to reach the inclusion and subsequently extract the fluid content of the inclusion. Figure 14.7b depicts an empty fluid inclusion whose content has been completely extracted as carried to the interface region of the ICP-MS and analyzed in the same manner and solid particles. Figure 14.8 is a time-resolved spectrum showing various stages of instrument response to ablation of a fluid inclusion-bearing sample. In terms of calibration technique, the initial work in this area was done at the British Geological Survey, where a set of homemade standards was made based on encapsulation of micro-droplets of standard solution in hydrophobic epoxy resin (Sheppherd and Chenery, 1995; Moissette *et al.*, 1996). Ghazi and coworkers (Ghazi *et al.*, 1996; McCandless *et al.*, 1997; Ghazi and Shuttleworth, 2000) from Georgia State

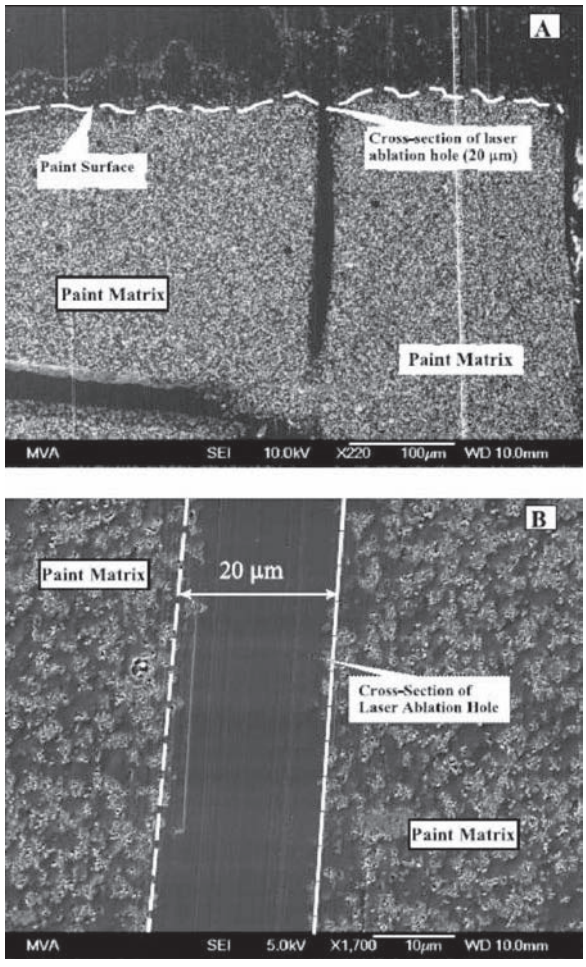


Figure 14.6

***A**. Scanning electron microscope (SEM) photomicrograph showing the depth of a laser ablation hole in cross-sectional view. As a result of interaction of laser beam with paint, which is relatively soft, it appeared that holes were slightly curved, thus the actual depth (~500 µm) of holes was greater than what it is shown in this diagram.* ***B**. Close-up SEM photomicrograph of the above ablation pit (20 µm diameter). Also note the porous nature of the paint matrix (Ghazi and Millette, 2004).*

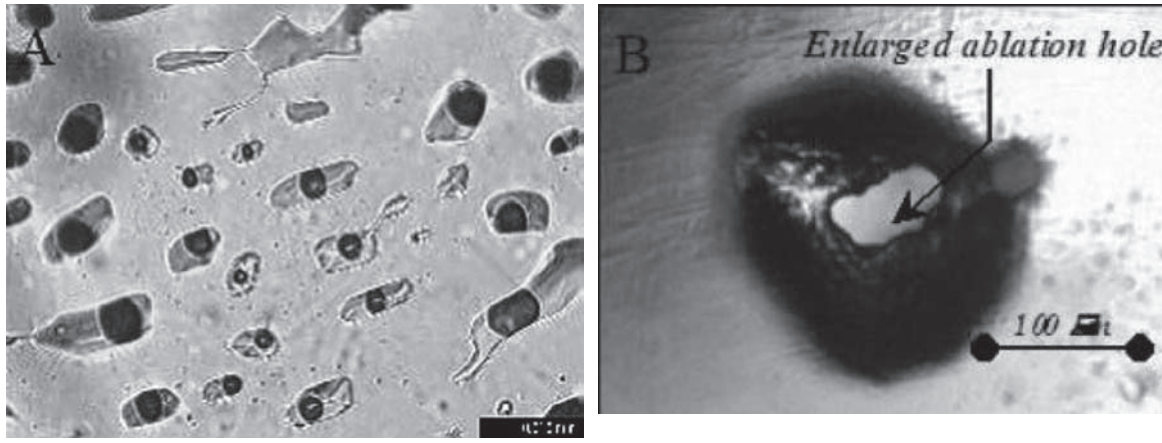


Figure 14.7

***A**. Photomicrograph of a two-phase (liquid+vapor) synthetic fluid inclusion in quartz host mineral.* ***B**. Photomicrograph of the emptied fluid inclusion showing irregular and enlarged opening of the laser hole.*

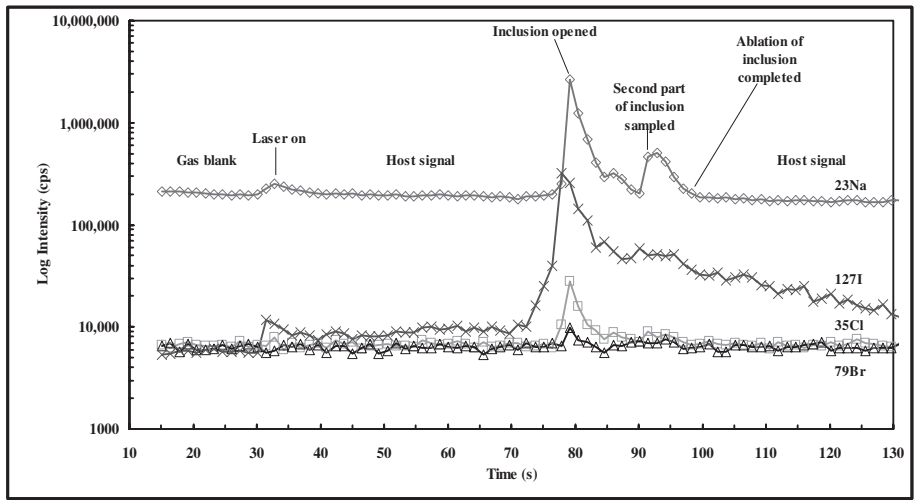


Figure 14.8
Time-resolved spectra from LA-ICP-MS analysis of a synthetic fluid inclusion with a Cl/Br weight ratio of 50. This inclusion was located at about 20–30 μm below the upper surface of the polished quartz host mineral, measured 30 μm × 50 μm and had two chambers. The large peak from 76 seconds to 85 seconds were from the main chamber, and the second peaks resulted from draining of the second chamber. Peaks for each isotope are offset because of minor variations in the time at which its respective mass window is scanned.

University developed a calibration technique using solutions sealed in microcapillary tubes. Günther *et al.* developed a quantification procedure using elements concentrations ratios calculated by referencing intensity integrals against signals from external standards (glass and solution) and an elemental internal standard (e.g., ^{23}Na) (Günther *et al.*, 1997a,b, 1998; Audétat *et al.*, 1998; Audétat and Günther, 1999; Heinrich *et al.*, 1999).

The application of laser ablation as an alternative mode of sample introduction is an established and proven technique for the analysis of solid samples without dissolution. However, perhaps the most challenging area in LA-ICP-MS technology has been in developing calibration procedures for routine analysis of nonsolid materials, namely fluid analysis. This is largely because of the lack of matrix matching since liquid matrix interacts differently with the laser beam. This is particularly important in single fluid inclusion analysis where the unknown matrix is in a liquid state, and most commonly used standards are in solid form (e.g., NIST glasses).

14.3.3 RASTERING ANALYSIS

In rastering analysis, the laser ablation sampling motorized chamber moves at a preset rate, usually microns per second; thus the laser beam produces a line of a fixed thickness equal to the radius of the beam. This type of sampling is called *rastering*. This analysis produces a shallow "trench" rather than a "pit." For example, Ghazi *et al.* (2000a,b, 2002a,b) in separate experiments used the rastering technique to study the rate of absorption and spatial distribution of various metals (Ni, Ag, and Ga) absorbed by tissue and teeth upon prolonged exposure to these metals. Figure 14.9 is a cross-section of a

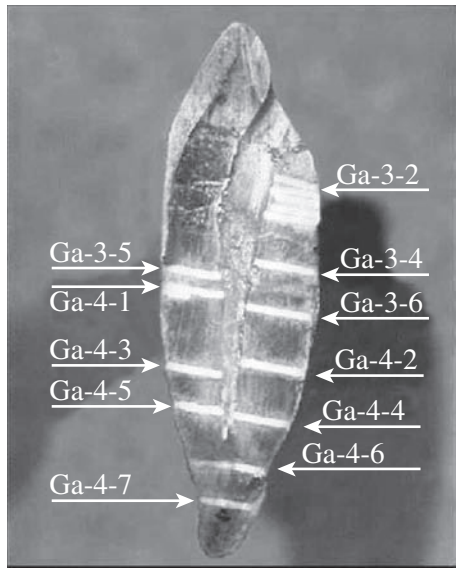


Figure 14.9

A photograph of a cross-section of gallium-treated human tooth showing the central root canal and the location of the rastering ablation tracks.

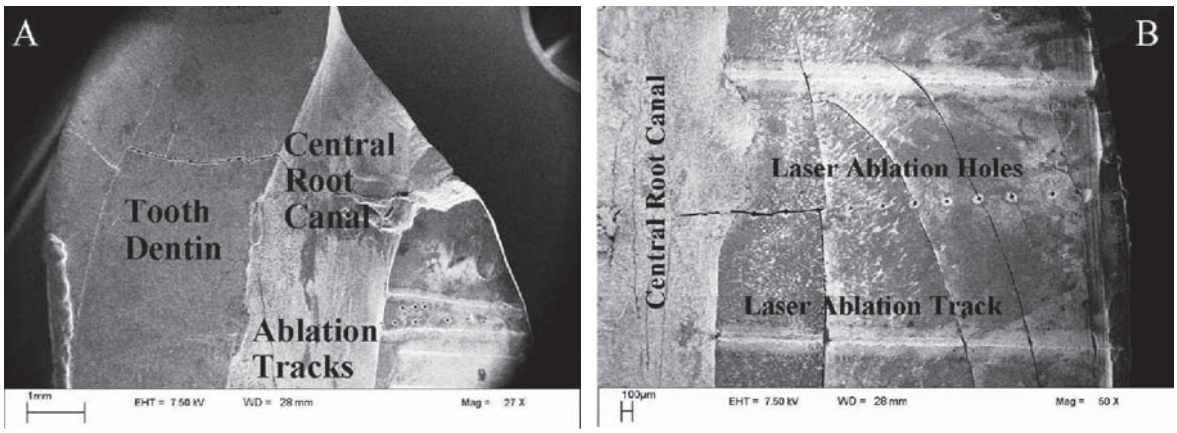


Figure 14.10

***A**. Scanning electron microscope (SEM) photomicrograph of the upper portion (cervical-third) of a gallium-treated tooth, showing the opening of the central canal in the middle and three laser ablation rastering tracks and individual laser ablation holes on the right. **B**. A close-up SEM photomicrograph showing two laser ablation rastering tracks above and below several individual laser ablation holes in the middle portion of the same gallium-treated tooth. The outer edge of the root on the right side of the micrograph is covered with a thin layer of cementum.*

human molar showing the central root canal and the location of the rastering ablation tracks. In this study a gallium solution was placed in the central canal of the tooth during a root canal operation and the objective was to determine the rate of diffusion of gallium into dentin (Ghazi *et al.*, 2000a,b). Raster sampling provides valuable information on spatial distribution of elemental concentration (see Figures 14.10 and 14.11). This method is an important

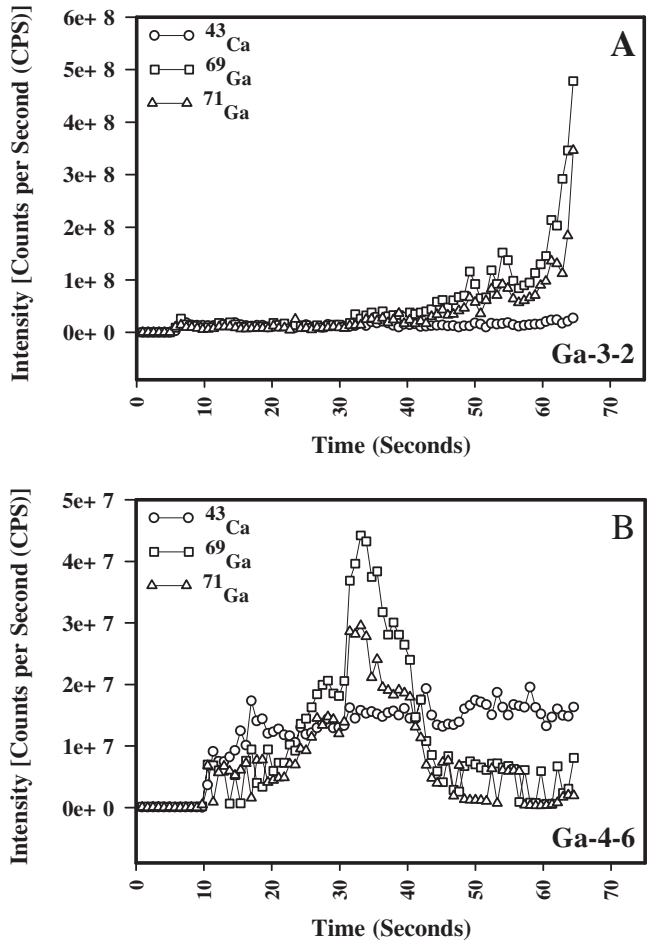


Figure 14.11

Profiles of the distribution of 43Ca, 69Ga, and 71Ga (cps) for two laser ablation rastering tracks from two different locations of the tooth. ***A****. Track Ga-3-2 is located in the upper portion of the sample.* ***B****. Track Ga-4-6 is located in the areas directly below the tip of the central canal, see Figure 14.9. Note that in both Figures 14.9A and 14.9B, rastering starts at time zero at the right egde of the sample and continues across the sample; however, in Figure 14.11A, rastering stops at the right edge of central canal, which expectedly corresponds to the highest absorbed concentration of gallium. Similarly, as shown in Figure 14.11B, the middle portion (30–40 second zone) is the area's closest zone to the central canal, thus shows the highest absorbed concentration. Also, note that that 43Ca was selected as the internal for quantification. The intensity for this isotope is relatively constant (1.4E7–2.0E7), which corresponds to uniform concentration of calcium in dentin, which is approximately 27%. As the movement of the specimen beneath the laser occurred at a constant rate (0.04 mm/sec), the time in seconds is also distance (i.e., 10 seconds = 0.4 mm) in both figures.*

tool for surface and shallow depth profiling and elemental mapping for comparative studies. Mason and Mank (2001) provide a detailed account on different sampling mechanisms with emphasis on depth analysis.

14.3.4 CALIBRATION STRATEGIES FOR QUANTITATIVE ANALYSIS

Calibration strategy for quantitative analysis typically is achieved by using a well-established solid standard reference material and comparing the response for a suitable element in a reference standard material and the unknown. If the counts per second measurement in both the standard and the unknown are the same for a selected element (e.g., Pb), then it can be assumed that the unknown contains the same concentration as the standard. However, typically this is not the case, and in order to perform a reliable quantitative analysis, three important parameters must be considered:

- The standard is well-homogenized with well-distributed composition of the analyte
- Matrix matching between standard and sample is as close as possible
- Selection of a correct internal standard element that exists in both standard and sample is required

Once these parameters have been controlled a complete quantitative procedure would include the following steps:

(1) Analysis of blank gas.
(2) Analysis of the reference material.
(3) Analysis of the unknown.
(4) Second analysis of the reference material in order to monitor the drift (sensitivity variation of the instrument through time (e.g., Figure 14.12).

The ablated materials will be transported by the carrier gas to the interface region, and the number of ions that are produced in the plasma, and subsequently reach the detectors, depends on

- The atomic proportions of the element in the source material
- The amount of material removed during ablation
- The ionization potential of the element in the plasma
- Isotopic abundance

The results obtained from laser ablation are always in the form of signal intensities and expressed as counts per seconds (cps). The equation

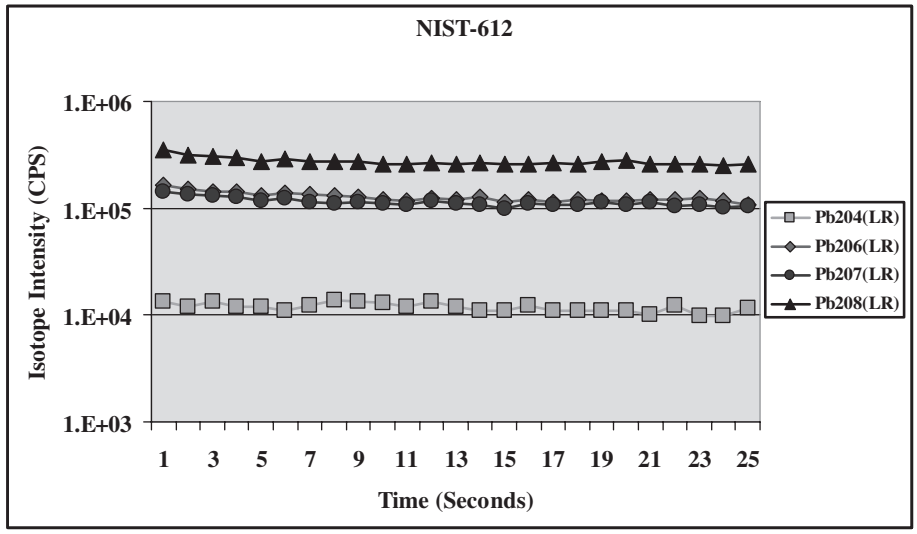


Figure 14.12
A 25-second time-resolved spectrum in counts per seconds (cps) for laser ablation sector field ICP-MS analysis for four isotope of lead (204Pb, 206Pb, 207Pb, 208Pb) in NIST-SRM-612 glass standard.

for calculating the concentration of an element in the sample accordingly is:

$$(C_E)_U = \{[(C_E)_{Ref}]*[(I_E)/(I_{IS})]_U\}\{[(I_E)/(I_{IS})_{Ref}]*[(C_{IS})_U/(C_{IS})_{Ref}]\} \quad (14.1)$$

where:

C = concentration (e.g., ppm, %)
I = intensity (cps)
E = element analyzed (the analyte)
IS = internal standard element
U = unknown
Ref = reference value (e.g., NIST standards, determined by electron microprobe (EMP) or secondary ion mass spectrometry (SIMS), or any other method).

14.4 APPLICATIONS OF LASER ABLATION ICPMS IN ENVIRONMENTAL STUDIES

Laser ablation ICP-MS instruments have produced trace element and isotopic ratio analyses of solid samples in a wide range of sample medium, including:

- Minerals and glasses (e.g., Arrowsmith, 1987; Jarvis *et al.*, 1992; Perkins *et al.*, 1992; Pearce *et al.*, 1992, 1997; Lichte, 1995; Günther *et al.*, 1997a)
- Soft tissues and dental materials (Ghazi *et al.*, 2000a,b, 2002a,b; Wataha *et al.*, 2001)
- Liquids such as fluid inclusions (Shepherd and Chenery, 1995; Ghazi *et al.*, 1996; Günther *et al.*, 1997b; McCandless *et al.*, 1997; Audétat *et al.*, 1998; Ulrich *et al.*, 1999; Ghazi and Shuttleworth, 2000)
- Forensic studies (Watling *et al.*, 1997; Watling, 1998)

Although geochemists are the primary users of laser ablation ICPMS, this instrument has gained popularity among other scientists including environmental scientists. Environmental scientists are attracted to the use of LA-ICP-MS because of the minimal sample size, minimal sample preparation, excellent spatial resolution, and ability to work with variety of samples medium.

In environmental forensic studies, where the analytes of concern are metals, LA-ICP-MS is the instrument of choice, because it offers high sample throughput, rapid, accurate, and relatively inexpensive analysis with minimal sample preparation time. For example LA-ICP-MS has been used for determination of metal contaminants and pollutant sourcing in the following areas:

- Dendrochemical studies of environmentally impacted biological samples (e.g., Rainbow and Philip, 1993; Outridge *et al.*, 1995; Hofman *et al.*, 1996a,b; Sinclair *et al.*, 1997; Veinott and Evans, 1999; Thorrold and Shuttleworth, 2000; Veinott, 2001)
- Source identification of anthropogenic lead in environment (McGill *et al.*, 2003; Ghazi and Millette, 2004, 2006)
- Age dating of environmental samples using U-Th isotopic series (Richards and Dorale, 2003; Eggins *et al.*, 2005).

14.4.1 DENDROCHEMICAL STUDIES

Laser ablation ICP-MS is particularly useful in the study of samples, which have incremental patterns of growth such as plants and animals where they record environmental changes in their structures. For example, tree rings and incremental growth of mollusk, shell, bones, teeth, and fish otoliths are excellent media to record and for monitoring environmental pollution in industrial areas. In these studies samples have been ablated and analyzed for their trace elements and isotopic composition (e.g., Pb isotopic ratio) to trace the source of pollution.

Shells and fish otoliths are by far the most widely used samples for analysis with LA-ICP-MS, principally because they are hard and calcified and can potentially provide information on the timing of exposure to a specific pollutant (e.g., monitor changes in their environment from seawater to river influences, sudden influxes of pollutant into aquatic environment, or to study the migratory patterns of certain species and can record past movements of the species. Outridge *et al.* (1995) and Veinott (2001) published reviews of the application of LA-ICP-MS to incremental structures, including a discussion on how various structures are formed. Mollusk shell and otolith have received the most attention for application of LA-ICP-MS in environmental science because these structures are believed to reflect environmental condition at the time of formation of a new layer and are not resorbed after deposition

(Rainbow and Philips, 1993). Usually the outer layers of these organisms are the most likely place for accumulation of metals (e.g., Pb, Cd, Cu, and Zn), thus an indicator of environmental changes. For example, Price and Pearce (1997), in analyzing growth rings on certain species of bivalves, reported spikes of high concentrations of Pb, Cu, and Zn. They contributed the presence of these spikes to short-term exposure of the organism to influx of extreme pollution events such as stream meltdown or sudden flood, which washed the metal-contaminated sediments into the water. Thorrold and Shuttleworth (2000) applied LA-HR-ICP-MS to the analysis of trace element and isotopic signatures in fish otoliths. They found Mn/Ca, Sr/Ca, and Ba/Ca profiles across juvenile Atlantic croaker otoliths, to show considerable spatial and temporal variations that were presumably reflective of variations in water chemistry over the same scales.

By contrast to laser ablation analysis of hard parts, which provide temporally and spatially highly resolved information, analysis of soft tissue can provide only time-integrated concentration values (i.e., bulk analysis) for metals, because concentration of metals in soft tissue reflects accumulation of metals over the entire life of the organism. Therefore, such bulk analysis would mask or dilute such spikes and limit the usefulness of any kind of shells with growth pattern as an environmental monitoring tool.

The study of corals is another research application of LA-ICP-MS. Studies have demonstrated that elements such as B, Mg, Mn, a, Sr, and U show seasonal variations that could be used as paleothermometers (Sinclair *et al.*, 1997, 1998). Similarly, ratios such as Sr/Ca, Mg/Ca in bone and fin of certain species of fish have been used as potential good indicators of environmental changes (Veinott and Evans, 1999).

In addition to dendrochemical applications of LA-ICP-MS in aquatic research, in environmental fields this instrument has been used for analysis of plant materials (tree ring, bark, and leaves). Similar to aquatic studies, the objective of studying individual tree growth rings is to obtain environmental information during the life span of trees (Hofman *et al.*, 1996a).

14.4.2 FINGERPRINTING THE SOURCE OF LEAD IN THE ENVIRONMENT

Lead has been one of the most important metals in human history with a wide range of applications. Historically, lead from the combustion of gasoline and burning coal are the primary contributors to such high levels of anthropogenic lead in our environment. In most cases, the suspected real pathway is soil that has been enriched with lead from old paint flaking from houses or lead that has been transported and distributed as aerosol from natural sources via air. Lead-base paint is another important environmental contaminant,

which is one of the major sources of lead poisoning. Ghazi and Millette (2006) published a comprehensive review of the presence of lead in the environment and a detailed account on various forensic applications of identifying the source of lead.

One of the most challenging issues, but fundamental to understating the effects of lead in environment and human health is the identification of correct sources and pathways of lead. Such understanding is also pivotal in decision-making and the implementation of the most cost-effective remedial actions. Traditionally, solution analysis using various types of ICP-MS have been the primary method of obtaining isotopic data from environmental samples. In more recent days the new generation of high resolution ICP-MS, in particular the multicollector instruments coupled with laser ablation, have generated results that are comparable to traditional thermal ionization mass spectrometry (TIMS). The usefulness of application of lead isotopic ratio studies to distinguish sources of lead (i.e., technique of lead isotope fingerprinting) has been successfully applied in solving lead-related environmental issues for establishing sources of lead. Examples include working with a variety of samples such as air, water, ice, soils, dust, and paint (Chiaradia *et al.*, 1997; Gulson *et al.*, 1997). Isotope ratio measurements using ICP-MS on lead-bearing artifacts and human remains have been used to determine the possible geographic origin of the ore and environmental impacts on historical sites (e.g., Ghazi, 1994; Ghazi *et al.*, 1994; Reinhard and Ghazi, 1994).

For environmental forensic applications, laser ablation ICP-MS has allowed lead isotopic measurements (i.e., $^{208}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$, and $^{206}Pb/^{204}Pb$). However, the usefulness of LA-ICP-MS in lead isotopic ratio studies to distinguish sources of lead depends largely upon the precision obtainable from the system, which depends on the composition and concentration of Pb in the sample. McGill *et al.* (2003) and Ghazi and Millette (2004, 2006) in separate studies have demonstrated unique environmental forensic applications of LA-HR-ICP-MS in source identification of lead found in brown-field samples and lead-based paint.

14.4.3 BASE PAINT

In the study of a layered paint, shown in Figure 14.4, Ghazi and Millette (2004) used laser ablation HR-ICP-MS to qualitatively identify each layer by determining the relative abundance of a suite of commonly used elements in the production of paint, including lead and mercury (see Figure 14.13), and to use the isotope systematic of lead to determine the origin of lead (see Figure 14.14). The data showed six identifiable paint layers. The data revealed that four of the layers contain copious amounts of lead (see

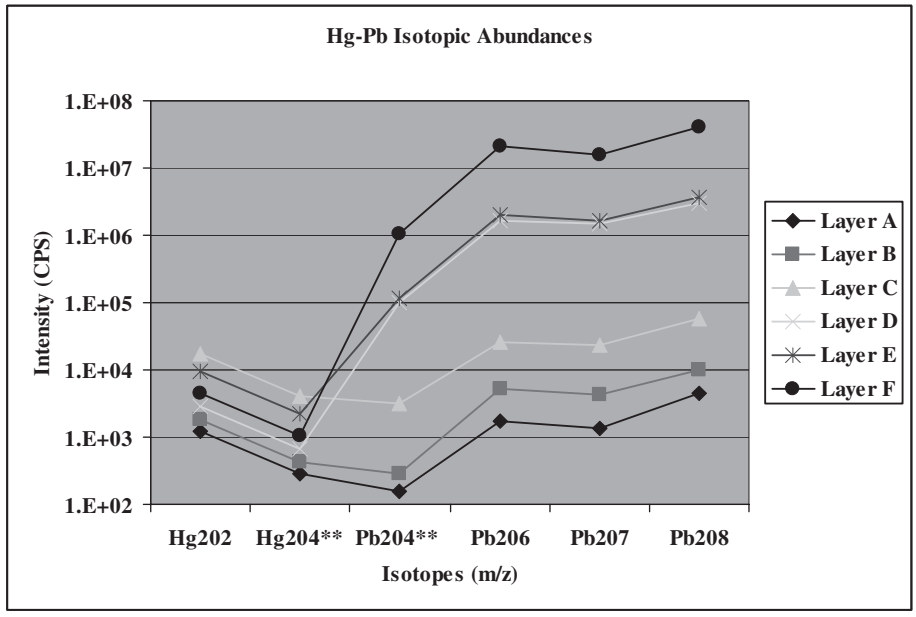


Figure 14.13

A plot of intensity profile in counts per seconds for isotopes (m/z) of lead (204Pb, 206Pb, 207Pb, 208Pb) and two isotopes of mercury (202Hg and 204Hg). The intensity (CPS) is proportional to concentration of Pb and Hg in each layer. Symbol (**) indicates calculated values for intensity of 204Hg and 204Pb, by calculating 204Hg intensity, using natural abundance ratio of 204Hg/202Hg = 6.85/29.8 = 0.229. 204Pb isotope corrections were made by subtracting calculated 204Hg values from total ions at m/z = 204.

Layers A and B are virtually Pb- and Hg-free. The concentration of Hg is highest in layer C, followed by layer E, and layer D also has virtually no Hg. The concentration profile of Pb in all six layers is as follows: A < B < C < D < E < F. However, concentration of Pb in layers D and E are basically identical. Note the crossing-cutting of the patterns for layers D, E, and F by that of layer C is a significant indication of different origin (i.e., mixture) for layer C. The cross-cutting between patterns for layers E and F is due to higher concentration of Hg in layer E.

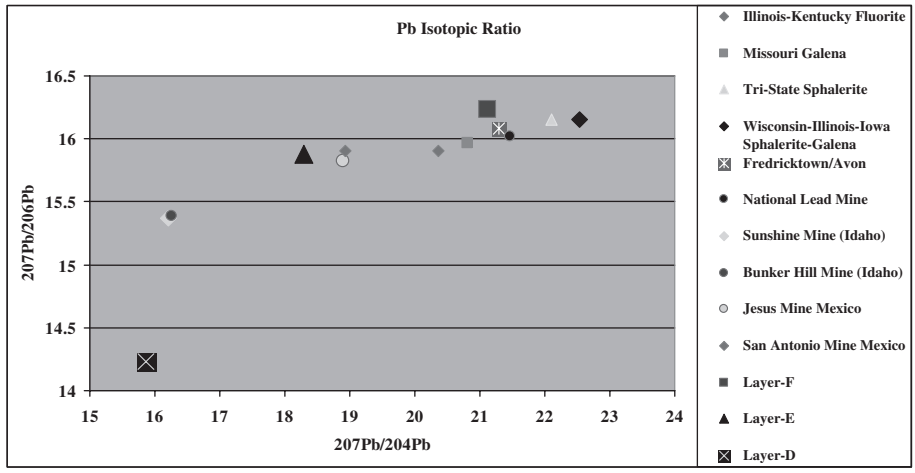


Figure 14.14

Pb isotope ratio correlation diagram for 206Pb/204Pb (vs) 207Pb/204Pb, showing isotope compositions for Pb from a number of Mississippi Valley-type deposit and non-Mississippi Valley-type deposits. Plotted in this diagram with larger fonts are the isotope composition of Pb in Layers D, E, and F (see text for discussion).

Figure 14.13). The analysis also yielded a set of intensity values for ^{202}Hg, ^{204}Hg, ^{204}Pb, ^{206}Pb, ^{207}Pb, and ^{208}Pb isotopes. The intensity data allowed qualitative (relative) determination of lead concentrations, but more importantly to calculate the isotopic composition of lead in each layer.

A challenge in measuring lead isotope ratio by ICP-MS is the elemental interference at $m/z = 204$ that involves ^{204}Hg and ^{204}Pb isotopes with 6.85% and 1.4% isotopic composition, respectively. This interference is particularly problematic when the concentration of mercury is relatively high and the isotope ratios are being made with respect to ^{204}Pb. In order to calculate the true ^{204}Pb values, it is a common practice to include another isotope of mercury in the list of isotopes (e.g., ^{202}Hg = 29.80%), then by using the natural isotopic composition ratio of ^{204}Hg and ^{202}Hg, make the appropriate isotopic correction (e.g., Ghazi, 1994).

The study by Ghazi and Millette (2004) showed that four of the six layers contained significant lead, and only layers A and B were free of lead (see Figure 14.13). Lead in layer F indicated a radiogenic signature, very similar to Pb-Zn mines of the central United States region (Mississippi Valley-type deposit), most likely from the eastern Missouri mining district of the Fredricktown/Avon area (see Figure 14.14). The isotope ratio data for layer E also indicates radiogenic signature, however this signature is more similar to that of Pb-Zn-Ag mining districts of Sierra Madre of eastern Mexico (see Figure 14.14). The Pb isotope ratio for layer D indicates a common (i.e., none radiogenic) lead signature (see Figure 14.14). They argued that the isotope ratio signature of layer D is significantly different than those for layers E and F, despite the fact that this paint layer has nearly similar levels of Pb contents as layer E. The authors concluded that the measurement of isotope ratio signature for layer C was severely impacted as result of high concentration of mercury, thus difficult to explain; however for the purpose of environmental forensic investigations the relative abundance of other elemental analysis can be combined with isotopic ration data to help identify the source for this layer of paint.

14.4.4 LEAD IN CONTAMINATED SOIL

Laser ablation IC-PMS also is used to characterize landfill and brownfield lead contamination. In a study, as part of the United Kingdom Natural Environment Research Council funded Urban Regeneration (URGENT) Thematic Program, a range of urban soils and contamination types were characterized by using a combination of imaging and high precision Pb-isotope fingerprinting (McGill *et al.*, 2003). In this study laser ablation Multicollector High Resolution ICP-MS (MC-HR-ICP-MS) was used to analyze press pellets (similar to those used for XRF analysis) of well homogenized highly contaminated

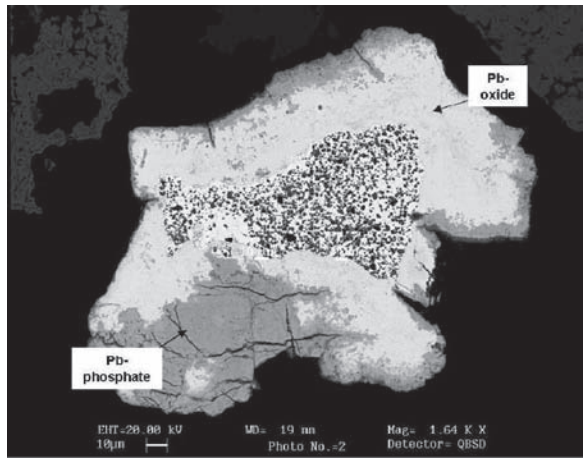


Figure 14.15
Sample WV6, 250–63 µm sinks—BSEM image of polished, section through an irregular grain with core of lead-rich inclusions of silica, calcite and graphite inside a Pb-oxide margin part altered to Pb-phosphate (modified after McGill et al., 2003).

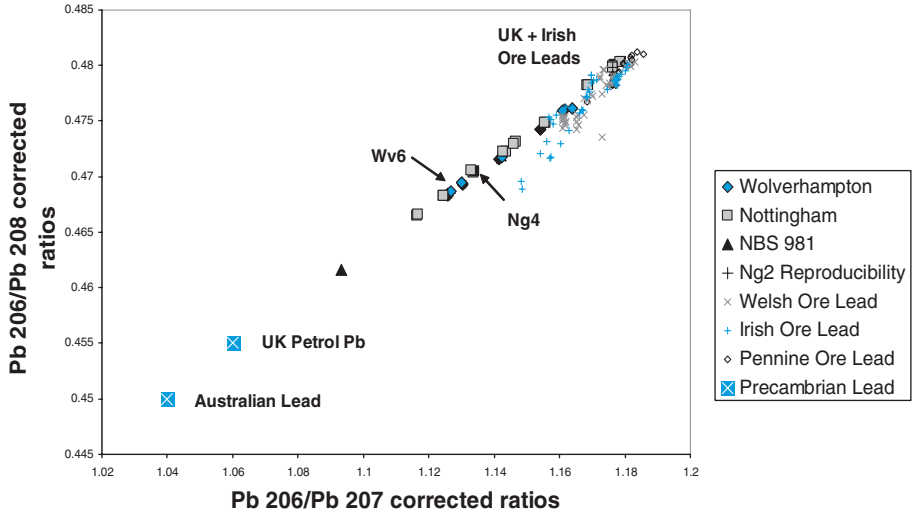


Figure 14.16
(206/207) Pb/(206/208)Pb Wolverhampton and Nottingham XRF Pellets (modified after McGill et al., *2003).*

soil samples from landfill and brown-field sites around Nottingham and Wolverhampton (see Figure 14.15). McGill *et al.* (2003) used $^{206}Pb/^{207}Pb$ versus $^{206}Pb/^{208}Pb$ diagram to plot the results from the soil powder pellets (see Figure 14.16). The results for each pellet represent a bulk soil lead-isotope signature, which constituted a mixture of all possible lead components in that soil. The authors proposed that the plot shows a range along an apparent mixing line suggestive of two dominant contamination sources (see Figure 14.16). The authors concluded it is possible to use LA-MC-ICP-MS to establish the extent of lead source variability in this type of environmental samples, as well as in determining the source of individual lead-bearing particles.

14.5 ENVIRONMENTAL IMPACTS TO METAL EXPOSURE IN TISSUE

Laser ablation can also be used in applications such as *in situ* determination and elemental profiling of tissue, teeth, and bones that have been exposed to metal contaminants. In a study at the University of Georgia School of Dentistry, LA-HR-ICP-MS was used to examine the spatial elemental profiling of nickel concentration in tissues exposed to nickel wire. The diffusion of nickel with time was observed, as well as its spatial distribution around nickel containing implants *in vivo* in rats (see Figure 14.17). They successfully analyzed tissues for nickel content and inflammatory response using ^{24}Mg and ^{60}Ni isotopes (Wataha *et al.*, 2001; Ghazi *et al.*, 2002a). Although the rastering technique was used for data acquisition (see Figure 14.17), ^{24}Mg was used as an internal standard element for quantification. The concentration of magnesium in tissue was determined by bulk analysis using graphite furnace atomic absorption analysis.

The data revealed a "significant penetration of nickel ions into tissues exposed to nickel wire implants." Within 96 hours, up to 30 micrograms per gram of nickel could be absorbed by tissues within the immediate vicinity of the implant site. However, the level of nickel falls exponentially to undetectable levels just 3 to 4 mm away from the implants (see Figure 14.18). The team points out that a further advantage of using this technique is that the distribu-

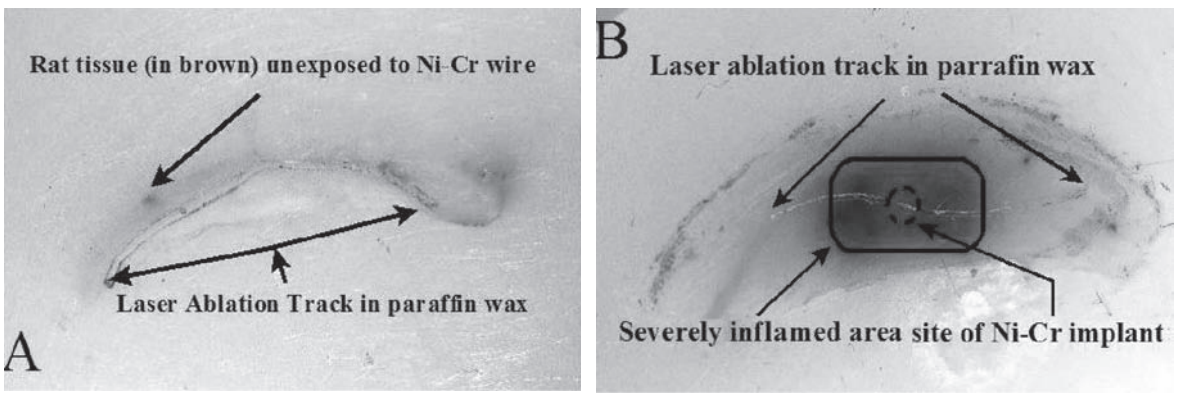


Figure 14.17

Photomicrograph of two tissues, showing laser ablation tracks and the difference between inflammation levels. Sample A is basically a blank and has never been exposed to Ni-implant; as a result no inflamation is present. Sample B shows inflammatory response by the tissue due to exposure to Ni-implant for a period of seven days. The laser ablation track indicates the general directions from the left to right analysis by going through the center of the specimen (i.e., dashed circle), which was the site of the Ni-implant that was subsequently removed and filled with wax. The dark brown areas in the photo show the tissues that have been destroyed or severely infiltrated by the overwhelming inflammatory response; even the muscle tissues are destroyed or damaged.

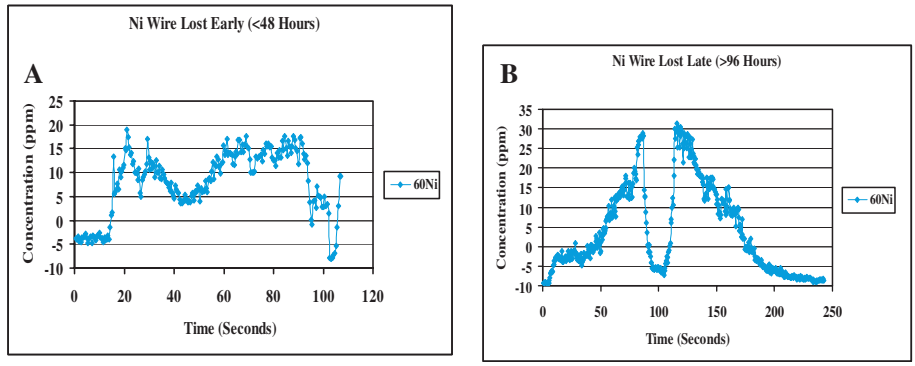


Figure 14.18

*Concentration profiles of distribution of nickel as 60Ni ion in tissue sample in response to subcutaneous implantation of nickel wire for periods of <48 hours (**A**) and >96 hours (**B**). The x axis shows the time of analysis in seconds and the analysis was performed from left to right (e.g., Figure 14.17). As the movement of the specimen beneath the laser occurred at a constant rate (0.04 mm per seconds), the time in seconds is also distance (i.e., 10 seconds = 0.4 mm). The large trough in each spectrum is the site for the nickel implant, which was removed after 96 hours and subsequently filled with paraffin during later stages of sample preparation. Nickel reached the concentrations of greater than 30 μg/g near the implant, then nearly decayed exponentially to the background levels 3–4 mm away (i.e., in both directions) from the site of the implant. Inflammatory responses were severe with necrosis, extending 4–5 mm from the implant site. Note that the site of the implant in all samples were taken as the reference point for the purpose of the nickel spatial distribution and inflammatory responses studies.*

tion of metal ions can be determined without disturbing the tissue structure, so the same sample might be examined again.

In another dental-related study, laser-ablation high resolution inductively coupled plasma mass spectrometry was used for *in situ* determination of silver concentration in treated human teeth (Ghazi *et al.*, 2002b). Silver nitrate has long been used to trace the presence of micro-sized gaps between the walls of dental restorations and teeth, and more recently, to trace the presence of nanometer-sized voids within resin-bonded restorations. This study examined the diffusion of silver nitrate into submicron voids beneath resin-bonded composite restorations (see Figures 14.19 and 14.20).

The silver treated-teeth were analyzed for ^{43}Ca, ^{107}Ag, and ^{109}Ag isotopes. Quantification was performed by assuming a uniform nominal calcium concentration value of 27 μg/g of dry dentin across the root and using the ^{43}Ca signal for internal calibration. The analysis revealed a significant penetration of resin-bonded dentin by silver, ranging from ~35,000 parts per million (ppm) at the very bottom of the resin-infiltrated demineralized dentin to only 10 s of ppm in the areas within the underlying mineralized dentin (see Figure 14.20). The analysis revealed that a significant amount of silver penetrated

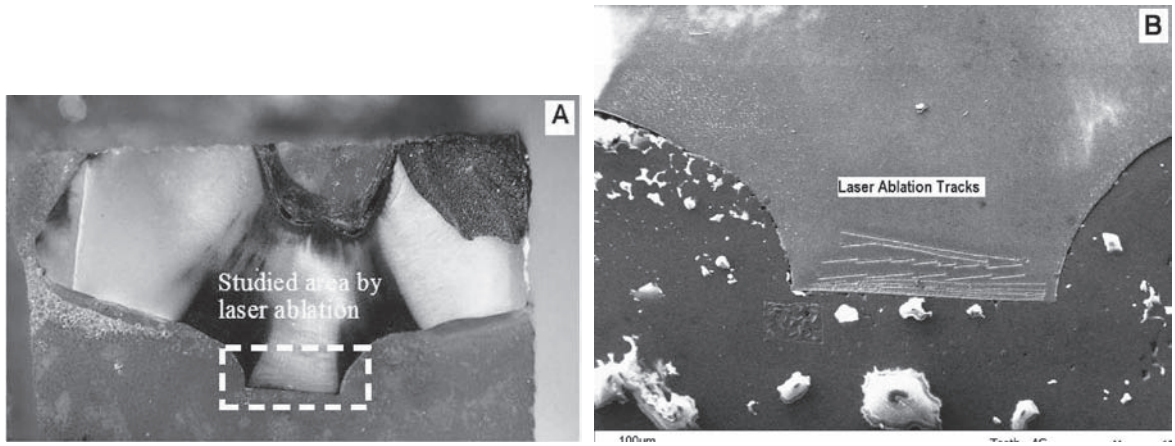


Figure 14.19

***A**. Photomicrograph of a 0.7 mm thick slab of a restored tooth that was serially sectioned longitudinally for analysis of the silver concentration profile. Each 0.7 mm thick slab was ground with an ultrafine diamond bar from both sides, to shape the specimen into hour-glass shape to permit uniform tensile stress inside the bonded interface for bond strength determination. Prior to measurement of tensile bond strength, the hour-glass-shaped specimens were coated with nail varnish except for 0.5 mm on either side of the bonded interface. The sample was immersed in 2.9 moles/liter AgNO3 and the black areas show the zones of more intense uptake of silver nitrate. The dashed rectangle on the photograph shows the area that was analyzed by LA-ICP-MS. **B**. Scanning electron microscope (SEM) photomicrograph of several laser ablation tracks along the edge of the tooth below the bonded surface also shown in Figure 14.20.*

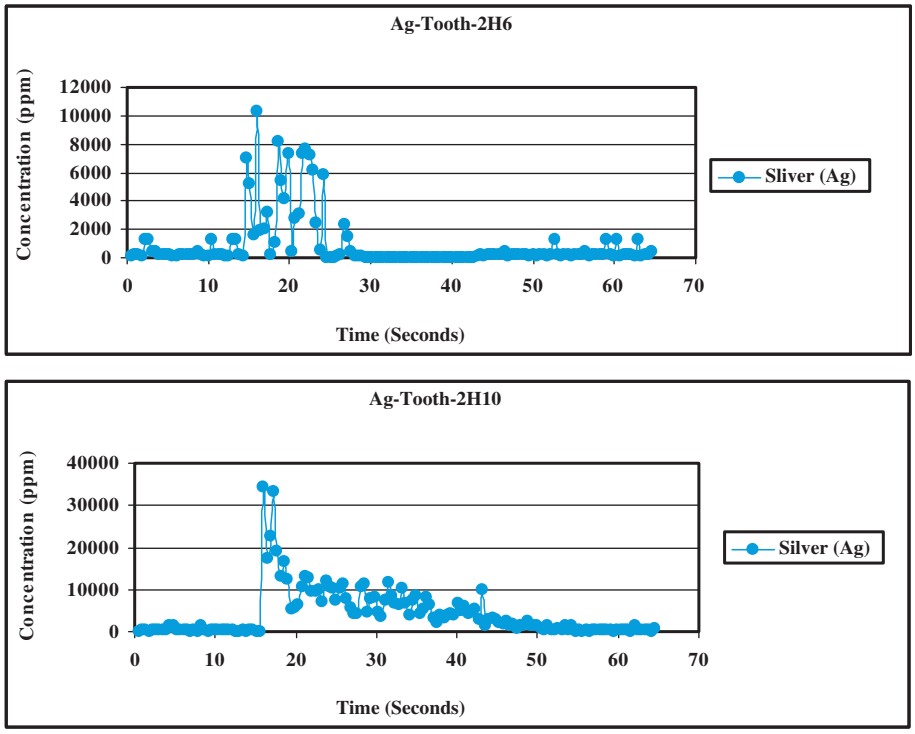


Figure 14.20

Profile of distribution of silver (107Ag and 109Ag) in tooth specimen shown in Figure 14.18 along laser ablation tracks 2H6 and 2H10. These two laser ablation tracks represent two different areas, i.e., intermediate (2H6) and closest (2H10) relative to the adhesive-bonded edge. The time axis can be converted to distance since the specimen moved under the laser beam at constant rate (0.04 mm per seconds), the time in seconds is also distance (i.e., 10 seconds = 0.4 mm) in all figures.

through resin–dentin bonds, and that the amount of silver increased with increased acid-etching time. The study suggests that current adhesive resin systems do not create perfect bonds, and that laser ablation high-resolution ICP-MS can be used to quantify the amount of nano-leakage of bonded restorations.

14.6 ENVIRONMENTAL FORMATION OF DIAMONDS

The physics and chemistry of diamonds and their mineral inclusions possibly can be used to determine the geographic source and environment of formation of rough diamonds. The science of diamonds includes their morphology and external features, nitrogen aggregation states and zonations, ^{13}C, ^{15}N, trace-element chemistry, and the nature of their rims and coatings (Taylor *et al.*, 2001). The science of the diamond inclusions (DIs) includes their mineralogy and chemical/isotopic compositions, the variation of mineral chemistry between diamonds and as multiple DIs within a single diamond, and the relations of DIs to geologic locales.

A method for discerning the locale of a diamond's origin involves the combined use of laser-ablation, inductively coupled plasma, mass spectrometry (LA-HR-ICP-MS), and Raman spectroscopy (Taylor *et al.*, 2001). These sophisticated instruments are combined in a study of the outermost rims of coated diamonds (see Figure 14.21). The premise to this endeavor is that these coatings contain the remnants of the latest stages of diamond formation, within the kimberlite as the diamonds are carried to the earth's surface (see Figure 14.21). The mantle of the earth is heterogeneous to some extent, although

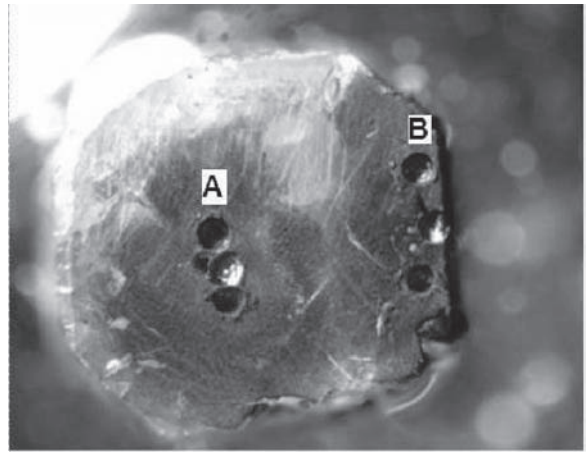


Figure 14.21

Photomicrograph showing laser ablation pits in a cross-section of a polished diamond. The pits on the right have been drilled in the rim area, whereas those on the left were drilled in the core. The spot size is 20 μm.

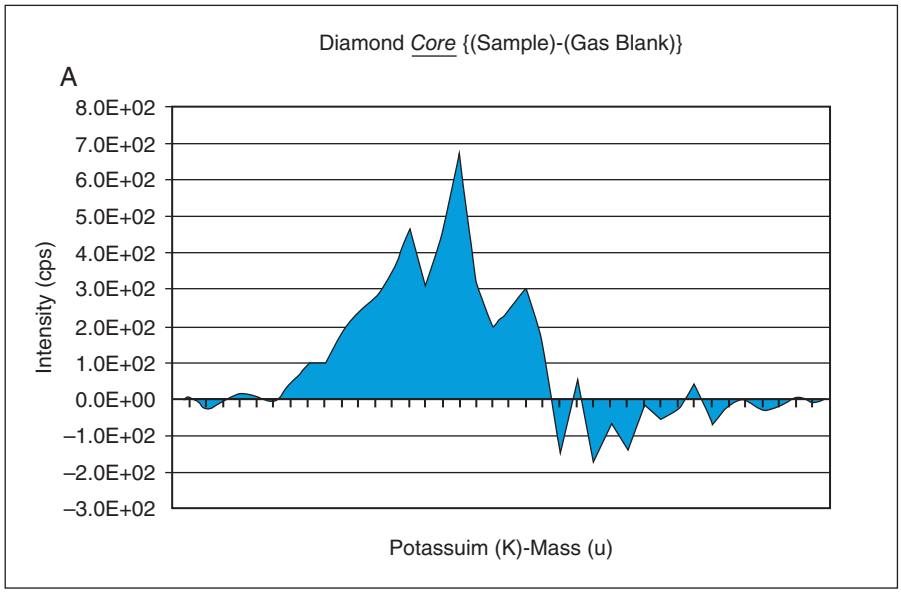


Figure 14.22
Time-resolved spectrum for one atomic mass unit at potassium (39K). The data is in count per seconds of gas blank-subtracted K in the CORE of the diamond sample shown in Figure 14.21.

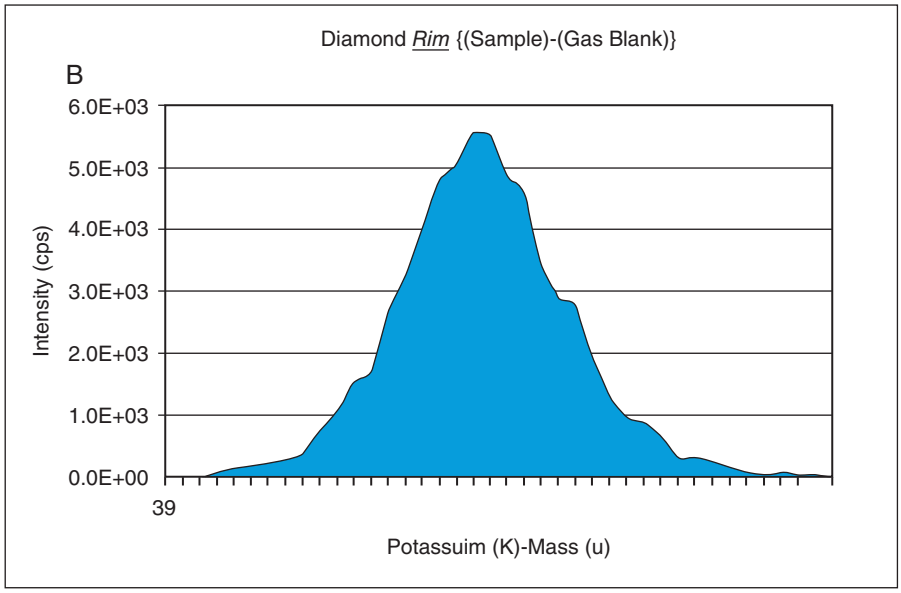


Figure 14.23
Time-resolved spectrum for one atomic mass unit at potassium (39K). The data is in count per seconds of gas blank-subtracted K in the RIM of the diamond sample shown in Figure 14.21.

globally it is quite similar, based upon DI studies. However, the chemistry of kimberlite magmas, as they traverse the upper mantle and lower crust on their way to the surface, can pick up late-stage characteristics that differ from place to place. Some of these definitive signatures may be passed on to the outermost chemistry and physics of the transported diamonds (see Figures 14.22 and 14.23).

14.7 SUMMARY

As shown in some of the aforementioned cases, environmental forensic studies that involve analysis of metals is very broad, ranging from analysis of lead-based paint to incremental growth patterns in plant and animal parts, to contaminated tissue studies, to contaminated soil studies, and perhaps to analysis of contaminated airborne particles. Laser ablation ICP-MS clearly has shown promise as a powerful method for rapid and accurate elemental and isotopic ratio analysis of trace elements in environmental samples. Both semi-quantitative and quantitative multielemental analysis can be achieved at relatively low cost. The capability of LA-ICP-MS to perform rapid isotopic ratio analysis at relatively low cost is a significant advantage of this technique for environmental forensic investigations. Moreover, laser ablation ICP-MS offers the added advantage of solid sampling capabilities, which significantly reduces the difficulties and the extra steps associated with sample preparation. More importantly, high spatial resolving power of laser ablation (in microns) allows measurements of spatial distribution of metal contaminants and elemental mapping of contaminated samples at microscopic scales.

In environmental forensic investigations, where the matrix can vary significantly (e.g., water, air, solid, tissue, bone, teeth), perhaps the greatest challenge facing the user is quantification of pollutant in order to be able to create defensible cases. Despite all the exciting and innovative applications that can be performed by using LA-ICP-MS there is still tremendous need for development of new standards and reference materials depending on the type of materials that are being analyzed. Nevertheless, as it was shown earlier, the capability of elemental mapping with a great resolving power and the ability in producing an enormous amount of data in virtually no time makes LA-ICP-MS technique arguably the most powerful analytical techniques among all available inorganic analytical instrumentation.

REFERENCES

Arrowsmith, P. (1987) Laser ablation of solids for elemental analysis by inductively coupled plasma mass spectrometry. *Analytical Chemistry* 59, 1437–1444.

Audétat, A., Günther, D., and Heinrich, C.A. (1998) Formation of a magmatic-hydrothermal ore deposit—Insights with LA-ICP-MS analysis of fluid inclusions. *Science* 279, 2091–2094.

Audétat, A. and Günther, D. (1999) Mobility and H2O loss from fluid inclusions in natural quartz crystals. *Contribution to Mineralogy Petrology* 137, 1–14.

Balhaus, C., Sylvester, P., and Ellis, D. (1997) Spatial distribution of platinum-group elements in magmatic sulphides. *RSES—Annual Report* 96–97.

Bea, F., Montero, P., Stoh, A., and Baasner, J. (1996) Microanalysis of minerals by an Excimer UV-LA-ICP-MS system. *Chemical Geology* 133, 145–156.

Becker, J.S. and Dietze, H.-J. (1997) Double-focusing sector field inductively coupled plasma mass spectrometry for highly sensitive multi-element and isotopic analysis. *Journal of Analytical Atomic Spectrometry* 12, 881–889.

Becker, J.S. (2002) State-of-the-art and progress in precise and accurate isotope ratio measurements by ICP-MS and LA-ICP-MS. *Journal of Analytical Atomic Spectroscopy* 17, 1172–1185.

Chaudhary-Webb, M., Paschal, D., Elliott, C.W., Hopkins, H., Ghazi, A.M., Ting, B., and Romieu, I. (1999) Determination of lead isotope ratios in whole blood and pottery-lead sources in Mexico City. *Atomic Spectroscopy* 19, 156–163.

Chaudhury-Webb, M., Paschal, D., Romeiu, I., Ting, B., Elliott, W.C., Hopkins, H.P. *et al.* (2003) Determining lead sources in Mexico using lead isotopes. *Revista Salud Pública de México* 45, S183–S188.

Chiaradia, M., Chenhall, B., Depers, A.M., Gulson, B.L., and Jones, B.G. (1997) Identification of historical lead sources in roof dusts and recent lake sediments from an industrialized area: Indications from lead isotopes. *Science Total Environment* 205, 107–128.

Clive, R.N., Davidson, J.P., and McKeegan, K.D. (1995) Geochemical analysis of small samples: Micro-analytical techniques for the nineties and beyond. In *U.S. National Report to International Union of Geodesy and Geophysics 1991–1994*.

Date, A.R. and Gray, A.L. (1983) Development progress in plasma source mass spectrometry. *Analyst* 108, 159–165.

Date, A.R. and Gray, A.L. (1989) Applications of inductively coupled plasma mass spectrometry. Blackie, Glasgow, 254.

Denoyer, E.R., Kenneth, J.F., and Hager, J.W. (1991) Laser solid sampling for inductively coupled plasma mass spectrometry. *Analytical Chemistry* 63, 445A–457A.

Ducreux-Zappa, M. and Mermet, J.-M. (1996) Analysis of glass by UV laser ablation inductively coupled plasma atomic emission spectrometry. Part1. Effects of the laser parameters on the amount of ablated material and the temporal behaviour of the signal for different types of laser. *Spectrochimica Acta* 51B, 321–332.

Eggins, S., Shelley, M., Grün, R., McCulloch, M.T., Kinsley, L., Pike, J.C. *et al.* (2005) In situ U-series dating by laser-ablation multi-collector ICPMS: New prospects for Quaternary geochronology. *Quaternary Science Reviews* 24, 2523–2538.

Eggins, S.M., Rudnick, R.L., and McDonough, W.F. (1998) The composition of peridotites and their minerals: A laser-abaltion ICP-MS study. *Earth and Planetary Science Letters* 154, 53–71.

Ghaderi, M., Campbell, I.H., and Sylvester, P.J. (1997) *Eu anomalies in scheelite*: Implications for the oxidation state of hydrothermal gold mineralisation in Archaean greenstone belts. *RSES—Annual Report*, 159–160.

Ghazi, A.M., Vanko, D.A., Roedder, E., and Seeley, R.C. (1993) Determination of rare earth elements in fluid inclusions by inductively coupled plasma mass spectrometry (ICP-MS). *Geochimica et Cosmochimica Acta* 57, 4513–4516.

Ghazi, A.M. (1994) Lead in archaeological samples: An isotopic study by ICP-MS. *Applied Geochemistry* 9, 627–636.

Ghazi, A.M., Reinhard, K.J., and Durrance, E.M. (1994) Brief communication, further evidence of lead contamination of Omaha skeleton: An application of ICP-MS. *Journal of American Physical Anthropologists* 95, 427–434.

Ghazi, A.M., McCandless, T., Ruiz, J., and Vanko, D.A. (1996) Quantitative elemental determination in fluid inclusions by laser ablation ICPMS: Applications to Sr and Rb measurements in fluid inclusions in halite. *Journal of Analytical Atomic Spectrometry* 11, 667–674.

Ghazi, A.M. and Shuttleworth, S. (2000) Quantitative trace elemental analysis of single fluid inclusions by laser ablation ICPMS (LA-ICPMS). *Analyst* 125, 205–210.

Ghazi, A.M., Shuttleworth, S., Angula, S.J., and Pashley, D.H. (2000a) Gallium diffusion in human root dentin: Quantitative measurements by pulsed Nd:YAG laser ablation ICPMS (LA-ICPMS). *Journal of Clinical Laser Medicine and Surgery* 18, 173–183.

Ghazi, A.M., Shuttleworth, S., Angula, S.J., and Pashley, D.H. (2000b) New applications for laser ablation high resolution ICPMS (LA-HR-ICPMS): Quantitative measurements of gallium diffusion across human root dentin. *Journal of Analytical Atomic Spectrometry* 15, 335–341.

Ghazi, A.M., Wataha, J.W., Shuttleworth, S., Simmons, R.B., O'dell, N.L., and Singh, B.B. (2002a) Quantitative concentration profiling of nickel in tissues around metal implants: A new biomedical application of laser ablation sector field ICP-MS. *Journal of Analytical Atomic Spectrometry* 17, 1295–1299.

Ghazi, A.M., Shuttleworth, S., Simmons, R.B., Paul, S.J., and Pashley, D.A. (2002b) Nanoleakage at the dentin adhesive interface vs μ-tensile bond strength: A new application for laser ablation high resolution ICPMS. *Journal of Analytical Atomic Spectrometry* 17, 682–687.

Ghazi, A.M. and Millette, J.R. (2006) Lead. In *Environmental Forensic, Contaminant Specific Guide*. R.D. Morrison and B.L. Murphy, eds. Academic Press, 55–74.

Ghazi, A. Mohamad and Millette, J.R. (2004) Pb-isotope fingerprinting for origin of lead-based paint: A new environmental forensic application for laser ablation high resolution ICPMS (LA-HR-ICPMS). *Environmental Forensics* 5, 97–108.

Ghazi, A.M. (In Press) Laser Ablation ICP-MS: A New Elemental and Isotopic Ratio Technique in Environmental Forensic Investigation. *Environmental Forensics*.

Gray, A.L. and Jarvis, K.E. (1989) *Applications of Inductively Coupled Plasma Mass Spectrometry*. Chapman & Hall, 324.

Gray, A.L., Houk, S., and Jarvis, K.E. (1992) *Handbook of Inductively Coupled Plasma Mass Spectrometry*. Chapman & Hall, 380.

Günther, D., Frischknecht, R., Heinrich, C.A., and Kahlert, H.J. (1997) Capabilities of an argon fluoride 193 nm excimer laser for laser ablation inductively coupled plasma mass spectrometry microanalysis of geological materials. *Journal of Analytical Atomic Spectrometry* 12, 939–944.

Günther, D., Frischknecht, R., Müschenborn, H.J., and Heinrich, C.A. (1997) Direct liquid ablation: A new calibration strategy for laser ablation-ICP-MS microanalysis of solids and liquids. *Fresenius Journal of Analytical Chemistry* 359, 390–393.

Günther, D., Audetat, A., Frischknecht, R., and Heinrich, C.A. (1998) Quantitative analysis of major, minor and trace elements using laser ablation inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* 13, 263–270.

Heinrich, C.A., Gunther, D., Audetat, A., Ulrich, T., and Frischknecht, R. (1999) Metal fractionation between magmatic brine and vapor, determined by microanalysis of fluid inclusions. *Geology* 27, 755–758.

Hoffmann, E., Ludke, C., and Stephanowitz, H. (1996) Application of laser-ICP-MS in environmental analysis. *Fresinius Journal of Analytical Chemistry* 355, 900–903.

Hoffmann, E., Stephanowitz, H., and Skole, J. (1996) Investigations of the migration of elements in tree rings by laser-ICP-MS. *Fresenius Journal of Analytical Chemistry* 355, 690–693.

Houk, R.S., Fassle, V.A., Flesch, G.D., Svec, H.J., Gary, A.L., and Taylor, C.E. (1980) Inductively coupled plasma as an ion source for mass spectrometric determination of trace elements. *Analytical Chemistry* 52, 2238–2289.

Jackson, S.E., Pearson, N., and Griffin, W.L. (2001) In situ isotopic determination using laser ablation (LA)-Magnetic Sector-ICPMS. In *Laser-Ablation-ICPMS in Earth Sciences Principle and Applications*, P. Sylvester, ed. Mineralogical Association of Canada, 105–120.

Jakubowski, N., Moens, L., and Vanhaecke, F. (1998) Sector field mass spectrometry in ICP-MS. *Spectrochimica Acta B* 53, 1739–1763.

Jarvis, I. and Jarvis, K.E. (1992) Plasma spectrometry in the earth sciences: Techniques, applications and future trends. *Chemical Geology* 95, 1–33.

Jeffries, T.E., Jackson, S.E., and Longerich, H.P. (1998) Application of a frequency quintupled Nd:YAG source (= 213 nm) for laser ablation inductively coupled plasma mass spectrometric analysis of minerals. *Journal of Analytical Atomic Spectrometry* 13, 935–940.

Kinsley, L. Shelley, J.M.G., and Eggins, S.M. (1997) Processes affecting laser ablation sampling of materials for analysis by ICP-MS. *RSES—Annual Report*, 137.

Koppenaal, D.W. (1992) Atomic mass spectrometry. *Analytical Chemistry* 64, 320R–340R.

Lahaye, Y. (2002) DFG Short Course Notes LA-ICP-MS (http://user.uni-frankfurt.de/~lahaye/html/icp.HTM).

Lichte, F.C. (1995) Determination of elemental content of rocks by laser ablation inductively coupled plasma mass spectrometry. *Analytical Chemistry* 67, 2479–2485.

Longerich, H.P., Jackson, S.E., Fryer, B.J., and Strong, D.F. (1993) The laser ablation microprobe-inductively coupled plasma-mass spectrometer. *Geoscience Canada* 20, 21–27.

McCandless, T.E., Lajack, D.J., Ruiz, J., and Ghazi, A.M. (1997) Trace element determination of single fluid inclusions in quartz by laser ablation ICP-MS. *Geostandards Newsletter* 21, 271–278.

Machado, N. and Simonetti, A. (2001) U-Pb dating and Hf isotopic composition of zircon by laser ablation –MC-ICP-MS. In *Laser-Ablation-ICPMS in Earth Sciences Principle and Applications*, P. Sylvester, ed. Mineralogical Association of Canada, 121–146.

McGill, R.A.R., Pearce, J.M., Fortey, N.J., Watt, J., Ault, L., and Parrish, R.R. (2003) Contaminant source apportionment by PIMMS lead isotope analysis

and SEM-image analysis. *Environmental Geochemistry and Health* 25, 25–32.

Mason, P. and Mank. Depth analysis by laser ablation ICP-MS. (2001) In *Laser-Ablation-ICPMS in Earth Sciences Principle and Applications*, P. Sylvester, ed. Mineralogical Association of Canada, 93–104.

Moissette, A., Shepherd, T.J., and Chenery, S.R. (1996) Calibration strategies for the elemental analysis of individual aqueous fluid inclusions by laser ablation inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* 11, 177–185.

Monna, F., Lancelot, J., Croudace, I.W., Cundy, A.B., and Lewis, J.T. (1997) Pb isotopic composition of airborne particulate material from France and the southern United Kingdom: Implications for lead pollution sources in urban areas. *Environmental Science and Technology* 31, 2277–2286.

Montaser, A. and Golightly, D.W. (1992) *Inductively Coupled Plasmas in Analytical Atomic Spectrometry*, 2e. John Wiley & Sons, 1017.

Montaser, A. (1998) *Inductively Coupled Plasmas Mass Spectrometry*. John Wiley & Sons, 1000.

Newman, A. (1996) Elements of ICP-MS. *Analytical Chemistry* 65, 46A–51A

Neal, C.R., Davidson, J.P., and McKeegan, K.D. (1995) Geochemical analysis of small samples: Micro-analytical techniques for the nineties and beyond. In *U.S. National Report to International Union of Geodesy and Geophysics 1991–1994*, (http://www.agu.org/revgeophys/neal00/neal00.html).

Outridge, P.M., Veinott, G., and Evans, R.D. (1995) Laser ablation ICP-MS analysis of incremental biological structures: Archives of trace-element accumulation. *Environmental Reviews* 3, 160–170.

Pearce, N.J.G., Perkins, W.T., Abell, I., Duller, G.A.T., and Fuge, R. (1992) Mineral microanalysis by laser ablation inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* 7, 53–57.

Pearce, N.J.G., Perkins, W.T., Westgate, J.A., Gorton, M.P., Jackson, S.E., Neal, C.R., and Chenery, S.P. (1997) A compilation of new and published major and trace element data for NIST SRM 610 and NIST SRM 612 glass reference materials. *Geostandards Newsletter* 21, 115–144.

Perkins, W.T., Pearce, N.J.G., and Fuge, R. (1992) Analysis of zircon by laser ablation and solution inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* 7, 612–616.

Perkins, W.T. and Pearce, N.J.G. (1995) Microprobe techniques in Earth Sciences, P.J. Potts *et al.,* eds. Chapman and Hall, London, 291–325.

Price, G.D. and Pearce, N.J.G. (1997) Bio-monitoring of pollution by *Cerastoderma edule* from British Isles: A laser ablation ICP-MS study. *Marine Pollution Bulletin* 34, 1025–1031.

Rainbow, P.S. and Philips, D.J.H. (1993) Cosmopolitan biomonitor of trace metals. *Marine Pollution Bulletin* 26, 593–601.

Reinhard, K.J. and Ghazi, A.M. (1992) Environmental lead contamination of a 17th century Omaha Indian population as evidenced by Inductively Coupled Plasma Spectrometric (ICP-MS) analysis of skeletal remains and artifacts. *Journal of American Physical Anthropologist* 89, 183–195.

Richards, D.A. and Dorale, J.A. (2003) Uranium-series chronology and environmental applications of speleothems. *Reviews in Mineralogy and Geochemistry* 52(1), 407–460.

Sargent, M. and Webb, K. (1993) Instrumental aspects of inductively coupled plasma-mass spectrometry. *Spectroscopy Europe* 5, 21–28.

Sheppherd, T.J. and Chenery, S.R. (1995) Laser ablation ICP-MS elemental analysis of individual fluid inclusions: An evaluation study. *Geochimica Cosmochimica Acta* 59, 3997–4007.

Shuttleworth, S. and Kemser, D.T. (1996) An assessment of laser ablation and sector field inductively coupled plasma mass spectrometry for elemental analysis of solid samples. *Journal of Analytical Atomic Spectrometry* 12, 412–422.

Sinclair, D.J. and McCulloch, M.T. (1997) The coral record of terrestrial fluxes into the Great Barrier Reef: Analysis of trace elements in inshore corals by laser-ablation ICPMS. *RSES—Annual Report*, 138–139.

Sinclair, D.J., Kinsley, L.P.J., and McCulloch, M.T. (1998) High resolution analysis of trace elements in corals by laser ablation ICP-MS. *Geochimica et Cosmochimica Acta* 62, 1889–1901.

Sylvester, P. (1997) New technique using laser ablation blasts into geochemical labs. *EOS Trans Am Geophys Union* 78, 117–120.

Sylvester, P. and Ghaderi, M. (1997) Trace element analysis of scheelite by excimer laser ablation-inductively coupled plasma-mass spectrometry (ELA-ICP-MS) using a synthetic silicate glass standard. *Chemical Geology* 141, 49–65.

Sylvester, P. (2001) Laser-ablation-ICPMS in earth sciences, principle and applications. *Mineralogical Association of Canada, Short Course Series Volume 29*, 240.

Taylor, L.A., Sobolev, N.V., Ghazi, A.M., Anand, M., and Bodnar, R. (2002) The science of diamonds and their inclusions: Can such data be used to establish diamond provenance? *EOS Trans Amer Geophys Union* 83, S369.

Thorrold, S.R. and Shuttleworth, S. (2000) In situ analysis of trace elements and isotope ratios in fish otoliths using laser ablation sector field inductively coupled plasma mass spectrometry. *Canadian Journal of Fishery and Aquatic Science* 57(6), 1232–1242.

Ulrich, T., Günther, D., and Heinrich, C.A. (1999) Gold concentrations of magmatic brines and the metal budget of porphyry copper deposits. *Nature* 399, 676–679.

Veinott, G.I. and Evans R.D. (1999) An examination of elemental stability in the fin ray of the white sturgeon with laser ablation sampling–inductively coupled plasma–mass spectrometry (LAM-ICP-MS). *Transaction American Fishery Society* 128, 1232–1242.

Veinott, G.I. (2001) The use of laser ablation in environmental sciences. In *Laser-Ablation-ICPMS in Earth Sciences Principle and Applications*, P. Sylvester, ed. Mineralogical Association of Canada, 213–224.

Walder, A.J., Abell, I.D., Platzner, I., and Freedman, P.A. (1993) Lead isotope ratio measurements of NIST 610 glass by laser ablation inductively coupled mass spectrometry. *Spectrochimica Acta* 48B, 397–402.

Walder, A.J. and Furuta, N. (1993) High-precision lead isotope ratio measurement by inductively coupled plasma multiple collector mass spectrometry. *Analytical Sciences* 9, 675–680.

Walder, A.J., Koller, D., Reed, N.M., Hutton, R.C., and Freedman, P.A. (1993) Isotope ratio measurement by inductively coupled plasma multiple collector mass spectrometry incorporating a high efficiency nebulization system. *Journal of Analytical Atomic Spectrometry* 8, 1037–1041.

Wataha, J.C., O'Dell, N.L., Singh, B.B., Ghazi, A.M., Whiteford, G.M., and Lockwood, P.E. (2001) Relating nickle-induced tissue inflammation to nickle release in vivo. *Journal of Biomedical Material Research* 49, 537–544.

Watling, R.J., Lynch, B.F., and Herring, D. (1997) Use of laser ablation inductively coupled plasma mass spectrometry for fingerprinting scene of crime evidence. *Journal of Analytical Atomic Spectrometry* 12, 195–203.

Watling, R.J. (••) Sourcing the provenance of canabis crops using inter-element association patterns "fingerprinting" and laser ablation inductively coupled plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry*.

CHAPTER 15

EMERGING FORENSIC TECHNIQUES

Jean-Christophe Balouet, Gil Oudijk, Ioana Petrisor, and Robert Morrison

15.1 Introduction to Emerging Technologies 672
15.2 Dendroecology 672
15.2.1 Introduction 672
15.2.2 Methodology 674
15.2.2.1 Sampling Considerations 675
15.2.2.2 Sample Preparation 675
15.2.2.3 Ring-Width Measurement 676
15.2.2.4 Chemical Measurement of Tree Rings 680
15.2.2.5 Chemical Microanalysis 680
15.2.2.6 Isotopic and Organic Compound Analyses of Tree Rings 681
15.2.2.7 Calibration 682
15.2.3 Interpretation of Tree Ring Data 682
15.2.4 Applications of Dendroecology to Site Assessment 683
15.2.5 Applications of Dendroecology in Environmental Forensic Investigations 684
15.2.6 Conclusions 685
15.3 Use of Atmospheric Contaminants for Age-Dating Contaminants in Groundwater 686
15.3.1 Introduction 686
15.3.2 Use of Tritium for Contaminant Age-Dating 687
15.3.3 Use of Chlorofluorocarbons (CFCs) for Age-Dating 691
15.3.4 Analytical Methodology 694
15.3.5 Limitations and Considerations 694
15.3.6 Conclusions 695
15.4 DNA Fingerprinting/Microbiological Techniques 695
15.4.1 Introduction 695
15.4.2 DNA (Deoxyribonucleic Acid) 696
15.4.3 Protein Synthesis and RNA (Ribonucleic Acid) 697
15.4.4 PCR (Polymerase Chain Reaction) 697
15.4.5 DNA Fingerprinting—Principles and Forensic Applications 699

15.4.5.1 Principles 699
15.4.5.2 Forensic Applications 700
15.5 Summary 716
References 716

15.1 INTRODUCTION TO EMERGING TECHNOLOGIES

The global advancement of environmental forensics has resulted in a proliferation of analytical and interpretative resources available to the forensic investigator. Although many environmental forensic techniques are mature and well established, many new and potentially significant methods have recently emerged. Three of these techniques (dendroecology, the use of atmospheric contaminants for age dating contaminants in groundwater, and microbiological methods) are presented to provide you with emerging tools available to you in environmental forensic investigations.

15.2 DENDROECOLOGY

15.2.1 INTRODUCTION

Dendroecology is the temporal study of forest changes depicted in tree rings. Tree rings often describe a pattern and history of growth suppression and release associated with specific time periods. An evaluation of tree rings can age-date events such as hurricanes, floods, fires, and other environmental stresses (Stokes and Smiley, 1996; Nash, 1999). Dendroecological research is extensive with over 800 papers dedicated to tree chemicals (http://web.utk.edu/~grissino/contents.htm and http://www01.wsl.ch/dendrobiblio/). Despite this body of literature, major forensic applications of dendroecology have been to age-date archeological sites (Baillie, 1982, 1995; Baillie *et al.*, 1995) or to reconstruct past climates (Fritts, 1971; Cook *et al.*, 1987; Fritts and Swetnam, 1989). Dendroecology draws upon associated scientific disciplines, including plant biology, physiology, chemistry, and pathology (McLaughlin *et al.*, 2002). The application of dendroecology for age-dating and contaminant source identification draws upon this same body of literature but with a different investigative purpose.

Dendroecological methods are commonly used to measure environmental stresses at regional or local scales (Eckstein and Krause, 1989; Banks, 1991; Cappellato *et al.*, 1998) such as:

- Acid rain (Cogbill, 1977)
- Atmospheric contaminants (Ashby and Fritts, 1972; Baes and McLaughlin, 1984; Boone *et al.*, 2004) such as SO_x plumes (Fox *et al.*, 1986; Wallner, 1998)



Figure 15.1
Wood structure. x-ray microdensitometric image of a maple core. Courtesy of Hakan Grudd, Dendrolab, Stockholm University.

- Anthropogenic lead enrichment in the environment (Watmough and Hutchinson, 2002)
- Urban city environments (Ragsdale and Berish, 1988; Jiang, 1996)
- Impacted surface waters, from both a contamination and flooding perspective (Faucherre and Dutot, 1976)
- Groundwater contamination (Vroblesky and Yanosky, 1990; LeBlanc and Loehle, 1993)
- Impacts to wetlands (Yanosky and Kappel, 1997)
- Impacts originating from wood preservation plants (Yanosky and Carmichael, 1993)
- Chlorinated solvent releases into soil and groundwater (Yanosky *et al.*, 2001)
- Site assessment investigations (EPA, 2005)
- Mining (Punshon *et al.*, 2003)
- Impacts from coal burning plants (Boone *et al.*, 2004)

Dendroecological methods have also been employed successfully by the police in criminal investigations (Baillie, 1982; Nash, 1999).

Dendroecology is a powerful forensic tool because of its ability to use trees as proxy-recorders, to identify when a contaminant entered a root system, resulting in its accumulation and preservation in the rings. Wood tissues (known as xylem) are formed annually by cell division within a thin layer called the cambium, which also forms the inner bark (known as the phloem). As xylem accumulates from the outer part of the tree, the inner portion (known as the pith) becomes the oldest ring. In gymnosperms (such as conifers) and angiosperms (for example, flowering trees), tree rings can be observed by contrasted colors or densities (see Figure 15.1). Tree growth is continuous from spring to fall, whereas during winter to spring, which is the dormant season, no new or very little xylem is formed. In temperate northern climates, the dormant season typically lasts from November to February. In tropical areas, the rainy season often dictates annual ring formation. As a consequence, each tree ring corresponds to one year. Early and late woods can be distinguished, by contrasted colors or location, allowing the subdivision of the tree rings into seasons.

The presence of contamination induces physical anomalies in the rings that can be identified. Such anomalies may be measured and the data compared with the chemical composition (spectroscopic methods can quantify elements in each ring). Thus, tree rings may provide a record of contaminant exposure

with a precision of a year, and sometimes even to a season (see Figures 15.3, 15.8, 15.9, 15.10).

15.2.2 METHODOLOGY

The principles governing tree growth and interpretation of ring width series are summarized by the following assumptions:

- The rate of plant processes are constrained by the environmental variable that is most limiting
- The pattern of a tree-ring series is dependent on several environmental factors over time where

$$R_t = A_t + C_t + D1_t + D2_t + E_t \tag{15.1}$$

where:

R_t = the measured tree-ring series
A_t = the age-size-related growth trend; outer most rings are narrower in old and large trees than in younger and smaller trees
C_t = climate
$D1_t$ = local stresses
$D2_t$ = stand-wide stresses
E_t = random processes not already accounted for or impacts that are not measurable but still contributes to tree-ring growth

- Matching patterns in ring widths or other ring characteristics (such as ring density patterns) among several tree-ring series allows for the identification of the year in which each ring was formed
- The environmental signal being investigated can be maximized, and the amount of noise minimized by sampling more than one core per tree, and more than one tree per site. Obtaining more than one core per tree reduces the amount of nondesirable signal impacting to each tree.

The first and second principles are important to the forensic investigator when a stress such as contamination ($D1_t$) impacts a tree and causes growth anomalies. The forensic method employs two steps: an evaluation of tree ring growth and chemical analysis of the tree rings (Balouet, 2005). The one-year precision may not always be possible, especially if ring(s) are missing or when cross-dating between live and dead trees is required.

Dendroecological methods have been successfully employed to estimate the time frames of contaminant releases impacting soil and groundwater. The use

of these methods to age-date chlorinated solvent or heavy metal contamination has been documented since the early 1990s. In particular, the United States Geological Survey (USGS) and other research institutions have published widely on this subject. More recently, the use of elemental markers has been employed to age-date petroleum releases (Balouet and Oudijk, 2006).

Several limitations of the method exist. The first limitation is the existence of trees within or above the contaminant plume, either in the soil or the groundwater. The tree must be older than the inception of the contamination although younger tree ring samples may provide the minimum age of the release. A second limitation is the availability of a control tree that is not impacted by the release.

Other limitations include, but are not limited to:

- The availability of trees proximate to a particular site, such as urban areas
- Landscaped areas where trees have been transplanted and do not reflect the on-site environmental conditions
- Trees with significant pruning or physical damage, which could depict significant stresses unrelated to contamination
- Double or missing rings, which reflect severe climatic stresses

15.2.2.1 Sampling Considerations

Samples should be collected from trees located within an area of known groundwater or soil contamination, as well as nearby trees of the same species (for comparison to unaffected control trees). Samples can be collected via chain-sawed slabs, or cores obtained with Pressler borers (see Figure 15.2). Most cores should be either 5 or 12 millimeters (mm) in diameter. The control tree sample should be of the same species as the trees sampled within the plume. Two or more cores per tree are recommended. When selecting the location of the cores, areas where branches or branch scars are visible on the bark, along with wound areas should be avoided. It is also recommended to orientate the cores toward the contaminant source.

Sampled trees should be located close to the contaminant source. Whenever possible, sampled trees should be healthy and absent of any significant trimming or damage.

15.2.2.2 Sample Preparation

Samples should be desiccated at ambient temperature, fixed on holding kits before they are planed and sanded (Balouet and Oudijk, 2006). This allows for observation of annual ring anomalies, optical reading of ring width, and to obtain the surface quality necessary for chemical microanalysis (see Figure 15.3).



Figure 15.2
Collecting a wood-core sample with a Pressler borer. Michael Obolensky of Triassic Technology, Inc. (Newton, New Jersey, USA). Slabs or cores are to be made on the stem at about four feet (1.3 m) above the ground surface. Coring needs to be performed perpendicular to the stem, and directed toward the pith (©Balouet).

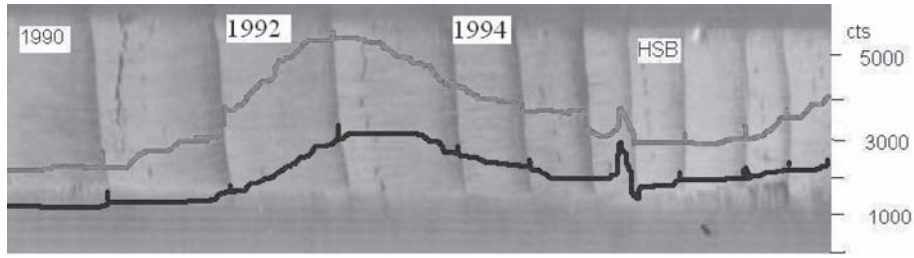


Figure 15.3
Dated tree ring series with associated chemical profiles for two elements showing contamination event dated 1992, and the consecutive growth anomaly dated 1994 (©Balouet).

15.2.2.3 Ring-Width Measurement

The measurement of tree ring widths allows a determination of growth anomalies caused by all environmental factors affecting the tree (such as contamination events, flooding, fires landslides, etc.) (see Figure 15.4). Comparison with a control tree (where contamination is not present) and confirmation of a particular contaminant effect by chemical analyses of rings will be further needed in forensic interpretations.

Tree ring widths normally are measured under direct light and with magnification equipment coupled with a measuring table (see Figure 15.5). However, ring-width measurements for some tree species or for sapwood can be difficult and other methods, such as transmitted light or microdensitometry techniques, may be needed (see Figure 15.1).

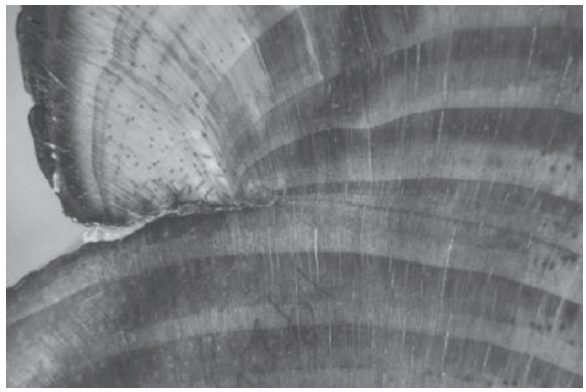


Figure 15.4
Dating a fire via tree-ring analysis. In this figure, the fire occurred five years before the tree was cut.

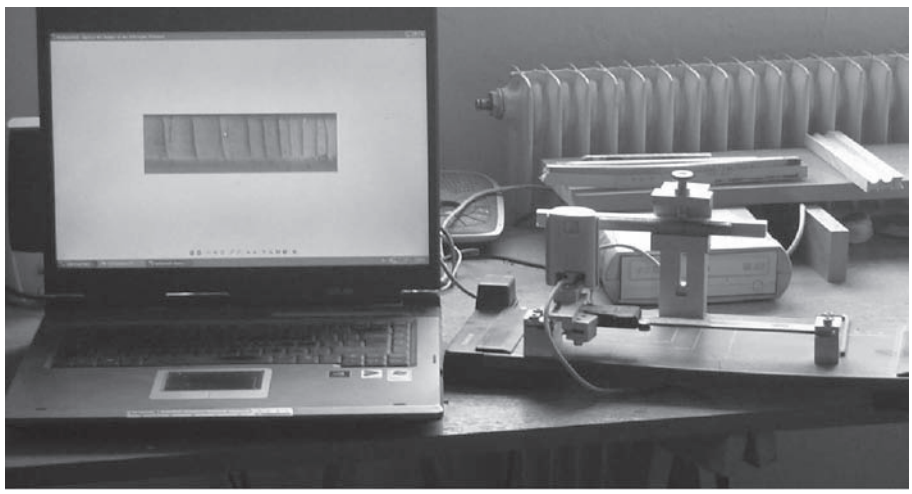


Figure 15.5
Ring width measuring equipment combining scientific camera, digital calliper, and visualization screen.

Ring widths can be measured with different precisions, typically between 1 micron for very slow growth trees, such as infra-millimetric rings, and 100 microns for faster growing trees. This, in general, complies with a precision of 20 to 50 microns, which usually meets the requirements for forensic applications (for example, ring widths of a couple mm or more, and a precision of around 1%). Once measured, the ring widths are tabulated and ring width series are verified, using pointer years defined by growth anomalies such as ring widths or intra-annual markers (by color or structure) (see Table 15.1). Figures 15.6 and 15.7 provide further illustration of how this information is then evaluated.

Annual variations are commonly observed and primarily depend on climatic pulses (At) when the age-size-related growth trend (Ct) of the aggregate model (Equation 15.1) remains, though expectedly altered by pollution pulses. To allow further interpretation of raw data, dendrologists use specific

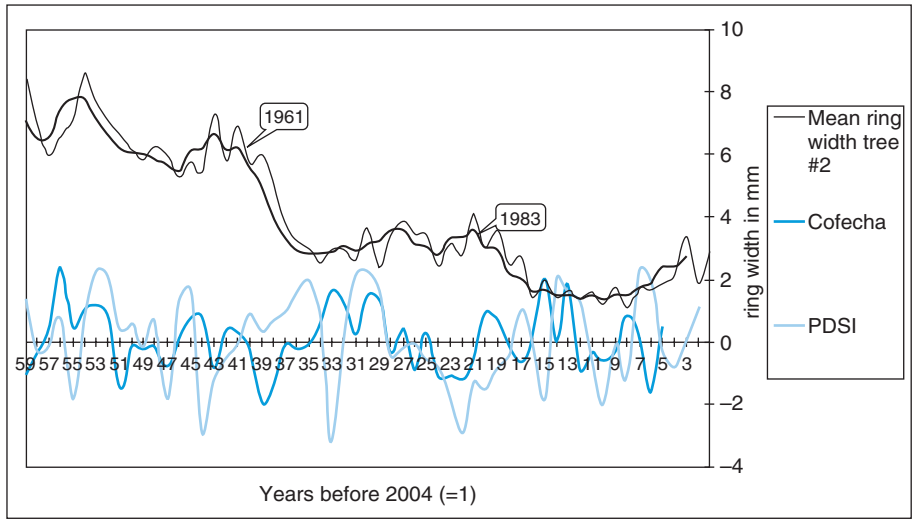
Table 15.1.

Example of a typical ring width series in millimeters (only a few years of this 60-year-old tree are presented). The 1999 Cofecha value of −1,653 is a negative, climate-related pointer year that is used for cross-dating.

Year	Tree 2 Core 2	Tree 2 Core 1	Tree 2 Core 3	Tree 2 Core 4	Tree 3 Core 2	Tree 3 Core 1	Cofecha Tulip Tree
2004	2,43	2,66	3,77	2,65	2,87	3,51	
2003	1,5	2,43	1,8	1,75	1,97	1,46	
2002	2,25	5,03	3,87	2,28	2,89	4,3	
2001	1,94	2,48	1,92	2	3,79	3,95	
2000	1,12	2,73	2,42	1,24	4,03	4,99	0,511
1999	1,44	1,98	1,9	1,31	3,3	3,46	−1,653
1998	1,31	1,79	2,33	1,28	4,26	5,33	0,081
1997	0,95	1,11	1,38	0,89	4,37	5,22	0,842
1996	1,56	1,94	1,85	1,58	4,68	6,87	−0,306
1995	1,05	1,53	1,23	0,98	2,83	2,68	−0,587
1994	1,42	2,14	1,63	1,22	3,17	2,69	−0,38
1993	1,15	1,46	1,65	1,07	3,14	2,05	−0,871
1992	1,06	1,38	2,02	2,13	3,56	2	1,828
1991	1,11	2,08	1,48	0,97	0,75	0,62	0,02

Figure 15.6

The mean of the four cores sampled from this tree reveals two growths anomalies dated 1961 and 1983. These anomalies refer to contamination events and not climatic pulses. The cofecha does not match PDSI climatic trends.



software, for example, the detrending and standardization software available from The Ultimate Tree Ring pages at http://web.utk.edu/~grissino/contents.htm.

Climate and age-size growth trend signals normally remain in an exposed tree ring series. It is essential to distinguish climate trends in exposed trees by comparing ring series to control samples or to information available from climate related sources such as PDSI Atlas or Cofecha databases (discussed later). To interpret ring-width data for given cores, the ring width anomalies that are attributable to climate growth trends (At) versus site disturbance (Ct)

is required. This is achieved by comparing the ring-width series to control trees, and to tree ring databases for the area (such as Palmer's Drought Severity Index (PDSI) or Cofecha (from the Spanish term *Fecha*, meaning time, moment) databases; these are relative numbers with no units). PDSI databases are available from the National Oceanographic and Atmospheric Administration's (NOAA's) web site at www.ncdc.noaa.gov/paleo/pdsiyear.html. PDSI indexes are calculated for each mapped location, based on multiple factors, including precipitation, water availability of soil, and dew point (thus temperature). Cofecha databases are based on indexed and standardized mean ring width over years, for a given tree species at specific locations.

Cofecha and PDSI databases reflect the mean impact of climate on tree growth, within a given area. The influence of climate is deduced from the growth trend. Positive numbers are correlated with positive growth trend, and negative numbers correspond to reduced growth. Because the data used in establishing such databases are indexed to exclude the influence of the individual tree age on ring growth and averaged from a large number of trees, Cofecha or PDSI can be used to identify pointer years (see Table 15.1 and Figure 15.6). The pointer years are then compared to the appropriate tree ring sequence, ensuring that there are no missing or false rings in the sequence.

Figure 15.6 depicts the multiple inflection points on the growth graph indicate that, despite an anticipated declining growth trend, other events caused two specific anomalies. Two distinct pollution events are observed: the first before 1961 and the second, possibly being a worsening of the release prior to 1983. The growth anomalies do not correlate to climatic factors. This supports the conclusion that the local pollution was the most limiting factor.

There are case-specific limitations to the use of growth trends (Nash *et a1.*, 1975). The first situation is when a contaminant impact coincides with negative growth caused by other environmental factors such as climate or disease. In some instances, growth decline caused by pollution may not be readily visible.

An alternative interpretation is that tree growth decline may appear only several years after uptake of the contaminant. The alteration of tree growth, even when caused by the same stresses, can vary by tree species.

The differences observed between particular (contaminated) tree growth series from the aggregated age-size-growth and climatic trends are impacted directly by the temporal and concentration gradients associated with the contaminant plume. If a sharp difference is noted, the release may have been severe (such as a spill or container/pipe failure, with significant amounts of pollutants dispersed in a short time).

15.2.2.4 Chemical Measurement of Tree Rings

The chemical measurement of tree rings is a recent development. Some early research used less precise methods, whereas more recent investigations did not fully consider the effects of internal biology on chemical patterns (Smith and Shortle, 1996). Analytical methods such as LIPS (Laser Induced Plasma Spectroscopy) or EDXRF (Energy Dispersive X-ray Fluorescence) can measure the concentration of different elements in each tree ring.

The prime tree elemental constituents are C, H, and O, accounting for about 98% of the xylem. Other elements, such as Ca, K, Mg, Mn, Fe, and Zn are important for tree existence, and usually recorded in the hundreds of milligrams per kilogram (mg/kg). Many other elements can be found, however, in trees in minute quantities, down to concentrations of less than 0.01 microgram per kilogram (μg/kg). An extensive bibliography that describes chemical compounds found in tree rings is available from the Swiss Federal Institute for Forest and its library, at http://www01.wsl.ch/dendrobiblio/, with an estimated minimum number of 800 publications recorded worldwide on the elemental chemistry of trees in relation to their environment.

15.2.2.5 Chemical Microanalysis

Chemical microanalysis of tree samples has been conducted with a variety of methods from classical wet chemistry to modern technologies such as gas chromatography coupled with mass spectrometry (GC/MS), electron microscopy, ion chromatography (Ferretti *et al.*, 1992), neutron activation, laser ablation-inductively coupled plasma mass spectrometry (LA-ICPMS) (Garbe-Schönberg *et al.*, 1997), proton induced x-ray emission (McClenahen *et al.*, 1989; Glass *et al.*, 1993; Harju *et al.,* 1996), or energy dispersive x-ray fluorescence (Gilfrich *et al.*, 1991; Lindeberg, 2004). EDXRF analysis has low detection limits (DL) (in the low mg/kg range) for elements with an atomic weight above 25, but it does not apply for very light elements (atomic weight less than 12) and has high DLs for elements with atomic numbers between 13 and 20 (DL in hundreds or thousands of mg/kg). EDXRF allows the selection of different sources to improve DL depending on searched elements. Counting times can also be adapted toward improving (lowering) detection limits as $1/\sqrt{\text{time}}$ (e.g., a nine-fold increase in counting time improves the detection limit by a factor of three).

Laser-induced plasma spectroscopy (Vaderno and Laserna, 2004) has been tested by the author. Sensitivity is in the mg/kg range, regardless of the atomic number (from 1 to 92). Laser ablation and plasma generation principles can attain low detection limits within one second. Amplified spectroscopic cameras used in laser-induced plasma spectroscopy provides the ability to distinguish

wavelengths of 0.02 nanometer (nm) in the visible electromagnetic spectrum (200 to 800 nm).

Energy dispersive x-ray fluorescence (Gilfrich *et al.*, 1991; Lindeberg, 2004) and laser-induced plasma spectroscopy are based on the irradiation of the sample surface followed by spectroscopy to measure the generated photon wavelengths and peak intensities. Both methods are performed at ambient pressure and there is no limitation in sample length. The samples are scanned and individual spot spectra are turned into elemental counts arranged in a spreadsheet.

For forensic investigations, the chemical line scanning of tree samples is the most appropriate approach, samples being analyzed at a given beam size, and for every given distance (increment). At common settings, beam sizes and increments are in the 100 μm, being able to be extended or diminished, as a function of mean ring widths, or expertise requirements. One hundred μm increments allow for scanning at 250 dpi (dot per inch), which is higher in definition than commonly computerized picture standards.

Criteria to be considered for selecting appropriate analytical methods include the following: it must not be sample destructive, and it should have a precision of at least one year, preferably up to a season. To accomplish this, the detection limits must be low (preferably in the mg/kg range), must allow for the detection of all elements or compounds at every measured spot, and should be compatible with the scanning of cores that commonly reach 8 inches (20 cm) without having to cut them in pieces. Vacuum is to be avoided as such treatment may eliminate some of the volatile or semivolatile compounds, and equipment should be able to run at temperatures below 0°C (for allowing analyses of frozen samples). Equipment availability and costs are critical. The equipment should allow for analysis of samples with limited to moderate preparation and should automatically generate spreadsheets of individual compounds that have been analyzed. Preferably, such equipment is also automated to reduce the analytical costs and allow for batch measurement of several samples.

15.2.2.6 Isotopic and Organic Compound Analyses of Tree Rings

In excess of 450 publications exist on the subject of isotopic analysis of tree rings. Over 300 of these articles refer to ^{14}C, however other radioisotopes such as ^{3}H, U or Pu have been measured (Edmands *et al.*, 2001). Stable isotopes of relevance include ^{2}H, N, S, Pb (Marcantonio *et al.*, 1998), oxygen, and carbon (Brenninkmeijer, 1983; Buhay and Edwards, 1993; Baillie *et al.*, 1995). Isotopes have been used to distinguish anthropogenic from natural sources of sulfur in trees. Most commonly $^{13}C/^{12}C$ or $^{15}O/^{18}O$ isotopic ratios have been evaluated for study of past climates. Other isotopes of relevance in environ-

mental forensics such as Cl may be further investigated. Unlike other EDXRF or LIPS methods, isotopic analysis is sample destructive. The amount of material needed is several cores or slabs, working on annual rings individually or combining several juxtaposed rings.

Most common volatile organic compounds, such as petroleum hydrocarbons, are readily metabolized or evaporated by trees after uptake through the root system. However, some stable compounds do exist, such as polychlorinated biphenyls (PCBs), which can be found in tree rings. In this case, the analytical methods are the conventional ones, based on wet chemistry. As in the case of isotopes, the method is sample destructive. Yet one slab or several core samples can prove sufficient for microanalysis while it is possible to preserve some sample for later purposes.

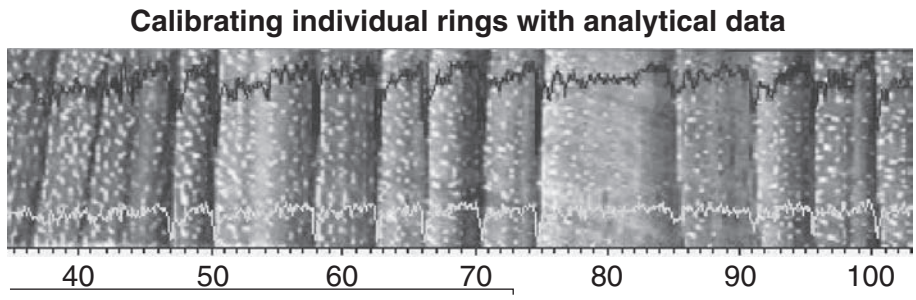


Figure 15.7
Calibration of annual rings and chemical data.

15.2.2.7 Calibration

The use of tree ring analysis for age-dating a contaminant release is dependent on correlating the chemical data to the ring year. As shown in Figure 15.7, there is some matrix effect relating to wood density, as shown by inter-ring anomalies. However, the inter-ring variations of total counts are by the percents, and amplitudes way below the intra-ring variation of some of the individually screened elements and way lower, several times fold, compared to the anomalies related to pollution events (see Figure 15.9). The marked inter-ring variation in EDXRF, coherent counts, of over 10% in this case, can be used for calibration of chemical data. For the wood section between 36 and 47 mm this check allows confirmation of false rings, as this core segment corresponds to one annual ring only, and not three as it may be first concluded from its intra-ring darker bands.

15.2.3 INTERPRETATION OF TREE RING DATA

Because tree physiology changes throughout the life of the tree, some gradual concentration changes can be observed between the pith and the outer rings. Annual fluctuations in chemical concentration often occur, depending on

tree species and the element or compound in question. Gymnosperms are good examples. The amplitude of annual fluctuation is often within a few tens of percent, and this amplitude remains basically constant over the heartwood, even when the tree is exposed to contaminants (see Figure 15.9). It is not uncommon for annual fluctuations to be of different magnitudes when comparing sapwood to heartwood.

Significant chemical anomalies can be found at the heartwood sapwood boundary (HSB) (Andrews and Siccama, 1995) and may be attributed to translocation mechanisms. The evaluation of such translocation mechanisms at the HSB is relevant when interpreting chemical profiles. The HSB is best measured on fresh core samples while in the field, and not always readily visible once the sample has dried.

The distribution of tree nutrients, such as Fe or Zn, can be influenced by environmental stresses at the stand. This often occurs when soils are acidified and the availability of ions is altered, a process known as ionic stress. Some other elements, such as potassium or calcium, unrelated to the contaminant plume, can also present time-correlated anomalies, and such anomalies can be caused by tree compensation mechanisms under the stress of pollution. In some instances, anomalies can exist for elements unrelated to the pollution, but consecutive to its release. Such anomalies can be attributed to contamination stress and caused by tree compensation mechanisms.

Alternatively, rigorously synchronous anomalies allow the characterization of the pollutant source. In Figure 15.8, the multielemental profiles allow us to determine that contaminant contained both S and P. Contamination related anomalies are of several magnitudes compared to background levels.

15.2.4 APPLICATIONS OF DENDROECOLOGY TO SITE ASSESSMENT

The chemical microanalysis of sapwood has allowed the mapping of chlorinated solvent plumes. Vroblesky *et al.* (1999) used tree-core analysis to delineate shallow groundwater contamination by chlorinated ethenes and in

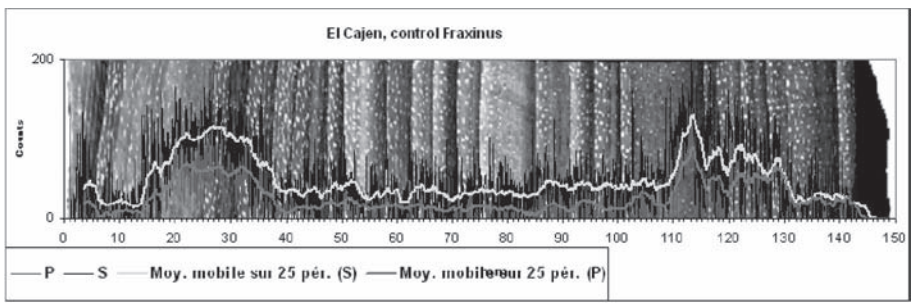


Figure 15.8
Potassium and sulfur profiles over a whole tree core.

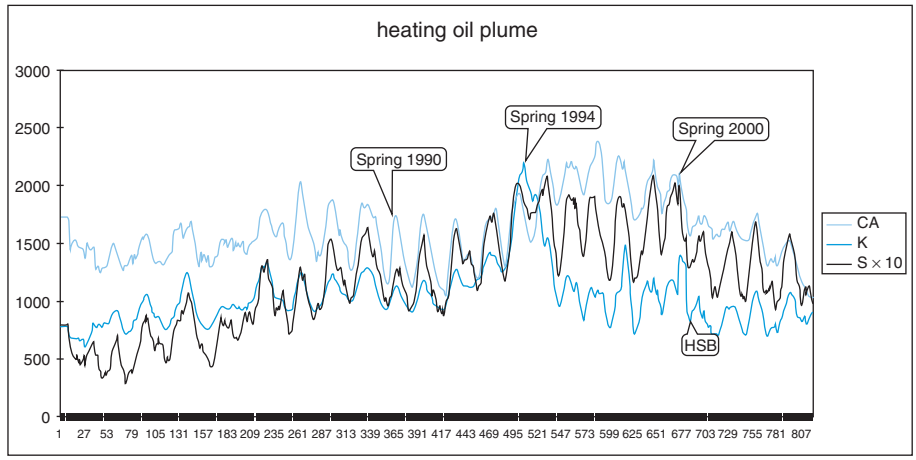


Figure 15.9
Heating oil impact on tree rings using sulfur as a marker.

general, obtained similar results as those obtained through soil and groundwater sampling. The trees grew over shallow groundwater contaminated with trichloroethylene (TCE) and *cis*-1,2-dichloroethylene (*cis*-DCE), a by-product of TCE degradation. Both *cis*-DCE and TCE were detected in the tree cores and their concentrations correlated with the plume present in the groundwater. Later on, USGS and the United States Environmental Protection Agency (USEPA) used the same method for subsurface chlorinated solvent contamination at the Riverfront Superfund site in New Haven, Missouri (Schumacher *et al.*, 2004; USEPA, 2005). The method proved successful in the delineation of a shallow tetrachloroethylene (PCE) and TCE plume in shallow groundwater. The results were used to design and focus the subsequent soil and groundwater investigation with substantial savings in time and cost.

15.2.5 APPLICATIONS OF DENDROECOLOGY IN ENVIRONMENTAL FORENSIC INVESTIGATIONS

Using trees as proxy-recorders has allowed the reconstruction of past pollution events. Using sulpher as a marker for heating oil (see Figure 15.9), the tank leak in this case has been progressive, ongoing from the very first years of site development, and worsening in 1994.

In another case (Balouet and Oudijk, 2006), a maintenance yard where 17 underground storage tanks containing petroleum hydrocarbon products were excavated in 1987, dendroecological investigation allowed characterization of three asynchronous contamination phases in a commingled plume

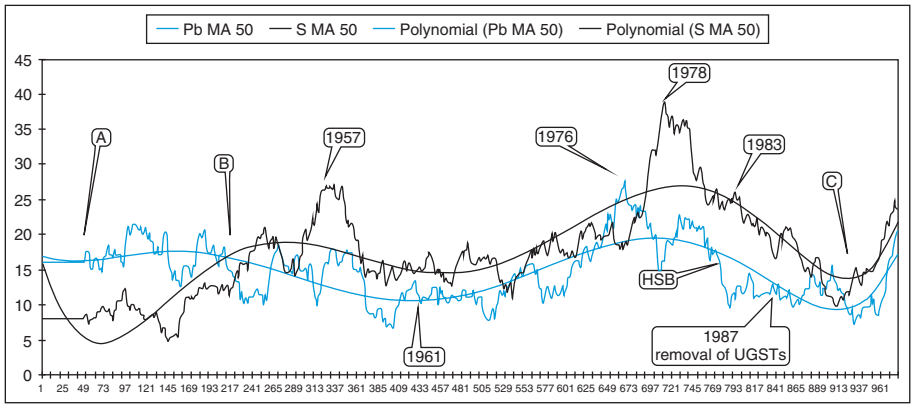


Figure 15.10
Comingled plume of sulfur and lead in groundwater depicted in tree rings.

(see Figure 15.10). A first event is dated 1957 and corresponds to heating oil, followed by a gasoline contamination as indicated by lead and dated 1976, with another marked sulpher anomaly dated 1978, representing a new leak of heating oil, or the worsening of the leak started in 1957. In this case, because of the time elapsed since inception and because the plume was discontinued by the nearby river, dendroecology has been the only applicable forensics method.

15.2.6 CONCLUSIONS

Dendrochronology, or the use of ring patterns to assess the ages of trees, has been in practice since the beginning of the 1900s and holds great potential for forensic applications. Without direct testimony or thorough record keeping, dendroecology may be the only method that can provide, as means to estimate the date of a contaminant with precision to within a year.

The tree-ring chronologies describe the pattern and history of growth suppression and release that can be associated with aging and natural disturbances. When contamination events occur, they are reflected in tree rings. A disturbance in growth is reflected in ring width and could be associated with the accumulation of a certain element or chemical in that particular ring. Subsequently, the simple measurement and chemical analysis of the tree rings may specifically point out the exact time when the contaminants entered the tree with a precision to the year and, in some cases, to the growth season. Wherever applicable, dendroecology can be used as an independent line of evidence in a forensic investigation.

15.3 USE OF ATMOSPHERIC CONTAMINANTS FOR AGE-DATING CONTAMINANTS IN GROUNDWATER

15.3.1 INTRODUCTION

A modification to an established approach of age-dating groundwater recharge via atmospheric compounds can be employed to estimate the minimum age of a dissolved contaminant plume in groundwater (Oudijk and Schmitt, 2000; Oudijk, 2003, 2005). As recharge passes through a zone of soil contamination, it will inherit that contaminant signature. By assessing the age of the recharge water, the minimum age of the contaminant may be estimated.

Traditional uses of atmospheric surrogates to age-date groundwater recharge include tritium (^{3}H), tritium/helium ($^{3}H/^{3}He$), krypton-85 (^{85}Kr), chlorofluorocarbons (CFCs), and sulfur hexafluoride (SF_6) (Thompson and Hayes, 1979; Smethie *et al.*, 1992; Cook and Solomon, 1995; Busenberg and Plummer, 2000; Cook and Herzog, 2000; Fulda and Kinzelbach, 2000). The groundwater age is defined as the time elapsed since the water entered the saturated zone and was isolated (through additional recharge). The age actually applies to the date of introduction of the chemical substance and not the age of the water (Glass, 2002). Atmospheric contaminants in groundwater also have been used for other purposes such as identifying sources of groundwater, whether river water or precipitation (Guay *et al.*, 2006), or characterizing arsenic contamination in Bangladesh (Klump *et al.*, 2006). These methods have been used by hydrologists—in particular, ground-water hydrologists—to investigate and explain ground-water flow and quality conditions (Philp, 2002). For example, ground-water flow paths, recharge dates, and residence times can be identified through these age-dating investigations. The important term used by hydrologists in this field is the apparent age, or simply the difference between sample date and recharge date (Burns *et al.*, 2003).

These atmospheric contaminants are also used to track groundwater or surface water flow paths (Cook and Böhlke, 2000). Because of their resistance to degradation, the tracers can also be artificially injected into surface and groundwaters to identify these pathways (Wanninkhof *et al.*, 2005).

The $^{3}H/^{3}He$ method is used to estimate groundwater recharge (Solomon *et al.*, 1993), estimate groundwater storage (Eriksson, 1958), infer groundwater flow directions (Cook and Böhlke, 2000; Tweed *et al.*, 2005), assess aquifer susceptibility to contamination (Nelms *et al.*, 2003), calibrate modeling studies of groundwater flow (Szabo *et al.*, 1996) and characterize plumes of sewage (Shapiro *et al.*, 1990).

For assessing the age of dissolved chemicals in groundwater, it is assumed that the age of the contaminant release can be estimated if a groundwater

sample is collected from the leading edge of the contaminant plume (both horizontally and vertically). Because recharge water migrates through the unsaturated zone and contaminants can dissolve into the water, the resulting impacted water will exhibit an age signature of tritium, or other atmospheric contaminants. Because the released chemical may degrade, disperse, or be adsorbed to the aquifer matrix, the estimated age is likely less than the actual age. However, longitudinal dispersion may have an offsetting effect on the calculated age of the impacted ground water.

To use this approach for age-dating a contaminant release, a sample must be collected from the leading edge of the contaminant plume. It is important that the sample does not contain water of significantly different ages. Because of the need for age-discrete sampling, some geologic formations will be more conducive to this method than others. Contamination within a fractured-rock system may be an inappropriate application of this approach due to the degree of mixing occurring within the fractures. Collecting a groundwater sample with the least amount of mixing is always desirable when using this approach (Weissmann *et al.*, 2002).

Valid estimates of contaminant age can be obtained only if the aquifer is not impacted by local sources of the atmospheric contaminant, the water sample is not impacted during collection, the water sample contains concentrations of the atmospheric contaminant that are representative of the amounts present in the aquifer, and the concentrations of the atmospheric contaminant are not modified by geochemical, biological, or hydrologic processes after entering the saturated zone (Busenberg and Plummer, 1992).

These methods have been used successfully to estimate the time frames of chlorinated solvent releases and leaks from underground storage tanks containing products such as home heating oil and gasoline with methyl-*tert*-butyl ether (MTBE). Case studies are provided in Oudijk and Schmitt (2000) and Oudijk (2003, 2005).

15.3.2 USE OF TRITIUM FOR CONTAMINANT AGE-DATING

Groundwater investigations frequently include the analysis of tritium (^{3}H), an isotope of hydrogen (Tolstikhin and Kamenskiy, 1970). The measurement of ^{3}H in surface and groundwater investigations has been employed since the late 1950s (Eriksson, 1958; Brown, 1961). ^{3}H was introduced into the hydrological cycle in the early 1950s through the explosion of hydrogen bombs (Schlosser *et al.*, 1988). Accordingly, after 1953, all atmospheric water contains some concentrations of ^{3}H. The input of ^{3}H to the atmosphere from thermonuclear explosions peaked around 1964 (see Figure 15.11). Because atmospheric water, through precipitation, percolates through the ground surface

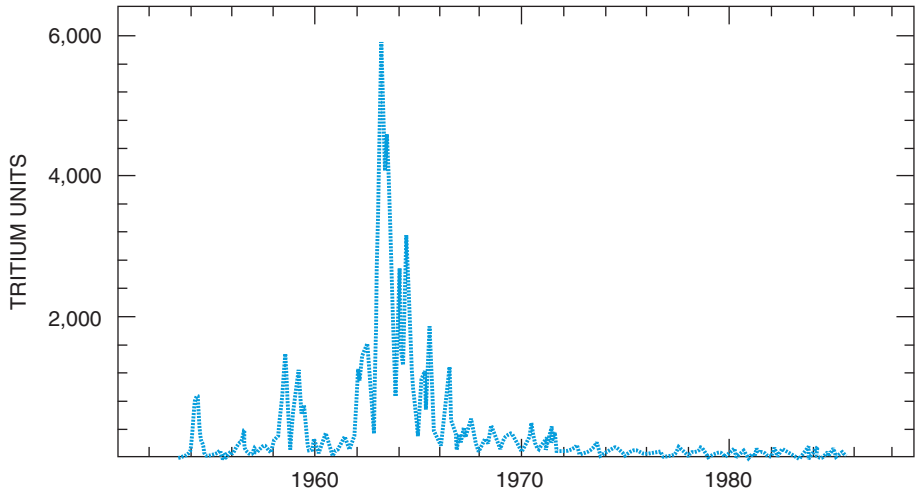


Figure 15.11
Tritium concentrations in the atmosphere versus time in Ottawa, Ontario, Canada. Redrawn from Cook and Herzog (2000).

and eventually recharges the water table, the presence of ^{3}H in groundwater relative to its decay products can be used to estimate the age of the water or when it was isolated within the saturated zone. The concentration of ^{3}H in the atmosphere has been tracked by governmental agencies (in the United States and worldwide) since the late 1950s through numerous monitoring stations such as the NOAA (Plummer *et al.*, 1993).

Atoms of ^{3}H are unstable and disintegrate radioactively into a helium isotope (^{3}He) at a half-life rate of 12.43 years (Mazor, 2004). The half-life is the time required for half of the isotope to decay into its daughter product. Therefore, after 12.43 years, 50% of the original ^{3}H is left and after 24.86 years, 25% of the ^{3}H is remaining. To determine the concentration of ^{3}H in the original rainwater, concentrations of ^{3}H and ^{3}He must be quantified. Because ^{3}He is prevalent in the atmosphere and is produced from the radioactive decay of uranium or lithium in rocks or rock fragments (Aeschbach-Hertig *et al.*, 1998; Schlosser *et al.*, 1989), the concentrations of the helium isotope (^{4}He) and neon (Ne) must also be quantified to determine the concentration of ^{3}He derived from ^{3}H. Based on known ratios of neon versus ^{3}He ($Ne/^{3}He$) and ^{4}He versus ^{3}He ($^{4}He/^{3}He$) in the atmosphere, the quantity of tritiogenic helium ($^{3}He_{tri}$) (or the ^{3}He derived solely from the radioactive decay of hydrogen bomb ^{3}H) can be determined (Takaoka and Mizutani, 1987).

Prior to 1953, rainwater and groundwater contained ^{3}H concentrations of less than 0.5 tritium units (TU) (one TU corresponds to one ^{3}H atom per 10^{18} atoms of hydrogen) (Clark and Fritz, 1997). Water containing ^{3}H concentrations of greater than 10 TU is considered to be of post-1953 age; water with ^{3}H concentrations of between 0.5 TU and 10 TU may be a mixture

of pre-1953 and post-1953 water (Schlosser *et al.*, 1989). For example, in 1963, precipitation on the Delmarva Peninsula, just south of New Jersey, contained ^3H concentrations of about 1600 TU (Szabo *et al.*, 1996). Because ^3H concentrations in the atmosphere are decreasing, use of the $^3H/^3He$ method is used more commonly where ^3He is the daughter product of ^3H decay. The use of ^3H and ^3He concentrations accounts for the more recent lower ^3H concentrations in the atmosphere and provides a more precise age estimate of the groundwater.

Because ^3H decays at a known and constant rate and because of the concentration of ^3H in rainwater, the age of the groundwater sample is estimated by quantifying the concentrations of ^3H and its daughter product, ^3He. The age of the groundwater is estimated by

$$t = (T_{1/2}/\ln 2)\ln(1 + {}^3He_{tri}/{}^3H) \qquad (15.2)$$

where t is the age of the water and $T_{1/2}$ is the half life of tritium (12.43 years) (Stute *et al.*, 1997). Methods for estimating values of $^3He_{tri}$ are provided in Schlosser *et al.* (1988) and Szabo *et al.* (1996).

There are two types of ages derived from these calculations: an *uncorrected age* and a *corrected age*. The uncorrected age assumes that all the $^3He_{tri}$ in the sample is derived from ^3H decay and from air. The corrected age accounts for the additional presence of terrigenic helium in the sample. It is normally assumed that the $^3He/^4He$ ratio of the terrigenic helium is 2×10^{-8} (if of radiogenic origin). The corrected age should be used only if the presence of terrigenic helium is indicated. If the percent of terrigenic helium in the sample is low, the uncorrected age and corrected age will be nearly identical.

In some cases, these calculations can provide a negative age. This is normally the result of some type of analytical and/or sampling problem. Loss of ^3He caused by inadequate confinement of the water sample following recharge can lead to anomalously young ages. This is often the case for samples collected near the water table. Fractionated samples can also exhibit negative ages. Small negative ages usually are interpreted as approximately modern. Limitations associated with the use of the ^3H method include the following:

- ^3H concentrations have been declining since a ban on atmospheric thermonuclear explosions was enacted. The Limited Test Ban Treaty was signed in 1963 by Britain, the United States, and the Soviet Union; countries such as France and the People's Republic of China did not sign the treaty. Accordingly, after about 1964, thermonuclear explosions lessened, but did not cease altogether.

- The method may not be applicable if the contaminant release occurred right at or below the water table.
- Use of the method in the general vicinity of nuclear power plants is not recommended because of potential, higher atmospheric ^{3}H concentrations.
- Some rocks with significant lithium concentrations possess ^{3}H concentrations not related to thermonuclear explosions. In most aquifers, this ^{3}H, known as geogenic ^{3}H, is at concentrations of less than 0.1 TU and would not be of concern (Clark and Fritz, 1997).
- Some igneous rocks, in particular granites, contain higher than usual ^{3}He concentrations, which may impact age-dating calculations.
- Use of the method may be invalid under extreme flow velocities or dispersion coefficients (Schlosser *et al.*, 1989).
- A phenomenon known as *fractionation* can occur, for example, if bubbles form during sampling. Gas bubbles can also occur naturally within the aquifer via denitrification or methanogenesis near the surface of the water table. Normally, in contaminant age-dating investigations, sampling will occur at some depth below the water table.
- Groundwater recharge in urban areas may not be uniform and, therefore, may impact the age-dating results. It would be simplistic to say that urban environments reduce recharge. Recent studies in Europe have revealed that urbanization may cause recharge to increase in some cases or decrease in others (Lerner, 1997; Barrett, 2004). Therefore, an understanding of the urban environment (underground utilities, foundations, hydraulic controls, etc.) would be needed to perform an age-dating investigation.
- Geologic and hydrologic conditions may affect the age calculated through ^{3}H concentrations. For example, deeper groundwater lacking in ^{3}H may flow upward into discharge zones causing a dilution of the ^{3}H concentration (Fritz *et al.*, 1991). Accordingly, a thorough understanding of the geologic and hydrologic conditions is needed to apply this method.
- All sampling work and transfers should be done outdoors with a specific ban on sampling personnel wearing wrist watches, particularly "glow in the dark" type watches. These watches contain tritium and may cross-contaminate samples. In addition, compasses or similar devices with luminescent dials, "Indiglow™," or so-called "beta" lights should not be worn. It is further recommended that sampling personnel not wear these types of devices for at least a week prior to sampling because the tritium is absorbed into the body (Clark and Fritz, 1997; University of Miami/RSMAS, 2003, 2005).

Information regarding appropriate sampling and analytical methods are available for helium and tritium, as described at http://water.usgs.gov/lab/3h3he/sampling/ or at http://water.usgs.gov/lab/cfc/.

15.3.3 USE OF CHLOROFLUOROCARBONS (CFCS) FOR AGE-DATING

The three types of CFCs present in the environment are CFC-11 (CCl_3F), CFC-12 (CCl_2F_2), and CFC-113 ($C_2CL_3F_3$). Chlorofluorocarbons (CFCs or Freons) are stable synthetic organic chemicals that were formulated in the United States since the early 1930s (Stoner *et al.*, 1997). Production of CFC-12 (dichlorodifluoromethane, CF_2Cl_2) began in 1931 followed by CFC-11 (trichlorofluoromethane, $CFCl_3$) in 1936 (Plummer and Busenberg, 1998). CFCs are used primarily as refrigerants, propellants, cleaning agents, solvents, and blowing agents. Their use in the United States and most European countries has declined because of restrictions imposed by the Montreal Protocols, although global atmospheric concentrations of CFCs continue to increase. Because CFC concentrations in the atmosphere over time are known and they are resistant to degradation, their concentrations in groundwater may be used to estimate the date of groundwater recharge (Busenberg and Plummer, 1992; Cook *et al.*, 1995).

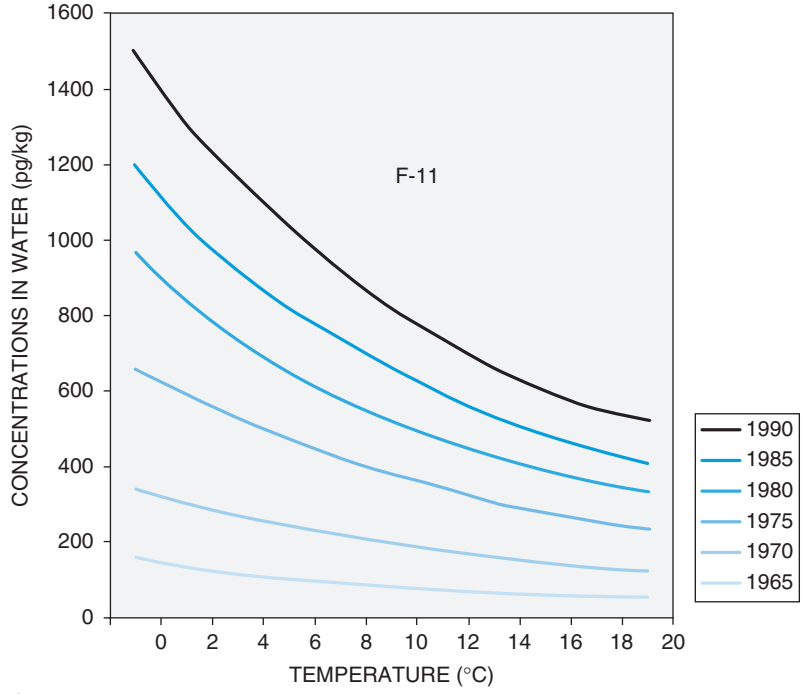
A comparison of the ratios between CFCs can provide insight as to whether the sample contains water of differing ages. CFC concentrations are then converted to atmospheric concentrations based on a known recharge temperature (Busenberg and Plummer, 1992). For example in the New York City area, the recharge temperature would be the average annual atmospheric temperature, or about 12˚C to 13˚C. The atmospheric concentrations are then compared to known published concentrations (see Figure 15.12). Busenberg and Plummer (1992) report that, in some cases within the unsaturated zones, the recharge temperature can be less than the mean annual temperature, because more recharge occurs during the colder months.

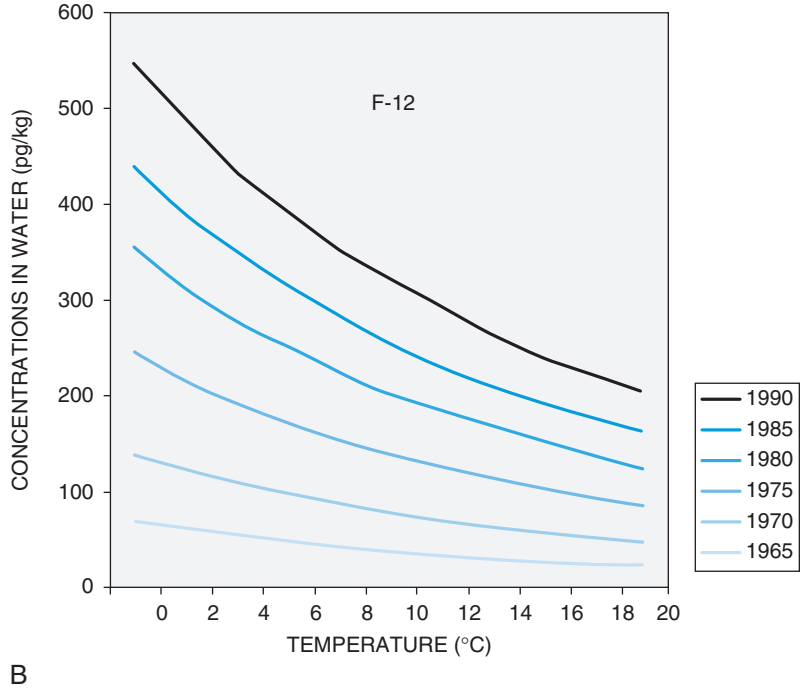
Limitations to the use of CFCs for age-dating a contaminant release include the following:

- CFCs are slightly degradable in groundwater under anoxic conditions (Oster *et al.*, 1996; Burns *et al.*, 2003). In particular, shallow groundwater and unsaturated zones exhibiting low oxygen contents have been found to be severely depleted in CFCs (Goode *et al.*, 1999). The United States Geological Survey performed studies on CFC degradation in surface waters and determined potential degradation rates (unpublished data provided at http://water.usgs.gov/lab/3h3he/research/norway/). However, based on research conducted in the United States and Europe, CFCs are resistant to degradation in aerobic groundwater systems (Ho *et al.*, 1998).
- CFC concentrations in the atmosphere are higher in industrial regions and highly populated areas (Ho *et al.*, 1998). Because CFC concentrations increase in the atmosphere, the age calculated with this method in an industrial region or highly populated area should be considered a minimum.

Figure 15.12

***A**. Concentration in the atmosphere of Freon 11 versus time and recharge temperature. Redrawn from Busenberg and Plummer (1992). **B**. Concentration in the atmosphere of Freon 12 versus time and recharge temperature. Redrawn from Busenberg and Plummer (1992).*





- In fine-grained sediments and fractured rock, the formation of air bubbles and their entrapment within the saturated zone can adversely affect the age calculations. Busenberg and Plummer (1992) report that these conditions typically will impact the age calculation by less than one year. To determine if excess air is an issue, analyses for nitrogen (N_2) and argon (Ar) are performed (Heaton and Vogel, 1981; Dunkle *et al.*, 1993). Nitrogen and argon concentrations can also be used to estimate recharge temperatures (Komor, 2002). An example from Australia of the problems associated with air bubbles in fractured rock and the associated apparent ages is described by Tweed *et al.* (2005).
- Knowledge of the recharge temperature is needed to estimate the groundwater age. Because of global climate change, calculation of these temperatures may become more difficult in the future. Colder winters may bias the recharge temperature downward, whereas warmer and heavier precipitation in the summers may bias the temperature upward (Ferguson and Woodbury, 2005). This limitation is also valid for tritium analyses.
- Groundwater in the vicinity of septic systems can contain abnormally high CFC concentrations due to the presence of CFC containing household plastics (Ho *et al.*, 1998).
- A thick unsaturated zone can cause a lag time between ages calculated with the ^3H and CFC methods (Johnston *et al.*, 1998).
- The uncertainty of age estimates provided by the CFC method increases with age because older recharge waters have lower CFC concentrations as well as a great possibility of CFC degradation (Stoner *et al.*, 1997).
- CFC-free sampling equipment must be used during the sample collection. Many soft, flexible plastics contain sufficient quantities of CFCs (Höhner *et al.*, 2003).

Sites most conducive to the use of this approach are those underlain by high-permeability, well-sorted, well-rounded sands; the least favorable sites are underlain by heterogeneous formations, such as glacial tills, interbedded sands, and clays or fractured rock. Thick unsaturated zones can also pose challenges due to the time needed to reach the water table. Furthermore, rural and undeveloped areas are preferred over urban or highly developed locations.

Care is required when interpreting test results to ensure that the sampling process does not bias the sample results. For example, the test results are considered invalid if bubbles are present. If bubbles are present in the sample, even though they were not detected during sampling, it is often considered as evidence that helium was stripped from the sample, potentially invalidating any possible interpretation of $^3H/^3He$ age.

15.3.4 ANALYTICAL METHODOLOGY

Samples analyzed for tritium are performed through electrolytical enrichment and scintillation counting. The precision is about ±0.6 TU. The sample is then analyzed for helium and neon with a noble gas mass spectrometer (Busenberg and Plummer, 1992; Solomon and Cook, 2000). The precision is normally between 1 and 2%. CFCs are analyzed by use of a purge-and-trap gas chromatography procedure with an electron capture detector (ECD). With ages younger than 1965, the uncertainty is ±1 year; for pre-1965 ages, the uncertainty may be as much as ±5 years.

15.3.5 LIMITATIONS AND CONSIDERATIONS

Considerations when using this approach include that the result provides a minimum age, and not an absolute age, for the contaminant release. Potential biodegradation of CFCs and the contaminant in question, dispersion of the plume, and adsorption to aquifer materials are factors that may influence the contaminant plume and the validity of the age-dating techniques. The age of the dissolved contaminants may be greater than the values calculated with these techniques. The effect of dispersion may present an offsetting effect although the magnitude is difficult to quantify.

The technique assumes that the leading edge of the plume can be identified. Samples collected at the water table may provide the age of water from the most recent precipitation. Accordingly, samples may need to be collected from the deepest portions of the plume, which may not always be possible. Monitoring wells with very short screens are needed to ensure that age-discrete samples are obtained. The short screen lengths may impart a low yield onto the monitoring well and prevent the collection of a sufficient sample volume. Furthermore, the low yield will impart greater drawdown and increase the possibility of air entering the pump and sample. The technique is invalid if a dense nonaqueous phase liquid (DNAPL) is present. DNAPLs are heavier than water, can migrate vertically through the aquifer and, for the most part, be unaffected by groundwater advection and dispersion. Therefore, samples collected under these circumstances would be biased to an older age. Accordingly, the contaminant release must be associated with a release of water whether it is precipitation or some other type of recharge. Sampling techniques with this method can be difficult and costly. Cross contamination of samples with the atmosphere is common and duplicate or triplicate sampling is highly recommended.

15.3.6 CONCLUSIONS

There are several contaminants present in the atmosphere that can be utilized to estimate the age of groundwater. These contaminants include ^{3}H and CFCs. Other atmospheric contaminants that can be used for the same purpose include SF_6 and ^{85}Kr. These methods are applicable to the age-dating of contaminant releases impacting groundwater. Through the use of short-screened monitoring wells, a thorough understanding of the geologic, hydrologic, and contaminant conditions and the laboratory analysis of ^{3}H and CFCs, the minimum ages of contaminant releases may be estimated, within the limitations described earlier.

15.4 DNA FINGERPRINTING/MICROBIOLOGICAL TECHNIQUES

15.4.1 INTRODUCTION

The use of microorganisms in environmental forensics depends on the existence of methods that are reproducible and specific enough to capture most of the microbial variability. Until recently, we were limited to those few microorganisms (about 5%) that can be cultured and grown in laboratory conditions. Today, this can be overcome by the advances in molecular microbiology reflected in the development of a series of methods usually called "DNA fingerprinting." These methods use the DNA structure from the DNA extracted directly from environmental samples (bypassing cultivation) to evaluate a microbial community's variability and to discern specific patterns and correlations. Such patterns are distinct in different environments, while for the same environment they may reflect the presence or absence of contamination.

DNA fingerprinting methods may be applied in two main types of environmental forensics investigations:

- Tracking the passage and source of contamination based on the changes in the structure of microbial groups (that can be observed at the DNA level) due to the presence of contamination (Flynn *et al.*, 2000; Macur *et al.*, 2004). The presence of a certain contaminant will influence the structure of microbial community from the environment where the contaminant is released. By extracting the DNA from that environment and creating a community DNA fingerprint, followed by comparison with the same environment without contamination, it may be possible to track changes due to certain contamination. The recorded changes (shifts) in microbial community will be trackable long after contamination has gone. One reason for this is that DNA persists in the environment after the death of microbial cells.

- Tracking the source of microorganisms in bioterrorism attacks. The existence of gene polymorphism between microbial species and within strains of the same species can be exploited to discern the exact source of a microorganism. Based on gene polymorphism, forensic methods and strategies for microbial identification can be developed and applied in the event of a bioterrorism attack or for any other forensic purpose, such as identifying the source of fecal pollution in water. More details on the methods and strategies used in bioterrorism forensics are not given here, but may be found in Petrisor *et al.* (2006).

15.4.2 DNA (DEOXYRIBONUCLEIC ACID)

Deoxyribonucleic acid (DNA) is the fundamental building block of all living cells. The characteristics, traits, and physical features of an organism are determined by the specific arrangement of DNA base-pair in the cell. DNA is a polymer consisting of a large repetition of monomer sequences. The monomer units of DNA are nucleotides. Each nucleotide consists of three components: (1) deoxyribose (a 5-carbon sugar), (2) a nitrogen-containing base attached to the sugar, and (3) a phosphate group. Deoxyribose and phosphate components are common (repeat) for all nucleotides, and the nitrogen containing bases may vary, being of four types. The nucleotides are given one-letter abbreviations (shorthand for the four bases): A for adenine; G for guanine; C for cytosine; and T for thymine. These bases belong to two main classes: purinic (A and G) and pyrimidinic (C and T). Their number is equal. It is this distinct arrangement of adenine, guanine, thymine, and cytosine that regulates the production of specific proteins and enzymes in a cell. Thus, DNA is the basic genotype (genetic identity) of an organism, which in turn determines the phenotype (physical features). Subsequently, particular DNA profiles can be ascribed to particular organisms. The DNA profile thus constitutes a unique fingerprint, specific to each individual (see *The Structure of DNA Double Helix,* from http://www.people.virginia.edu/~rjh9u/dnareq2.html).

A DNA molecule consists of two strands that are coiled around each other in a double helix. The two DNA strands run in opposite directions and form a helical spiral, winding around a helix axis in a right-handed spiral. The bases of the individual nucleotides are on the inside of the helix, stacked on top of each other like the steps of a spiral staircase; the phosphates are on the outside. The bases are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side. One of the pair must be a purine and the other a pyrimidine for bonding to occur. Only specific pairs of bases can bond together. These pairs are adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

The information responsible for the proper function of a cell and organism is stored into DNA, as a recipe in a recipe book. DNA encodes the information determining the protein synthesis in an organism. In this process of protein synthesis DNA is using another polymeric molecule (RNA, ribonucleic acid) to transfer genetic information.

15.4.3 PROTEIN SYNTHESIS AND RNA (RIBONUCLEIC ACID)

The sequence of the three nucleotides in the polymeric chain of DNA codes for an amino acid in the polymeric chain of proteins. The protein synthesis process involves the reading of the genetic code stored in DNA, which is accomplished by transcribing DNA into another molecule called RNA, or ribonucleic acid. The RNA is similar to the DNA molecule except it contains U (uracyl) as a base in the nucleotide structure instead of T (timine). The RNA inserts itself inside the DNA molecule that has partially unzipped to reveal the protein code (nucleotide sequence). It reads the code by matching its molecules with the corresponding DNA molecules in a process called *transcription*. With the exact code for a particular protein transcribed in its structure, the RNA proceeds to a cellular structure called the ribosome, which is the place where proteins are synthesized by putting together the amino acids according to the code within RNA. More amino acids continue to accumulate until the end of the RNA code is reached. There are distinct molecules of RNA that contain the transcribed information and those that bring and put together the amino acids into the protein chain. With the amino acids linked together, the end result is a protein.

15.4.4 PCR (POLYMERASE CHAIN REACTION)

PCR is a powerful technique, which effectively amplifies DNA. The technique was developed relatively recently (in 1987) by a young scientist named Kary Mullis. Dr. Mullis shortly afterward was awarded the Nobel Prize for his achievement and PCR was hailed as a monumental discovery.

The use of PCR requires tiny amounts of sample DNA, which makes it of great potential for forensic investigations usually limited by the amount of sample that can be recovered. PCR can produce multiple copies of DNA segments from an initial very limited amount of DNA (as little as 50 molecules), enabling a DNA fingerprint to be made from a single hair, for example. Generally, PCR is used to amplify a known sequence of DNA. Once the sequence to be amplified is selected, then the primers (small DNA snippets conserved in most species) are designed and flank the sequence of

interest that is to be amplified by the PCR reaction. Thus, PCR leads to the amplification of a particular segment (gene) of DNA. Another advantage of PCR is its simplicity.

Primers are small snippets of DNA homologous to different regions in a target gene. For any targeted DNA gene at least two primers are necessary to delimitate a DNA variable segment targeted to be amplified. This segment composed at each end by a primer and amplified during the PCR reaction is called *amplicon*. The selection of appropriate primers represents the key for a successful PCR amplification of a target gene. In order to design such primers, a collection of sequences for the target gene from many different genetic backgrounds is necessary. At least two conserved sequence regions must exist in the gene of interest in order to provide priming sites in genes from a broad range of organisms (Kitts, 2001). The primers should be spaced far enough apart for obtaining a multiplied DNA fragment with sufficient sequence divergence. At the same time, the primer regions should not be too far apart, because the long resulted amplicons may create, after digestion with restriction enzymes within DNA fingerprinting applications, fragments too large to be analyzed. An optimal amplicon length is between 400 and 700 base pairs (bp) (Kitts, 2001).

The PCR process requires three main steps that are repeated (each time doubling the sequence sampling):

(1) Denaturation of the sample DNA at 94°C by heating the DNA sample to unravel and split or unzip the double DNA helix.
(2) Annealing at 54°C—the selected primers, along with primase and polymerase enzymes are used to identify DNA sequences and produce a copy of them; ionic bonds are constantly formed and broken between the single-stranded primer and the single-stranded template. On the piece of double-stranded DNA, the polymerase attaches itself and starts copying the template.
(3) Extension at 72°C—the bases complementary to the template are coupled to the primer on the 3′ side. The polymerase adds dNTP's (nucleotides) from 5′ to 3′, reading the template from 3′ to 5′ side, bases are added complementary to the template.

PCR does not discriminate between DNA from living or dead organisms. Any organism in the sample can and will be analyzed, given that it contains the target gene (primer sites). This is of great advantage in forensic investigations since changes in microbial communities induced while contaminants are in contact with the microbiological community are visible in DNA fingerprints (of released DNA) after the contamination is no longer present and the living microbial community shifted back to normal conditions. Standard PCR

reaction is illustrated at http://avery.rutgers.edu/WSSP/StudentScholars/project/archives/onions/rapd.html.

15.4.5 DNA FINGERPRINTING—PRINCIPLES AND FORENSIC APPLICATIONS

15.4.5.1 Principles

The first technique of this type was developed in England by the geneticist Alec Jeffreys in the mid-1980s, who was looking at techniques to screen for hereditary diseases and genetic defects. However, when asked by the local constabulary to investigate a rape case, he successfully identified the rapist from his DNA profile. Within a few years, DNA profiling-based techniques were used by forensic scientists around the world.

Unique DNA sequences are present in all species and are used as biomarkers for the detection of cells from that species. These DNA sequences can most easily be detected using the polymerase chain reaction (PCR)-based DNA fingerprinting methods. All DNA fingerprinting techniques produce a pattern or profile of nucleic acids amplified (by PCR) from a sample and that pattern reflects the microbial community structure from an environment (Kitts, 2001). DNA fingerprinting or profiling comprises any DNA-based techniques that identifies the DNA from a certain individual or group of individuals within a community of organisms. The DNA fingerprints may be used as a tool for determining the identity of a specific DNA sample, or to assess the relatedness between samples. Most DNA fingerprinting methods are highly specific, highly sensitive, and largely independent on the physiological or growth state of the organisms, which make them of particular interest for forensic investigations.

The principle of these techniques consists of screening for certain DNA sequences that are found in some individuals, but not in others, thus visualizing DNA polymorphisms between samples and producing specific DNA fingerprints. Each individual has a DNA profile as unique as a fingerprint. In the case of humans, for instance, over 99% of all nucleotides are identical among all individuals. However, for every 1000 nucleotides inherited there is one site of variation, called polymorphism, in the population. These DNA polymorphisms induce the change in the length of the DNA fragments produced by digestion with restriction enzymes in the course of a fingerprinting technique. So, DNA polymorphism determines the resulting fragments to be of different length and/or sequence. Gel electrophoresis is used to separate and determine the size of the resulted fragments. This is because DNA is a charged molecule, so in an applied electric field it moves toward an electrode. The exact number and size of fragments produced by a specific restriction enzyme

digestion varies between individuals. The fragments are arranged in order of length and tagged with radioactive probes emitting x-rays, so when the sample is photographed they show up. This produces the fingerprint represented by a series of black lines corresponding to the DNA sequences present.

Many DNA fingerprinting methods are available. DNA fingerprinting techniques with potential applicability in environmental forensics are listed in Table 15.2. They differ methodologically using various techniques to look into the polymorphism of certain genes. All these techniques function based on the same principle described earlier, and they all result in producing specific patterns (fingerprints) of microbial communities in the environment.

DNA fingerprinting has been increasingly applied to determine the ancestry of plants, animals, and other microorganisms. Another application relates to tracking genomic mutations. Molecular diagnosis of complex inherited disorders, population screening of genetic diseases, studies of the genetic basis of variable drug response (pharmacogenetics), as well as discovery and investigation of new drug targets (pharmacogenomics) involving screening for mutations in multiple DNA samples are only a few examples of the DNA fingerprinting application pool. The primary applications of the DNA fingerprinting techniques described before are listed in Table 15.3.

15.4.5.2 Forensic Applications

Microoganisms respond to contamination, and such response may be recorded by DNA fingerprinting methods. The response to heavy metal contamination was well studied and could be the basis for forensic applications of microorganisms (Petrisor *et al.*, 2006). Biological processes are particularly sensitive to soil heavy metal loadings (Kelly and Tate, 1998). The negative impact of heavy metals results from their toxicity to biological processes. The population size of rhizobia has been found to decrease as a result of metal contamination. (Brookes *et al.*, 1986; Lorenz *et al.*, 1992; Chaudri *et al.*, 1993; Dahlin *et al.*, 1997). Several studies that used plate count techniques have demonstrated an increase in gram-negative bacteria (Doelman and Haanstra, 1979; Barkay *et al.*, 1985) and a shift in the composition of fungal species toward a more metal-tolerant community in metal-contaminated soils (Jordan and Lechevalier, 1975). Therefore, it is postulated that the soil microbial community could serve as an indicator of losses in soil quality due to heavy metal contamination (Kelly and Tate, 1998). Such indications may provide the basis for developing innovative forensic techniques for tracking metal contamination passage through the environment and identifying the source of contamination.

From different DNA fingerprinting methods presented, probably Terminal Restriction Fragment Length Polymorphism (TRFLP) holds the greatest potential for application in environmental forensics studies. Many laboratories

Table 15.2

DNA fingerprinting techniques with potential environmental forensics applicability.

Technique	Main Steps	Advantages	Limitations	References
Terminal Restriction Fragment Length Polymorphism (TRFLP)	DNA extraction. PCR amplification of the target gene with fluorescently labeled primers (only used if automated). Digestion of the amplified fragments with one or more restriction enzymes. Separation and visualization of terminal-labeled fragments (using automated capillary-based electrophoresis). Data analysis for peak identification (species identification in some cases).	Rapidity of analyzing large amounts of information due to available automated systems, generating hundreds of reproducible TRF patterns. A resolution previously unachievable. The potential to use TRF data to search existing databases for matching sequences and identification of individual organisms in a community profile. Ability for unprecedented opportunities to correlate the structure and diversity of microbial communities to the physical-chemical parameters of their surroundings.	A tendency of PCR to differentially amplify templates preventing quantification of relative abundance in community. Individual TRF detection is limited by electrophoresis. Patterns resulted from a single enzyme digest do not result in accurate community characterizations. Problems with incomplete digestion of amplicons creating artificial peaks in TRF patterns. A possibility of TRF length to overlap exists. The complexity of identification of the organisms responsible for a particular element in a profile due to the fact that TRFP are destructively sampled.	Cano and Borucki, 1995; Brunk *et al.*, 1996; van der Maarel *et al.*, 1998; Greenblatt *et al.*, 1998; Clement *et al.*, 1998; Liu *et al.*, 1998; Marsh *et al.*, 1998; Lambert *et al.*, 1998; Flyn *et al.*, 2000; Fey and Conrad, 2000; Osborn *et al.*, 2000; Leuders and Friedrich, 2000; Kitts, 2001; Kaplan *et al.*, 2001; Franklin *et al.*, 2001; Derakshani *et al.*, 2001; Park *et al.*, 2002; Dickie *et al.*, 2002; Horswell, *et al.*, 2002; LaMontagne *et al.*, 2003; Brodie *et al.*, 2003; Kaplan and Kitts, 2004

Table 15.2
Continued

Technique	Main Steps	Advantages	Limitations	References
Amplified Fragment Length Polymorphism (AFLP)	DNA extraction. DNA digestion with (usually two) restriction enzymes. Ligation of adapters to the restriction fragments. PCR amplification of the restricted fragments using primers complementary to each of the adapter sequences; selective nucleotides are added in the extension of restriction fragment at 3′ends determining the selectivity and complexity of the amplification. Separation and visualization of the resulted AFLP fragments.	The large number of polymorphisms that the method generates, being highly sensitive to polymorphism detection at the total-genome level. No sequence information is required. The PCR technique is fast. A high multiplex ratio is possible. Provides in general enhanced performance in terms of reproducibility, resolution, and time efficiency; can be automated.	Homology is perhaps the greatest problem in AFLP analysis. It is often assumed that comigrating bands are homologous. However, there is no *a priori* reason to accept this. Bias introduced by dominance, reproducibility problems, the effect of polyploids, as well as practical problems. The choice of either restriction enzyme or primer can affect the number of AFLP polymorphisms detected.	Petrisor *et al.*, 2006
Single-Stranded Conformation Polymorphism Analysis (SSCP)	DNA extraction. Amplification of DNA fragments by PCR using a phosphorylated and a nonphosphorylated primer for the target gene (usually rRNA). Denaturation of the amplified fragments to a single-strand form by exonuclease digestion. Separation of the denaturated amplified fragments, based on their polymorphisms by polyacrylamide gel-electrophoresis.	Polymorphisms at any of several sequence positions may be detectable by this technique, without requiring precise knowledge of the sequence polymorphism. Potential applicability to all unique sequences, both within genes and nongenic regions; if one amplified region does not show polymorphism, one can always move up- or downstream to another.	The amount of mobility differences is not correlated to the amount of sequence differences; thus, the only information that can be gained from SSCP is if PCR amplicons are "identical" or different. The optimal amplicon size for detection of most point mutations is rather small, involving complicated strategies to deal with this limitation.	Suomalainen *et al.*, 1992; Sarkar *et al.*, 1992; Sunnucks *et al.*, 2000; Xie *et al.*, 2002; Wagner, 2002; Melcher, 2003

Table 15.2
Continued

Technique	Main Steps	Advantages	Limitations	References
	Visualization of the separated fragments by silver staining or autoradiography and interpretation of the resulted pattern by hybridization with either DNA fragments or with RNA copies synthesized on each strand as probes.	PCR amplicons are smaller (200–300 bp) and thus easier to amplify and when polymorphisms are detected; the amplicons are of optimal size to sequence using an ABI automated sequencer.	Single-stranded DNA mobility is also dependent on temperature and pH. The fragment length may also affect SSCP analysis. For optimal results, DNA fragment size should fall within the range of 150 to 300 bp.	
Thermal and Denaturating Gradient Gel Electrophoresis (TGGE, DGGE)	DNA extraction. Amplification of DNA fragments by PCR using primers for the target gene (usually rRNA). Incomplete denaturation of the amplified fragments (using two main modalities as described before corresponding to DGGE and TGGE, respectively) and separation (based on fragment polymorphism) in gels containing a linear gradient of DNA denaturing agent. Visualization of the separated fragments.	Almost all single-base substitutions can be detected in longer amplicons; it is an ideal technique to search for rare variant alleles among individuals within a population. Applies best for finding intraspecific variation within one genus (insects) or family (vertebrates) of organisms. The resulting electrophoresis bands may be recovered from the gel for further analysis such as sequencing the specific DNA fragments of interest.	DGGE requires that one of the PCR primers include a "GC-clamp" (about 40 bp of G's and C's) and additional specialized equipment, which is expensive. DGGE does not scale up easily. DGGE requires the determination of optimal gradient conditions before screening individuals for variation.	Heuer and Smalla, 1997; Bruggemann *et al.*, 2000; Viglasky *et al.*, 2000; Petrisor *et al.*, 2006.

Table 15.2
Continued

Technique	Main Steps	Advantages	Limitations	References
Amplified Ribosomal DNA Restriction Analysis (ARDRA)	Cultivation of the targeted microorganisms → microbial isolates. DNA extraction from microbial isolates. DNA amplification by PCR, using 15S rRNA universal primers. Digestion of amplified DNA fragments by endonuclease restriction with one or more restriction enzymes (selection of restriction enzymes should be on the basis of simulated digests of the complete 15S rRNA gene sequences of chosen reference strains). Separation and visualization of the digested DNA fragments based on their polymorphisms, by agarose gel electrophoresis (gels stained with ethidium bromide). DNA was visualized by transillumination with UV light after the gels were stained with ethidium bromide.	The automated culture technique, in combination with ARDRA provides accurate, practically applicable, and wide range identification of mycobacterial species. The existing identification library covers most species and can be updated easily when new species are studied or described.	Culture-dependent. Unsuitable for typing analysis, as it overlooks intraspecific differences.	Massol-Deya *et al.*, 1995; Seifert *et al.*, 1997; De Ventura *et al.*, 2000; Gich *et al.*, 2000; Baere *et al.*, 2002; Rahman *et al.*, 2003.

Table 15.2
Continued

Technique	Main Steps	Advantages	Limitations	References
Randomly Amplified Polymorphic DNA (RAPD)	DNA extraction. Amplification of DNA fragments by PCR using randomly chosen primers (about 10 bp) at low annealing temperatures → amplification at multiple loci. Separation of the amplified fragments based on their polymorphisms by agarose gel electrophoresis (two-dimensional gel). Visualization of the separated fragments. Possible cloning and sequencing of the DNA band of interest → develop longer specific PCR primer pairs targeting reproducible amplification and detection of the differences.	The relative speed (a complete analysis in one day) and ease with which results can be obtained. Does not include expensive techniques such as cloning and sequencing. More rapid and easier to perform as compared to ribotyping and pulsed-field gel electrophoresis. Useful to analyze large amounts of strain collections from a monitoring standpoint. Cloning of RAPD fragments can be rapidly accomplished and cloned. RAPD loci will not contain repetitive sequences.	Lacks specificity, due to low annealing temperatures and easier reaction conditions. A polymorphism detected between one pair of strains may not translate into use for another pair of strains. On average, only half of the RAPD polymorphisms detected between two strains would be mapped. Most of RAPD markers are dominant and this makes it difficult to use the data in population genetics studies. RAPD is sometimes not reproducible.	Shianna *et al.*, 1998; Benecke, 1998; Krane *et al.*, 1999; Romalde *et al.*, 2002; Dergam *et al.*, 2002; Soliman *et al.*, 2003; Manners *et al.*, 2004.
Restriction Fragment Length Polymorphisms (RFLP) (non-PCR technique)	DNA extraction. Digestion of DNA by endonuclease restriction. Separation of the resulted fragments by agarose gel-electrophoresis. Transference of the fragments to a nylon filter by Southern blotting. Hybridization of fragments to a locus specific radio labeled DNA probe.	Higher degree of polymorphism and the possibility of obtaining an infinite number of probes, evenly distributed over the whole genome. High stability and reproducibility, giving constant results over time, and location. High genomic abundance. Random distribution throughout the genome.	The results do not specifically indicate the chance of a match between two organisms. The process is very expensive, laborious, and technically demanding. Large quantities of purified, high molecular weight DNA required (5–10 μg). Additional development costs in case suitable probes are unavailable.	Petrisor *et al.*, 2006

Table 15.2
Continued

Technique	Main Steps	Advantages	Limitations	References
	Visualization of the polymorphisms by autoradiography.	Band profiles interpretation in terms of loci and alleles. Codominance of alleles. Producing of semidominant markers, allowing determination of homozygosity, or heterozygosity.	Not amenable to automation. Long methodology before results are gained. Many probes not available depending on species. Too many polymorphisms may be present for a short probe.	
Ribotyping (with or without PCR)	Cultivation of targeted organism (optional step). DNA extraction. Digestion of extracted DNA with restriction enzymes. Analysis of the digested DNA fragments by Southern blotting or PCR followed by an electrophoresis gel. Discriminant analysis of the resulted banding patterns for strain identification.	Availability of an automated ribotyping system, the DuPont Qualicon RiboPrinter, which provides highly reproducible and standardized ribotype data. Uses ribosomal DNA, which is very stable in a population of bacteria, providing a limit to genotype variability that may overcome geographical restrictions. Routinely analyses inexpensive ($375–600 per sample).	A lack of standardization. High cost per isolate. Discrimination is dependent upon the restriction enzyme and the probe used. Time-consuming procedure. A need for highly skilled technical staff. Requires access to specialized equipment.	Ge and Taylor, 1998; Parveen *et al.*, 1999; Carson *et al.*, 2001; Oliveira and Ramos, 2002; Scott *et al.*, 2002, 2003, 2004; Fontana *et al.*, 2003.

Table 15.3

Forensic and other applications of DNA fingerprinting techniques.

Technique	General Application	Specific Applications
Terminal Restriction Fragment Length Polymorphism (TRFLP)	Ability to characterize the microbial communities from different environments, allowing the determination of spatial and temporal shifts in community structures—potential application in environmental forensics.	Studies focused on *Eubacteriaceae* (Brunk *et al.*, 1996; Liu *et al.*, 1998; Flynn *et al.*, 2000; Franklin *et al.*, 2001). Characterization and comparison of subsurface and surface soil bacterial communities in California grassland (LaMontagne *et al.*, 2003).
	Used in criminal forensics to link different soil samples (from crime scene and sole of a shoe or clothing).	Discerning bacterial population changes in feces during feeding rats with *L. acidophilus* NCFM (Kaplan *et al.*, 2001). Characterization of the debrominating methanogenic consortia.
	Able to characterize functional diversity in bacterial communities, including reports on the functional diversity of important processes such as nitrification, nitrogen fixation, denitrification, and mercury resistance investigated in a large variety of environments from marine sediments to termite intestines.	Characterization and comparison of bacterial communities in deer fecal pellets, petroleum hydrocarbon-contaminated sands, and pristine sand (Clement *et al.*, 1998). Observation of taxonomic diversity in other microbial groups, such as archaebacterial communities in soil and fish intestines (van der Maarel *et al.*, 1998; Fey and Conrad, 2000; Leuders and Friedrich, 2000). Revealing *Nitrosospira*-like organisms as one of the major contributors to ammonia oxidation in a full-scale aerated-anoxic Orbal reactor (Park *et al.*, 2002). Fungal community studies—to point alterations in soil micro-fungal community structure across a transect between a seminatural upland grassland and an agriculturally improved enclosure (Brodie *et al.*, 2003). Establishing the source of bacteria isolated from amber and age-dating amber (Cano and Borucki, 1995; Greenblatt *et al.*, 1998). Investigation of bacterial community structure and dynamics during a bioremediation experiment in a land treatment unit (LTU) established at a site contaminated with highly weathered petroleum hydrocarbons at the Guadelupe Oil Field (central California Coast) (Kaplan and Kitts, 2004).

Table 15.3
Continued

Technique	General Application	Specific Applications
Amplified Fragment Length Polymorphism (AFLP)	Used for "basic" diversity and genetic variation studies. Reliable and efficient method for detecting molecular markers. May be used for defined applications such as polymorphism screening, QTL analysis, or genetic mapping. An efficient and reliable technique for evolutionary studies. Applications in paternity analyses.	Investigation of the genetic variation of *Calycophyllum spruceanum* (Rubiaceae), a fast growing pioneer tree of the Amazon Basin (Russell *et al.*, 1999). Investigation of introgression and hybridization (Beismann *et al.*,1997; Rieseberg *et al.*, 1999). Generation of a large number of markers that are linked to target genes (Xu *et al.*, 1999). AFLP markers have been used at the individual level for application in paternity analyses and gene-flow investigations (Krauss and Peakall, 1998; Krauss, 1999). AFLP markers are used in phylogenetic analyses (Heun *et al.*, 1997; Kardolus *et al.*, 1998; Aggarwal *et al.*,1999; Mace *et al.*, 1999). Frequently applied to mapping studies (Bradshaw *et al.*, 1998; Zhu *et al.*, 1998; Xu *et al.*, 1999). Gene-flow experiments and for plant variety registration (Law *et al.*, 1998).
Single-Stranded Conformation Polymorphism Analysis (SSCP)	A powerful tool for the analysis of microbial community diversity in environment, with a large pool of potential applications in microbial ecology, environmental biotechnology, and environmental forensics. Can detect DNA polymorphisms and mutations at multiple places in DNA fragments. Can be used in cloning studies.	Analysis of microbial community diversity in environmental samples (Wagner, 2002). Sensitive in determining genetic variation (Sunnucks *et al.*, 2000). Effective and rapid search for point mutations or polymorphisms in mitochondrial DNA (Sarkar *et al.*, 1992; Suomalainen *et al.*, 1992). The analysis of polymorphisms at single loci, especially when used for medical diagnoses (Sunnucks *et al.*, 2000).

Table 15.3
Continued

Technique	General Application	Specific Applications
		Identification and cloning of nine aquaporin genes in cucumber (Xie *et al.*, 2002).
Thermal and Denaturating Gradient Gel Electrophoresis (TGGE, DGGE)	Popular tool in microbial ecology. Ideal for analyzing complex bacterial communities. Able to detect mutations and in general genetic variations. DGGE is highly efficient for detection of DNA sequence differences. DGGE is a convenient tool for analyzing community shifts that involves only a small region (~200–300 bp) of the 15S rRNA gene. Monitoring microbial communities with potential in Environmental Forensics studies.	Detection of actinomycetes in the environment, monitoring actinomycete community changes in potato rhizosphere, and investigation of actinomycete diversity in different soils (Heuer *et al.*, 1997). Monitoring microbial communities in metal contaminated environments (Petrisor *et al.*, 2006). TGGE technique was used as a tool for detecting various conformational modifications of plasmid DNAs (Viglasky *et al.*, 2000).
Amplified Ribosomal DNA Restriction Analysis (ARDRA)	Successful in species differentiation but not strain differentiation as it relies on the conserved nature of rDNA. Able to evaluate microbial community structure changes—potential for environmental forensics applications.	Rapid and unambiguous species differentiation within the genus *Acinetobacter* (Seifert *et al.*, 1997). Reliable and rapid method for identifying *Lactobacillus* species from intestinal and vaginal microflora at species and subspecies level (Ventura *et al.*, 2000). Detection and identification of all mycobacterial species at once, instead of attempting direct PCR-based detection from clinical samples of *M. tuberculosis*. Characterization of thermophilic bacteria associated with the subsurface soil environment in Northern Ireland (Rahman *et al.*, 2003).

Table 15.3
Continued

Technique	General Application	Specific Applications
		Detection of differences in activated sludge bacterial communities fed on domestic or industrial wastewater, and subjected to different operational conditions (Gich *et al.*, 2000).
Randomly Amplified Polymorphic DNA (RAPD)	Typically used for genetic mapping and studies of population genetic structure. Able to produce a biochemical fingerprint of a particular species. Capable of screening the differences in DNA sequences of two species of plants. Rarely used to infer phylogeny, as it is difficult to know if bands of the same size are homologous across species. Applied in criminal forensic studies and potential for environmental forensics.	Comparison of genetic material from four females and four unknown male trees of Egyptian date palm (Soliman *et al.*, 2003). Determining the identity and relationships of old roses (Manners *et al.*, 2004). Determining whether variations exist among bovine strains isolated from a localized geographic area (watershed of the Red River) (Shianna *et al.*, 1998). Testing for genetic differentiation in fish populations within Rio Dolce basin of southeastern Brazil (Dergam *et al.*, 2002). Correlation of RAPD profile-based measures of genetic diversity in crayfish with environmental impacts (Kranee *et al.*, 1999). Used for quick identification of arthropods and support classical morphological and medico-legal analysis of maggots on a human corpse (Benecke, 1998).

Table 15.3
Continued

Technique	General Application	Specific Applications
Restriction Fragment Length Polymorphisms (RFLP) (non-PCR technique)	Able to identify the origins of a particular species, with the aim of mapping its evolution. Able to characterize a large number of individuals, given preexisting knowledge of the sequence variation.	Using the sequence data of three species of nematodes (from a particular genus) to determine a simple way to diagnose the species based on restriction digestion (Petrisor *et al.*, 2006). Survey of a large sample of individual nematodes (even life history stages that might not be morphologically distinguishable) and enumerate the species composition of the sample. Gased Ribotyping) (Oliveira and Ramos, 2002)
Ribotyping (with or without PCR)	Able to be used in subtyping of foodborne pathogens and of microorganisms important for fermentation and food spoilage. Excellent tool for indicating major sources of biological pollutants. Effective tool for bacterial source tracking (BST), e.g., tracking the source of fecal pollution in water. Effective tool for discriminating between human vs. nonhuman forms of *Escherichia coli.*	Tracking *E. coli* that are regrowing and proliferating in a watershed. Resolving a "true-life" water quality and management problem consisting in tracking sources of fecal pollution in a watershed in South Carolina (Scott *et al.*, 2004). Differentiation between human and nonhuman sources of fecal pollution (generally *E. coli*) in water (Parveen *et al.*, 1999). Identification of individual host sources of fecal *Escherichia coli* and distinguishing between fecal *E. coli* of human and nonhuman origin (Carson *et al.*, 2001). Differentiation of *E. coli* between human and nonhuman sources when applied to organisms isolated from a large geographic region (Scott *et al.*, 2003). Automated ribotyping and pulsed-field gel electrophoresis were used for the rapid identification of multidrug-resistant *Salmonella* serotype Newport (Fontana *et al.*, 2003).

routinely use TRFL patterns to characterize the microbial communities from different environments, allowing the determination of spatial and temporal shifts in community structure. Such shifts may be used in forensic studies as evidence for the presence of certain contaminants through an environment, or to track their past passage. Also, the toxicity of contaminants to bacterial communities, coupled with their use as nutrient sources (such as oil that may be used by bacteria as carbon source) are likely selective pressures on bacterial community composition and diversity. Thus, the characterization of such microbial diversity and its response to contamination through simple automated techniques such as TRFLP could help predict the fate of contaminants in the environment.

TRFLP applications in criminal forensics generally based on significant differences between DNA profiles of different soils are well established. Thus, by setting up mock crime scenarios, Horswell *et al.* (2002) established that bacterial community DNA profiling can be used to compare soil samples of the size likely to be encountered in real forensic situations. In the first scenario, the shoe print from a hypothetical crime scene was compared with soil from the scene, and in a second scenario, the jeans were impregnated (knee level) with soil from the scene. A simple comparison index, the Sorenson's Index, was used to compare profiles, in which a value of 1 indicates identical profiles, and 0 indicates that the profiles share no common fragment sizes (peaks). The microbial community profiles for soil from the shoe print and soil collected from the shoe itself produced very similar electropherograms and a very high (>0.9) similarity index. In contrast, there were major differences between the profiles of the reference soils and those of the crime scene and the suspect's shoe, with a similarity index of <0.6. When the soil sample collected eight months later (A3) was compared with the microbial community DNA profile of soil collected at the time of original sampling (A2), differences were observed, but overall, there was a considerable similarity between the profiles and a similarity index of 0.70 was obtained. In the second scenario, the microbial community profiles of the soil from the left knee impression in the soil and the soil extracted from the left knee of the jeans were very similar with a Sorenson's index of 0.82.

Several environmental studies have successfully used the TRFLP method to study and monitor natural remediation processes. TRFLP techniques were used to investigate bacterial community structure and dynamics during a bioremediation experiment in a land treatment unit (LTU) established at a site contaminated with highly weathered petroleum hydrocarbons (C_{10}-C_{32} range) at the Guadalupe Oil Field (central California Coast) (Kaplan and Kitts, 2004). The different shifts in the bacterial community structure were linked to the degradation rate of petroleum hydrocarbons. Thus the TRF patterns

revealed a series of sample clusters describing bacterial succession during the study, and the data suggested that specific polytypes of bacteria are associated with different phases of petroleum degradation in the land treatment unit. Such findings supported the possible use of TRFP technique to monitor, better understand, and improve the bioremedial treatment. On the other hand, the sensitive insight into microbial community generated may allow association with different contaminants and sources in a forensic investigation.

The net effect of naturally occurring oil on marine sediment bacterial communities was also evaluated at the Guadalupe Oil Field. Sediment piston-cores were collected at an active marine hydrocarbon seep located in the Coal Oil Point area at the depth of 66 m in the Santa Barbara Channel, and community profiles were determined by TRFLP of 15S rDNA genes PCR-amplified using eubacterial primers. TRFLPs obtained by restriction with *Hha* I suggested that the bacterial communities collected in the center of the seep are distinct from those located 10 meters outside the area of active hydrocarbon release. Additionally, diversity indices calculated from the number of peaks were the lowest in surface sediments (0–2 mm) collected in the center of the seep. These consistent differences in surface sediment communities suggested a negative effect of hydrocarbons on bacterial diversity and the existence of a unique bacterial community within hydrocarbon seep fields.

Amplified Fragment Length Polymorphism (AFLP) is generally used for "basic" diversity and genetic variation studies. The method is also reliable and efficient for detecting molecular markers. AFLP markers have been used at the level of the individual, for application in paternity analyses and gene-flow investigations; for example, Krauss and Peakall (1998) analyzed paternity in natural populations of *Persoonia mollis* (Proteaceae), a long-lived fire-sensitive shrub from southern Australia. Another application in paternity analysis was described by Krauss (1999).

Due to its methodological simplicity and sensitivity, Single-Stranded Conformation Polymorphism Analysis (SSCP) is a powerful tool for the analysis of microbial communities with a large pool of potential applications in microbial ecology and environmental biotechnology, as well as in environmental forensics. SSCP analysis offers an inexpensive, convenient, and sensitive method for determining genetic variation (Sunnucks *et al.*, 2000). The method can be applied for the cultivation and independent analysis of microbial community diversity in environmental samples, based on PCR-amplified small subunit (SSU) rRNA gene sequences from directly extracted DNA. Under optimal conditions, approximately 80 to 90% of the potential base exchanges are detectable by SSCP (Wagner, 2002).

Analysis of PCR amplified 15S rDNA fragments from environmental samples by denaturing gradients of chemicals or heat within polyacrylamide gels is a

popular tool in microbial ecology (Bruggemann *et al.*, 2000). These techniques are ideal for analyzing complex bacterial communities. A thorough analysis of microbial communities on several levels ranging from the community structure of dominant populations using conserved sequence primers, to phylogenetic sequence analysis of single bands from individual community members is generated. Similar to SSCP, Denaturing/Thermal Gradient Gel Electrophoresis (D/TGGE) techniques are used successfully to detect mutations and in general genetic variations, and hold great potential for environmental forensics use by monitoring microbial communities in contaminated environments. Thus, TGGE was used effectively to monitor bacterial communities in soil (http://risk.kan.ynu.ac.jp/rmg/WS99PDF/99ITOH.PDF). These methods monitored actinomycete community changes in potato rhizosphere and investigated actinomycete diversity in different soils. The use of DGGE and TGGE techniques in monitoring microbial communities in metal-contaminated environments is reviewed in detail by Petrisor *et al.* (2006).

The assessment of microbial community structure changes by amplified ribosomal DNA restriction analysis was also successfully performed and may be the basis for developing innovative forensic techniques tracking the path of contaminants in environment. Gich *et al.* (2000) evaluated the suitability of this method for detecting and monitoring changes in activated sludge systems. The method was applied efficiently to detect differences in activated sludge bacterial communities fed on domestic or industrial wastewater, and subjected to different operational conditions. Rahman *et al.* (2003) used molecular and culture-based methods to characterize thermophilic bacteria associated with the subsurface soil environment in Northern Ireland. The bacteria were screened by amplified ribosomal DNA restriction analysis prior to 15S rRNA gene sequencing.

The Randomly Applied Polymorphic DNA (RAPD) method also holds potential for environmental forensics applications. This method was already applied in criminal forensic studies. Benecke (1998) used RAPD to permit quick identification of arthropods and support classical morphological and medico-legal analysis of maggots on a human corpse. It was determined that maggots found on the inside of a body bag were identical with maggots found on the outside of the bag, and with the pupae found on the floor under the corpse.

Ribotyping represents a molecular subtyping method with widespread applications in bacterial source tracking. Ribotyping was shown to be an effective tool for bacterial source tracking (BST) and is an excellent tool for indicating major sources of biological pollutants since most BST projects only need to determine broad groups of contamination. In the context of BST investigations, ribotyping is an effective tool for discriminating between human and nonhuman forms of *Escherichia coli*. It has also been shown effective in distinguishing *E. coli* from major animal groups. The statistical proba-

bilities for distinguishing the major animal groups are greatly enhanced when comparison samples are submitted.

A broad range of forensic applications of ribotyping are related to tracking the source of fecal pollution in water. Fecal pollution affects the quality and safety of water systems used for drinking, recreation, and in the harvesting of seafood. Fecal coliform can originate from numerous sources, including sewage treatment plant discharges, failing septic systems, agricultural and urban runoff, improper disposal of wastes from boats, and wildlife. Ribotyping is now commercially used to track *E. coli* that are regrowing and proliferating in watersheds (Scott *et al.*, 2004). The ability of ribotyping to differentiate between human and nonhuman sources of fecal pollution (generally *E. coli*) in water was also demonstrated in the study of Parveen *et al.* (1999), who applied ribotyping on a total of 238 *E. coli* isolates from human sources (HS) and nonhuman sources (NHS) collected from Apalachicola National Estuarine Research Reserve, from associated sewage plants, and directly from animals. The analysis of ribotyping profiles showed that 97% of the HS isolates and 100% of the animal fecal isolates were correctly classified.

Carson *et al.* (2001) also used ribotyping to successfully identify individual host sources of fecal *E. coli*, and distinguish between fecal *E. coli* of human and nonhuman origin, such as fecal *E. coli* ribotype patterns from humans and seven nonhuman hosts. Such an application could assist in formulation of pollution reduction plans. In another study also related to sources of fecal pollution of water, Scott *et al.* (2003) applied ribotyping analysis to *E. coli* isolates obtained from humans, beef cattle, dairy cattle, swine, and poultry from different locations in Florida. The study indicated that using a single restriction enzyme (*Hind*III) ribotyping cannot differentiate *E. coli* isolates from different animal species. A possible conclusion of the study was that a combination of geographic and environmental variation exert a significant role in affecting the ability of ribotyping for identification of sources of *E. coli* in the environment. The study indicated and confirmed, however, that ribotyping procedure can be used effectively to differentiate *E. coli* between human and nonhuman sources when applied to organisms isolated from a large geographic region.

A project conducted at the University of Georgia, Athens, (http://alpha.marsci.uga.edu/coastalcouncil/hartel_ribotyping.htm) used ribotyping to determine the host origin of fecal contamination in Georgia's coastal waters. As a result of the project, four hundred *Enterococcus faecalis* ribotypes (200 each from birds and humans) were also added to this host origin database.

Routine microbial source tracking analysis (ribotyping) is commercially available at relatively low costs. Such methods and applications include:

- DNA analysis of presence or absence of *E. faecium*
- DNA analysis and quantification of *E. faecium* from human sources

- DNA analysis of presence or absence *Bacteroides* from human, DNA analysis of presence or absence human fecal viruses (i.e., enteroviruses)
- DNA analysis of presence or absence of *E. coli* from cattle sources
- DNA analysis of presence or absence of *Bacteroides* from cattle sources
- DNA analysis of presence or absence of cattle fecal viruses (i.e., bovine enteroviruses)
- DNA ribotyping (i.e., DNA fingerprinting) analysis of five or more *E. coli* isolates per sample
- DNA comparative analysis of the *E. coli* isolates from fecal and water or sediment sources
- DNA fingerprinting of *Giardia* species and strains

15.5 SUMMARY

The various applications of DNA fingerprinting methods indubitably show the great potential of these techniques in developing reliable innovative forensic methods, adjustable to any contaminant and in basically any environmental conditions. We hope that this chapter will help understanding the many opportunities that microbial diversity and versatility in nature offer us for the development of reliable, sensitive, and comprehensive environmental forensic methods.

REFERENCES

Aeschbach-Hertig, W., Schlosser, P., Stute, M., Simpson, H. J., Ludin, A., and Clark, J. (1998) A $^{3}H/^{3}He$ study of ground water flow in a fractured bedrock aquifer. *Ground Water* 36(4), 661–670.

Aggarwal, R., Brar, D., Nandi, S., Huang, N., and Khush, G. (1999) Phylogenetic relationships among Oryza species revealed by AFLP markers. *Theoretical and Applied Genetics* 98, 1320–1328.

Andrews, J. and Siccama, T. (1995) Retranslocation of calcium and magnesium at the heartwood-sapwood boundary of Atlantic white cedar. *Ecology* 76(2), 569–663.

Ashby, W. and Fritts, H. (1972) Tree growth, air pollution, and climate near LaPorte, Ind. *Bulletin of the American Meteorological Society* 53(3), 246–251.

Baes, III, C. and McLaughlin, S. (1984) Trace elements in tree rings: Evidence of recent and historical air pollution. *Science* 224, 494–497.

Baillie, M. (1982) *Tree-Ring Dating and Archaeology*. The University of Chicago Press, Chicago, IL.

Baillie, M. (1995) *A Slice through Time: Dendrochronology and Precision Dating*. B. T. Batsford, Ltd., London, UK.

Baillie, M., Pilcher, J., Pollard, M., and Ramesh, R. (1995) Climatic significance of D/H and 13C/12C ratios in Irish oak cellulose. In *Tree Rings: From the Past to the Future*, S. Ohta, T. Fujii, N. Okada, M. Hughes, and D. Eckstein, eds. Proceedings of the International Workshop on Asian and Pacific Dendrochronology. Forestry and Forest Products Research Institute Scientific Meeting Report 1, 205–208.

Balouet J. (2005) Dendroecology, a key forensic age dating tool. *Environmental Forensics* 6(1), 3–4.

Balouet J. and Oudijk, G. (2006) The use of dendroecological methods to estimate the time frame of environmental releases. *Environmental Claims Journal* 18(10), 35–52.

Banks, J. (1991) A review of the use of tree-rings for the quantification of forest disturbances. *Dendrochronologia* 9, 51–70.

Barkay, T., Tripp, S., and Olson, B. (1985) Effect of metal-rich sewage sludge application on the bacterial communities of grass-lands. *Applied Environmental Microbiology* 49, 333–33••.

Barrett, M. (2004) Characteristics of urban groundwater. In *Urban Groundwater Pollution*, D. N. Lerner, ed. A. A. Balkema Publishers: Lisse, The Netherlands.

Beismann, H., Barker, J., Karp, A., and Speck, T. (1997) AFLP analysis sheds light on distribution of two Salix species and their hybrid along a natural gradient. *Molecular Ecology* 6, 989–993.

Benecke, M. (1998) Random Amplified Polymporphic DNA (RAPD) typing of necrophageous insects (diptera, coleoptera) in criminal forensic studies: Validation and use in practice. *Forensic Science International* 98, 157–168.

Boone, R., Tardif, J., and Westwood, R. (2004) Radial growth of oak and aspen near a coal-fired station, Manitoba, Canada. *Tree-Ring Research* 60(1), 45–58.

Bradshaw, J., Hackett, C., Meyer, R., Milbourne, D., McNichol, J., Philips, M., and Waugh, R. (1998) Identification of AFLP and SSR marker associated with quantitative resistance to Globodera pallida (Stone) in tetraploid potato (Solanum tuberosum subsp. tuberosum) with a view to marker-assisted selection. *Theoretical and Applied Genetics* 97, 202–210.

Brenninkmeijer, C. (1983) Deuterium, oxygen-18 and carbon-13 in tree rings and peat deposits in relation to climate. Ph.D. dissertation, University of Groningen, The Netherlands, 82.

Brodie, E., Edwards, S., and Clipson, N. (2003) Soil fungal community structure in a temperate upland grassland soil. *FEMS Mirobiology Ecology* 45(2), 105–114.

Brookes, P., McGrath, S., and Heiijnen, C. (1986) Metal residues in soils previously treated with sewage sludge and their effects on growth and nitrogen fixation by blue-green algae. *Soil Biology and Biochemistry* 18, 345–353.

Brown, R. (1961) Hydrology of tritium in the Ottawa Valley. *Geochimica and Cosmochimica* 21, 199–215.

Bruggemann, J., Stephen, J., Chang, Y., Macnaughton, S., Kowalchuk, G., Kline, E., and White, D. (2000) Competitive PCR-DGGE analysis of bacterial mixtures: An internal standard and an appraisal of template enumeration accuracy. *Journal of Microbiological Methods* 40(2), 111–123.

Brunk, C., Avaniss-Aghajani, E., and Brunk, C. (1996) A computer analysis of primer and probe hybridization potential with bacterial small subunit rRNA sequences. *Applied Environmental Microbiology* 62, 872–879.

Buhay, W. and Edwards, T. (1993) Reconstruction of Little Ice Age climate in southwestern Ontario, Canada, from oxygen and hydrogen isotope ratios in tree rings. In *Isotope Techniques in the Study of Past and Current Environmental Changes in the Hydrosphere and the Atmosphere*. International Atomic Energy Agency, Vienna, Austria, 407–417.

Burns, D., Plummer, L., McDonnell, J., Busenberg, E., Casile, G., Kendall, C. *et al.* (2003) The geochemical evolution of riparian ground water in a forested piedmont catchment. *Ground Water* 41(7), 913–925.

Busenberg, E. and Plummer, L. (1992) Use of chlorofluorocarbons (CCl_3F and CCl_2F_2) as hydrologic tracers and age-dating tools: The alluvium and terrace system of central Oklahoma. *Water Resources Research* 28(9), 2257–2283.

Busenberg, E. and Plummer, L. (2000) Dating young groundwater with sulfur hexafluoride: Natural and anthropogenic sources of sulfur hexafluoride. *Water Resources Research* 36(10), 3011–3030.

Cano, R. and Borucki, M. (1995) Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Science* 268, 1060–1064.

Cappellato, R., Peters, N., and Meyers, T. (1998) Above-ground sulfur cycling in adjacent coniferous and deciduous forests and watershed sulfur retention in the Georgia Piedmont, USA. *Water, Air and Soil Pollution* 103(1–4), 151–171.

Carson, C., Shear, B., Ellersieck, M., and Asfaw, A. (2001) Identification of fecal *Escherichia coli* from humans and animals by ribotyping. *Applied and Environmental Microbiology*, 1503–1507.

Chaudri, A., McGrath, S., Giller, K., and Sauerbeck, D. (1993) Enumeration of indigenous Rhizobium leguminosarum biovar trifolii in soils previously treated with metal-contaminated sewage sludge. *Soil Biology and Biochemistry* 25, 301–309.

Clark, I. and Fritz, P. (1997) *Environmental Isotopes in Hydrogeology*. Lewis Publishers, Boca Raton, Florida.

Clement, B., Kehl, L., DeBord, K., and Kitts, C. (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. *Journal of Microbiological Methods* 31, 135–142.

Cogbill, C. (1977) The effect of acid precipitation on tree growth in eastern North America. *Water, Air, and Soil Pollution* 8, 89–93.

Cook, E., Johnson, A., and Blasing, T. (1987) Forest decline: Modeling the effect of climate in tree rings. *Tree Physiology* 3, 27–40.

Cook, P. and Solomon, D. (1995) The transport of atmospheric trace gases to the water table: Implications for groundwater dating with chlorofluorocarbons and Krypton-85. *Water Resources Research* 31(2), 263–270.

Cook, P. and Böhlke, J. (2000) Determining timescales for groundwater flow and solute transport. In *Environmental tracers in subsurface hydrology*, P.G. Cook, and A. Herczeg, eds. Kluwer Academic Press, Boston, MA, 1–30.

Cook, P. and Herzog, A. (2000) *Environmental Tracers in Subsurface Hydrology*. Kluwer Academic Press, Boston, MA.

Dahlin, S., Witter, A., Martensson, A., Turner R., and Baath, E. (1997) Where's the limit? Changes in the microbiological properties of agricultural soils at low levels of metal contamination. *Soil Biology and Biochemistry* 29, 1405–1415.

De Baere, T., de Mendonça, R., Claeys, G., Verschraegen, G., Mijs, W., Verhelst, R. *et al.* (2002) Evaluation of amplified rDNA restriction analysis (ARDRA) for the identification of cultured mycobacteria in a diagnostic laboratory. *BMC Microbiology* 2, 4.

Derakshani, M., Lukow, T., and Liesack, W. (2001) Novel bacterial lineages at the (sub)division level as detected by signature nucleotide-targeted recovery of 15S rRNA genes from bulk soil and rice roots of flooded rice microcosms. *Applied Environmental Microbiology* 67, 623–631.

Dergam, J., Paiva, S., Schaeffer, C., Godinho, A., and Vieira, F. (2002) Phylogeography and RAPD-PCR variation in Hoplias malabaricus (Bloch, 1794) (Pisces, Teleostei) in southeastern Brazil. *Genetics and Molecular Biology* 25(4), 1415–4757.

Dickie, I., Xu, B., and Koide, R. (2002) Vertical niche differentiation of ectomycorrhizal hyphae in soil as shown by T-RFLP analysis. *New Phytologist* 156, 527–535.

Doelman, P. and Haanstra, L. (1979) Effects of lead on the soil bacterial microflora. *Soil Biology and Biochemistry* 11, 487–491.

Dunkle, S., Plummer, L., Busenberg, E., Phillips, P., Denver, J., Hamilton, P. *et al.* (1993) Chlorofluorocarbons (CCl_3F and CCl_2F_2) as dating tools and hydrologic tracers in shallow groundwater of the Delmarva Peninsula, Atlantic Coastal Plain, United States. *Water Resources Research* 29(12), 3837–3860.

Eckstein, D. and Krause, C. (1989) Dendroecological studies on spruce trees to monitor environmental changes around Hamburg. *IAWA Bulletin* 10(2), 175–182.

Edmands, J., Brabander, D., and Coleman, D. (2001) Uptake and mobility of uranium in black oaks: Implications for biomonitoring depleted uranium-contaminated groundwater. *Chemosphere* 44, 789–795.

Environmental Protection Agency (EPA). (2005) Tree core analysis brings savings to site assessments. *Technology News and Trends*, November 2005, 21, 3–5.

Eriksson, E. (1958) The possible use of tritium for estimating groundwater storage. *Tellus* 10(4), 472–478.

Faucherre, J. and Dutot, A. (1976) Le zinc et le cadmium des anneaux de croissance des végétaux en tant que traceurs chronologiques de pollutions d'un cours d'eau. Comptes Rendus des Séances de l'Académie des Sciences, Série D 282, 187–190.

Ferguson, G. and Woodbury, A. (2005) The effects of climate variability on estimates of recharge from temperature profiles. *Ground Water* 43(6), 837–842.

Ferretti, M., Udisti, R., and Barbolani, E. (1992) A new approach to experimental data evaluation from tree ring analysis: Ion-chromatographic determination of P, Cl, and S. *Fresenius' Journal of Analytical Chemistry* 343(7), 607–611.

Fetter, C. (1993) *Contaminant Hydrogeology*. MacMillan Publishing Company, New York.

Fey, A. and Conrad, R. (2000) Effect of temperature on carbon and electron flow and on the archaeal community in methanogenic rice field soil. *Applied Environmental Microbiology* 66(1), 4790–4797.

Flynn, S., Loffler, F., and Tiedje, J. (2000) Microbial community changes associated with a shift from reductive dechlorination of PCE to reductive dechlorination of cis-DCE and VC. *Environmental Science and Technology* 34, 1056–1061.

Fontana, J., Stout, A., Bolstorff, B., and Timperi, R. (2003) Automated Ribotyping and Pulsed-Field Gel Electrophoresis for Rapid Identification of Multidrug-Resistant Salmonella Serotype Newport. Emerging Infections Diseases 9(4); http://www.cdc.gov/ncidod/EID/vol9no4/02-0423.htm.

Fox, C., Kincaid, W., Nash, III, T., and Fritts, H. (1986) Tree ring variation in western larch (Larix occidentalis) exposed to sulfur dioxide emissions. *Canadian Journal of Forest Research* 15(2), 283–292.

Franklin, R., Garland, J., Bolster, C., and Mills, A. (2001) Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations and batch culture experiments. *Applied Environmental Microbiology* 67, 702–712.

Fritts, H. (1971) Dendroclimatology and dendroecology. *Quaternary Research* 1, 419–449.

Fritts, H. and Swetnam, T. (1989) Dendroecology: A tool for evaluating variations in past and present forest environments. *Advances in Ecological Research* 19, 111–188.

Fritz, S.J., Drimmie, R., and Fritz, P. (1991) Characterizing shallow aquifers using tritium and ^{14}C: Periodic sampling based on tritium half-life. *Applied Geochemistry* 6, 17–33.

Fulda, C. and Kinzelbach, W. (2000) Sulfur hexafluoride (SF_6) as a new age dating tool for shallow groundwater: Methods and first results. In *Tracers and Modelling in Hydrogeology* (Proceedings of the TraM'2000 Conference), IAHS Publication No. 262, 181–186.

Garbe-Schönberg, C., Reimann, C., and Pavlov, V. (1997) Laser ablation ICP-MS-analyses of tree-ring profiles in pine and birch from N-Norway and NW-Russia—A reliable record of the pollution history of the area. *Environmental Geology* 32(1), 9–15.

Ge, Z. and Taylor, D. (1998) Helicobacter pylori: Molecular genetics and diagnostic typing. *British Medical Bulletin* 54(1), 31–38.

Gich, F., Amer, E., Figueras, J., Abella, C., Balaguer, M., and Poch, M. (2000) Assessment of microbial community structure changes by amplified ribosomal DNA restriction analysis (ARDRA). *International Microbiology* 3, 103–106.

Gilfrich, J., Gilfrich, N., Skelton, E., Kirkland, J., Qadri, S., and Nagel, D. (1991) X-ray fluorescence analysis of tree rings. *X Ray Spectrometry* 20(4), 203–208.

Glass, R. (2002) Ground-water age and its water-management implications, Cook Inlet Basin, Alaska. U. S. Geological Survey Fact Sheet 022-02.

Glass, Sr., G., Hasenstein, K., and Chang, H. (1993) Determination of trace element concentration variations in tree rings using PIXE. *Nuclear Instruments and Methods in Physics Research* B 79, 393–396.

Goode, D., Busenberg, E., Plummer, L., Shapiro, A., and Vroblesky, D. (1999) CFCs in the unsaturated zone and in shallow ground water at Mirror Lake, New Hampshire. U. S. Geological Survey Water-Resources Investigations Report 99-4018C.

Greenblatt, C., Davis, A., Clement, B., Kitts, C., Cox, T., and Cano, R. (1998) Diversity of microorganisms isolated from amber. *Microbial Ecology* 38, 58–68.

Guay, B., Eastoe, C., Bassett, R., and Long, A. (2006) Identifying sources of groundwater in the lower Colorado River Valley, USA, with $\delta^{18}O$, δD and 3H: Implications for river water accounting. *Journal of Hydrogeology* 14, 146–158.

Guyette, R. and Cutter, B. (1994) Barium and manganese trends in tree-rings as monitors of sulfur deposition. *Water, Air and Soil Pollution* 73(1–4), 213–223.

Hagemeyer, J. (1993) Monitoring trace metal pollution with tree rings: A critical reassessment. In *Plants as Biomonitors. Indicators for Heavy Metals in the Terrestrial Environment*, B. Markert, ed. VCN Weinheim, New York, 541–563.

Hagemeyer, J., Schaefer, H., and Breckle, S.W. (1994) Seasonal variations of nickel concentrations in annual xylem rings of beech trees (Fagus sylvatica L.). *Science of the Total Environment* 145, 111–118.

Harju, L., Lill, J., Saarela, K., Heselius, S., Herberg, F., and Lindroos, A. (1996) Study of seasonal variations of trace-element concentrations within tree rings by thick-target PIXE analyses. *Nuclear Instruments and Methods in Physics Research* B 110, 536–541.

Heaton, T. and Vogel, J. (1981) "Excess air" in groundwater. *Journal of Hydrology* 50, 201–215.

Heuer, H. and Smalla, K. (1997) Application of denaturing gradient gel electrophoresis and temperature gradient gel electrophoresis for studying soil microbial communities. In *Modern Soil Microbiology*, J.D. Van Elsas *et al.*, eds. Marcel Dekker, Inc., New York, 353–373.

Heuer, H., Krsek, M., Baker, P., Smalla, K., and Wellington, E. (1997) Analysis of actinomycete communities by specific amplification of genes encoding 15S rRNA and gel-electrophoretic separation in denaturing gradients. *Applied Environmental Microbiology* 63(8), 3233–3241.

Heun, M., Schäfer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., and Salamini, F. (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278, 1312–1314.

Ho, D., Schlosser, P., Smethie, W., and Simpson, H. (1998) Variability in atmospheric chlorofluorocarbons (CCl_3F and CCl_2F_2) near a large urban area: Implications for groundwater dating. *Environmental Science and Technology* 32(15), 2377–2382.

Höhner, P., Werner, D., Balsiger, C., and Pasteris, G. (2003) Worldwide occurrence and fate of chlorofluorocarbons in groundwater. *Critical Reviews in Environmental Science and Technology* 33(1), 1–29.

Horswell, J., Cordiner, J., Mass, E., Martin, T., Sutherland, B., Speier, T. *et al.* (2002) Forensic comparison of soils by bacterial community DNA profiling. *Journal of Forensic Sciences* 47, 350–353.

Innes, J. (1992) Forest decline. *Progress in Physical Geography* 15(1), 1–64.

Jiang, G. (1996) Tree ring analysis for determination of pollution history of Chengde City, north China. *Journal of Environmental Sciences* 8(1), 77–85.

Johnston, C., Cook, P., Frape, S., Plummer, L., Busenberg, E., and Blackport, R. (1998) Ground water age and nitrate distribution within a glacial aquifer beneath a thick unsaturated zone. *Ground Water* 36(1), 171–180.

Jordan, M. and Lechevalier, M. (1975) Effects of zinc-smelter emissions on forest soil microflora. *Canadian Journal of Microbiology* 21, 1855–1865.

Kaplan, C. and Kitts, C. (2004) Bacterial succession in a petroleum land treatment unit. *Applied and Environmental Microbiology*, March 2004, 1777–1786.

Kaplan, C.W., Astaire, J.C., Sanders, M.E., Reddy, B.S., and Kitts, C.L. (2001) 15S Ribosomal DNA terminal restriction fragment pattern analysis of bacterial communities in feces of rats fed lactobacillus acidophilus NCFM. *Applied and Environmental Microbiology*, Apr. 2001, 1935–1939.

Kardolus, J.P., van Eck, H.J., and van den Berg, R.G. (1998) The potential of AFLPs in biosystematics: A first application in Solanum taxonomy (Solanaceae). *Plant Systematics and Evolution* 210, 87–103.

Karp, A., Kresovich, S., Bhat, K., Ayand, W., and Hodgkin, T. (1997) Molecular tools in plant genetic resources conservation: A guide to the technologies. IPGRI Technical Bulletin No. 2, International Plant Genetic Resources Institute, Rome, Italy. Available at http://198.93.227.125/publicat/techbull/TB2.pdf.

Katoh, T. (1998) Environmental assessment by plant indicators. *Japanese Journal of Toxicology and Environmental Health* 44(2), 55–74.

Kelly, J.J. and Tate, R.L. (1998) Effects of heavy metal contamination and remediation on soil microbial communities in the vicinity of a zinc smelter. *Journal of Environmental Quality* 27, 609–617.

Kitts, C.L. (2001) Terminal restriction fragment patterns: A tool for comparing microbial communities and assessing community dynamics. *Current Issues in Intestinal Microbiology* 2(1), 17–25.

Klump, S., Kipfer, R., Cirpka, O.A., Harvey, C.F., Brennwald, M.S., Ashfaque, K.N. *et al.* (2006) Groundwater dynamics and arsenic mobilization in Bangladesh assessed using noble gases and tritium. *Environmental Science and Technology* 40, 243–250.

Komor, S. (2002) Ground-water age dating in community wells in Oswego County, New York. U. S. Geological Survey Open-File Report 01-232.

Krane, D., Sternberg, D., and Burton, G. (1999) Randomly amplified polymorphic DNA profile-based measures of genetic diversity in crayfish correlated with environmental impacts. *Environmental Toxicology and Chemistry* 18, 504–508.

Krauss, S. and Peakall, R. (1998) An evaluation of the AFLP fingerprinting technique for the analysis of paternity in natural populations of Persoonia mollis (Proteaceae). *Australian Journal of Botany* 46, 533–546.

Krauss, S. (1999) Complete exclusion of nonsires in an analysis of paternity in a natural plant population using amplified fragment length polymorphism (AFLP). *Molecular Ecology* 8, 217–226.

Lambert, L., Cox, T., Mitchell, K., Rosello-Mora, R., Del Cueto, C., Dodge, D. *et al.* (1998) Staphylococcus succinus sp. nov., isolated from Dominican amber. *International Journal of Systematic Bacteriology* 48, 511–518.

LaMontagne, M., Schimel, J., and Holden, P. (2003) Comparison of subsurface and surface soil bacterial communities in California grassland as assessed by terminal restriction fragment length polymorphisms of PCR-amplified 15S rRNA genes. *Microbial Ecology* 46, 215–227.

Law, J., Donini, P., Koebner, R., Jones, C., and Cooke, R. (1998) DNA profiling and plant variety registration III: The statistical assessment of distinctness in wheat using amplified fragment length polymorphisms. *Euphytica* 102, 335–342.

LeBlanc, D. and Loehle, C. (1993) Effect of contaminated groundwater on tree growth: A tree-ring analysis. *Environmental Monitoring and Assessment* 24, 205–218.

Lerner, D. (1997) Too much or too little: Recharge in urban areas. In *Groundwater in the Urban Environment, Volume 1, Problems Processes and Management*, J. Chilton, ed. A. A. Balkema, Rotterdam, The Netherlands, 41–48.

Leuders, T. and Friedrich, M. (2000) Archaeal population dynamics during sequential reduction process in rice field soil. *Applied Environmental Microbiology* 66, 2732–2742.

Lindeberg, J. (2004) *X-Ray Based Tree-Ring Analysis*. Acta Universitatis Agriculturae Sueciae (ISBN 91-576-6533-8), 110.

Liu, W., Marsh, T., and Forney, L. (1998) Determination of the microbial diversity of anaerobic-aerobic activated sludge by a novel molecular biological technique. *Water Science and Technology* 37, 417–422.

Lorenz, S.E., McGrath, S.P., and Giller, K.E. (1992) Assessment of free-living nitrogen fixation activity as a biological indicator of heavy metal toxicity in soil. *Soil Biology and Biochemistry* 24, 601–606.

Mace, E., Gebhardt, C., and Lester, R. (1999) AFLP analysis of genetic relationships in the tribe Datureae (Solanaceae). *Theoretical and Applied Genetics* 99, 634–641.

Macur, R., Jackson, C., Botero, L., McDermott, T., and Inskeep, W. (2004) Bacterial populations associated with the oxidation and reduction of arsenic in an unsaturated soil. *Environmental Science and Technology* 38, 104–111.

Manners, M.M., Morvillo, N., Frederick, C., and Wagner, A. (2004) RAPD-PCR answers some long-standing questions about rose identification. *Acta Horticulturae (ISHS)* 634, 85–89; http://www.actahort.org/books/634/634_10.htm.

Marcantonio, F., Flowers, G., Thien, L., and Ellgaard, E. (1998) Lead isotopes in tree-rings: Chronology of pollution on bayou Trepagnier, Louisiana. *Environmental Science and Technology* 32, 2371–2376.

Marsh, T., Liu, W., Forney, L., and Cheng, H. (1998) Beginning a molecular analysis of the eukaryal community in activated sludge. *Water Science and Technology* 37, 455–460.

Massol-Deya, A., Odelson, D., Hickey, R., and Tiedje, J. (1995) Bacterial community fingerprinting of amplified 15S and 15-23S ribosomal DNA gene sequences and restriction endonuclease analysis (ARDRA). In *Molecular Microbial Ecology Manual*, A.D.L. Akkermans, J.D. van Elsas, and F.J. de Brujin, eds. Kluwer Academic Publications, Dordrecht, 1–8.

Matthes, M., Daly, A., and Edwards, K. (1998) Amplified fragment length polymorphism (AFLP). In *Molecular Tools for Screening Biodiversity*, A. Karp, P.G. Isaac, and D.S. Ingram, eds. Chapman and Hall, London, 183–190.

Mazor, E. (2004) *Chemical and Isotopic Groundwater Hydrology*, 3Ed. New York, New York, Marcel Dekker, Inc.

McClenahen, J., Vimmerstedt, J., and Scherzer, A. (1989) Elemental concentrations in tree rings by PIXE: Statistical variability, mobility, and effects of altered soil chemistry. *Canadian Journal of Forest Research* 19(7), 880–888.

McClenahen, J. and Vimmerstedt, J. (1993) Soil, climate, and atmospheric deposition relationships with elemental concentrations in annual rings of tulip tree. *Journal of Environmental Quality* 22(1), 23–32.

McLaughlin, S., Shortle, W., and Smith, K. (2002) Dendroecological applications in air pollution and environmental chemistry: Research needs. *Dendrochronologia* 20, 133–158.

Melcher, U. (2003) SSCPs. http://opbs.okstate.edu/~melcher/MG/MGW1/MG11129.html.

Morrison, R. and Murphy, B. (2002) Chlorinated solvents: Chemistry, history and utilization for source identification and age dating. Chapter 7. In *Introduction to Environmental Forensics*. Elsevier Publishing Company, Oxford, UK, 261–310.

Nash, III, T., Fritts, H., and Stokes, M. (1975) A technique for examining non-climatic variation in widths of annual tree rings with special reference to air pollution. *Tree-Ring Bulletin* 35, 15–24.

Nash, S. (1999) *Time, Trees and Prehistory.* University of Utah Press, Salt Lake City, UT.

Nelms, D., Harlow, G., Plummer, L., and Busenberg, E. (2003) Aquifer susceptibility in Virginia, 1998–2000. U. S. Geological Survey Water-Resources Investigations Report 03-4278.

Oliveira, A. and Ramos, M. (2002) PCR-based ribotyping of *Staphylococcus aureus*. *Brazilian Journal of Medical Biological Research* 35(2), 175–180.

Osborn, A., Moore E., and Timmis, K. (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environmental Microbiology* 2, 39–50.

Oster, H., Sonntag, C., and Münnich, K. (1996) Groundwater age dating with chlorofluorocarbons. *Water Resources Research* 32(10), 2989–3001.

Oudijk, G. and Schmitt, L. (2000) Age dating of a chlorinated solvent plume in groundwater. In *Tracers and Modelling in Hydrogeology* (Proceedings of the TraM'2000 Conference), IAHS Publication No. 262:255–262.

Oudijk, G. (2003) Estimating the minimum age of a chlorinated solvent plume in ground water with chlorofluorocarbon (CFC) and tritium methodologies: A case study. *Journal of Environmental Forensics* 4(1), 81–88.

Oudijk, G. (2005) The use of atmospheric contaminants to estimate the minimum age of environmental releases impacting ground water. *Journal of Environmental Forensics* 6(4), 345–354.

Park, H.-D., Regan, J.M., and Noguera, D.R. (2002) Molecular analysis of ammonia-oxidizing bacterial populations in aerated-anoxic Orbal processes. *Water Science and Technology* 46(1–2), 273–280.

Parveen, S., Portier, K., Robinson, K., Edmiston, L., and Tamplin, M. (1999) Discriminant analysis of ribotype profiles of *Escherichia coli* for differentiating human and nonhuman sources of fecal pollution. *Applied and Environmental Microbiology* 65(7), 3142–3147.

Petrisor, I., Parkinson, R., Horswell, J., Waters, J., Burgoyne, L., Catcheside, D. *et al.* (2006) Microbial forensics. Chapter 18. In *Environmental Forensics Contaminant Specific Guide*, R.D. Morrison and B.L. Murphy, eds. Elsevier Publishing Company, Oxford, UK, 227–258.

Philp, R. (2002) Application of stable isotopes and radioisotopes in environmental forensics. Chapter 5. In *Introduction to Environmental Forensics*, B. Murphy and R. Morrison, eds. Academic Press, San Diego, CA, 99–136.

Plummer, L., Michel, R., Thurman, E., and Glynn, P. (1993) Environmental tracers for age dating young ground water. In *Regional Ground-Water Quality*. Van Nostrand Reinhold Publishers, New York, 255–294.

Plummer, L. and Busenberg, E. (1998) Chlorofluorocarbons background. In *Environmental Tracers in Subsurface Hydrology*, P. Cook and A. Herczeg, eds. Kluwer Academic Press, Dordrect.

Punshon, T., Bertsch, P.M., Lanzirotti, A., McLeod, K., and Burger, J. (2003) Geochemical signature of contaminated sediment remobilization revealed by spatially resolved X-ray microanalysis of annual rings of Salix nigra. *Environmental Science and Technology* 37, 1766–1774.

Rafalski, J., Vogel, J., Morgante, M., Powell, W., Andre, C., and Tingey, S. (1996) Generating and using DNA markers in plants. In *Non-Mammalian Genomic Analysis: A Practical Guide*, B. Birren and E. Lai, eds. Academic Press, London, 75–134.

Ragsdale, H. and Berish, C. (1988) The decline of lead in tree rings of Carya spp. in urban Atlanta, GA, USA. *Biogeochemistry* 6, 21–29.

Rahman, T.J., Marchant, R., and Banat, I.M. (2003) Distribution and molecular investigation of highly thermophilic bacteria associated with cool soil environments. *Biochemical Society Transactions* (2004) 32, 209–213 (Printed in Great Britain).

Ridout, C. and Donini, P. (1999) Use of AFLP in cereals research. *Trends in Plant Science* 4, 76–79.

Rieseberg, L., Kim, M., and Seiler, G. (1999) Introgression between the cultivated sunflower and a sympatric wild relative, Helianthus petiolaris (Asteraceae). *International Journal of Plant Science* 150, 102–108.

Romalde, J., Castro, D., Magarinos, B., Lopez-Cortes, L., and Borrego, J. (2002) Comparison of ribotyping, randomly amplified polymorphic DNA, and pulsed-field gel electrophoresis for molecular typing of vibrio tapetis. *Systematic and Applied Microbiology* 25(4), 544–550.

Russell, J.R., Weber, J.C., Booth, A., Powell, W., Sotelo-Montes, C., and Dawson, I.K. (1999) Genetic variation of Calycophyllum spruceanum in the Peruvian Amazon Basin, revealed by amplified fragment length polymorphism (AFLP) analysis. *Molecular Ecology* 8, 199–204.

Sarkar, G., Yoon, H., and Sommer, S. (1992) Screening for mutations by RNA single-strand conformation polymorphism (rSSCP): Comparison with DNA-SSCP. *Nucleic Acids Research* 20(4), 4871–4878.

Schlosser, P., Stute, M., Dörr, H., Sonntag, C., and Münnich, K.O. (1988) Tritium/^{3}He dating of shallow groundwater. *Earth and Planetary Science Letters* 89, 353–362.

Schlosser, P., Stute, M., Sonntag, C., and Münnich, K.O. (1989) Tritiogenic ^{3}He in shallow groundwater. *Earth and Planetary Science Letters* 94, 245–256.

Schumacher, J., Struckhoff, G., and Burken, J. (2004) Assessment of subsurface chlorinated solvent contamination using tree cores at the Front street site and a former dry cleaning facility at the Riverfront Superfund site in New Haven, Missouri, 1999–2003, USGS SIR 2004-5049, 41.

Scott, T., Rose, J., Tracie, M., Farrah, S., and Lukasik, J. (2002) Microbial source tracking: Current methodology and future directions. *Applied and Environmental Microbiology*, December 2002, 5796–5803.

Scott, T., Parveen, S., Portier, K., Rose, J., Tamplin, M., Farrah, S. *et al.* (2003) Geographical variation in ribotype profiles of *Escherichia coli* isolates from humans, swine, poultry, beef, and dairy cattle in Florida. *Applied and Environmental Microbiology*, February 2003, 1089–1092.

Scott, T., Caren, J., Nelson, G., Jenkins, T., and Lukasik, J. (2004) Tracking sources of fecal pollution in a South Carolina watershed by ribotyping *Escherichia coli*: A case study. *Environmental Forensics* 5, 15–19.

Seifert, H., Dijkshoorn, L., Gerner-Smidt, P., Pelzer, N., Tjernberg, I., and Vaneechoutte, M. (1997) Distribution of Acinetobacter species on human skin: Comparison of phenotypic and genotypic identification methods. *Journal of Clinical Microbiology* 35, 2819–2825.

Shapiro, S., LeBlanc, D., Schlosser, P., and Ludin, A. (1990) Characterizing a sewage plume using the ^{3}H–^{3}He dating technique. *Ground Water* 37(6), 861–878.

Shianna, K., Rytter, R., and Spanier, J. (1998) Randomly amplified polymorphic DNA PCR analysis of bovine Cryptosporidium parvum strains isolated from the watershed of the Red River of the North. *Applied and Environmental Microbiology* 64(6), 2262–2265.

Smethie, W., Solomon, D., Schiff, S., and Matthieu, G. (1992) Tracing groundwater flow in the Borden aquifer using krypton-85. *Journal of Hydrology* 130, 279–297.

Smith, K. and Shortle, W. (1996) Tree biology and dendrochemistry. In *Tree Rings, Environment and Humanity*, J. Dean, D. Meko, and T. Swetnam, eds. Radiocarbon, Tucson, Arizona, 629–635.

Soliman, S., Ali, B., and Ahmed, M. (2003) Genetic comparisons of Egyptian date palm cultivars (Phoenix dactylifera L.) by RAPD-PCR. *African Journal of Biotechnology* 2(4), 86–87.

Solomon, D.K., Schiff, S.L., Poreda, R.J., and Clarke, W.B. (1993) A validation of the $^{3}H/^{3}He$ method for determining groundwater recharge. *Water Resources Research* 29(9), 2951–2962.

Stokes, M. and Smiley, T. (1996) *An Introduction to Tree-Ring Dating*. University of Arizona Press, Tucson, Arizona.

Stoner, J., Cowdery, T., and Puckett, L. (1997) Ground-water age dating and other tools used to assess land-use affects on water quality. U. S. Geological Survey Water-Resources Investigations Report 97–4150.

Stute, M., Deák, J., Révész, K., Böhlke, J., Deseö, É., Weppernig, R., and Schlosser, P. (1997) Tritium/^{3}H dating of river infiltration: An example from the Danube in the Szigetköz area, Hungary. *Ground Water* 35(5), 905–911.

Sunnucks, P., Wilson, A., Beheregaray, L., Zenger, K., French, J., and Taylor, A. (2000) SSCP Is not so difficult: The application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology* 9, 1599–1710.

Suomalainen, A., Ciafaloni, E., Koga, Y., Peltonen, L., DiMauro, S., and Schon, E. (1992) Use of single strand conformation polymorphism analysis to detect point mutations in human mitochondrial DNA. *Journal of Neurological Sciences* 111(2), 222–226.

Sutherland, E. and Martin, B. (1990) Growth response of Pseudotsuga menziesii to air pollution from copper smelting. *Canadian Journal of Forest Research* 20, 1020–1030.

Szabo, Z., Rice, D.E., Plummer, L.N., Busenberg, E., Drenkard, S., and Schlosser, P. (1996) Age dating of shallow groundwater with chlorofluorocarbons, tritium/

helium 3 and flow path analysis, southern New Jersey coastal plain. *Water Resources Research* 32(4), 1023–1038.

Takaoka, N. and Mitzutani, Y. (1987) Tritiogenic ^{3}He in groundwater in Takaoka. *Earth and Planetary Science Letters* 85, 74–78.

Thompson, G. and Hayes, J. (1979) Trichlorofluoromethane in groundwater—A possible tracer and indicator of groundwater age. *Water Resources Research* 15(3), 546–554.

Tolstikhin, I. and Kamenskiy, I. (1970) Determination of groundwater age by the T-^{3}He method. *Geochemistry International.* No. 7–8:810–811.

Truper, H. (1992) Prokaryotes—An overview with respect to biodiversity and environmental importance. *Biodiversity and Conservation* 1, 227–236.

Tweed, S.O., Weaver, T.R., and Cartwright, I. (2005) Distinguishing groundwater flow paths in different fractured-rock aquifers using groundwater chemistry: Dandenong Ranges, southeast Australia. *Hydrogeology Journal* 13, 771–786.

University of Miami/RSMAS. (2003) Tritium Laboratory: Environmental Level Measurements of Tritium, Helium and Chlorofluorocarbons (CFCs): http://www.rsmas.miami.edu/groups/tritium/, 3.

University of Miami/RSMAS. (2005) Advice on Sampling: http://www.rmas.miami.edu/groups/tritium/prices.html.

Vaderno, J. and Laserna, J. (2004) Laser induced plasma spectrometry, truly a surface analytical tool. *Spectrochemical Acta Part B* 59, 147–151.

Van der Maarel, M., Artz, R., Haanstra, R., and Forney, L. (1998) Association of marine archaea with the digestive tracts of two marine fish species. *Applied and Environmental Microbiology* 64, 2894–2898.

Ventura, M., Casas, I.A., Morelli, L., and Callegari, M.L. (2000) Rapid amplified ribosomal DNA restriction analysis (ARDRA) identification of Lactobacillus spp. isolated from fecal and vaginal samples. *Systematic and Applied Microbiology* 23(4), 504–509.

Víglasky, V., Antalík, M., Adamcík, J., and Podhradský, D. (2000) Early melting of supercoiled DNA topoisomers observed by TGGE. *Nucleic Acids Research* 28(11), E51.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M. *et al.* (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* 23, 4407–4414.

Vroblesky, D.A. and Yanosky, T.M. (1990) Use of tree-ring chemistry to document historical ground-water contamination events. *Ground Water* 28(5), 677–684.

Vroblesky, D.A., Nietch, C.T., and Morris, J.T. (1999) Chlorinated ethenes from groundwater in tree trunks. *Environmental Science and Technology* 33, 510–515.

Wallner, G. (1998) Elements in tree rings of Norway spruce (Picea abies (L.) Karst.) as indicators for SO2 polluted sites at the East-Erzgebirge (Germany). *Journal of Radioanalytical and Nuclear Chemistry* 238(1–2), 149–153.

Wagner, J. (2002) Screening Methods for Detection of Unknown Point Mutations. <http://www-users.med.cornell.edu/~jawagne/screening_for_mutations.html#Single-Strand.Conformational.Polymorphism>.

Wanninkhof, R., Sullivan, K., Damman, W., Proni, J., Bloetscher, F., Soloviev, A., and Carsey, T. (2005) Farfield tracing of a point source discharge plume in the coastal ocean using sulfur hexafluoride. *Environmental Science and Technology* 39(2), 8883–8890.

Watmough, S.A. and Hutchinson, T.C. (2002) Historical changes in lead concentrations in tree-rings of sycamore, oak and Scots pine in north-west England. *Science of the Total Environment* 293(1–3), 85–96.

Weissmann, G., Zhang, Y., LaBolle, E., and Fogg, G. (2002) Dispersion of groundwater age in an alluvial aquifer system. *Water Resources Research* 38(10), 1198–1211.

Xie, J., Wehner, T., and Conkling, M. (2002) PCR-based single-strand conformation polymorphism (SSCP) analysis to clone nine aquaporin genes in cucumber. *Journal of American Society of Horticulture Science* 127(6), 925–930.

Xu, M., Melchinger, A., Xia, X., and Lübberstedt, T. (1999) High-resolution mapping of loci conferring resistance to sugarcane mosaic virus in maize using RFLP, SSR and AFLP markers. *Molecular and General Genetics* 261, 574–581.

Yanosky, T. and Kappel, W. (1997) Effects of solution mining of salt on wetland hydrology as inferred from tree rings. *Water Resources Research* 33(3), 457–470.

Yanosky, T., Hansen, B., and Schening, M. (2001) Use of tree rings to investigate the onset of contamination of a shallow aquifer by chlorinated hydrocarbons. *Journal of Contaminant Hydrology* 50, 159–173.

Yanosky, T.M. and Carmichael, J.K. (1993) Element concentrations in growth rings of trees near an abandoned wood-preserving plant site at Jackson, Tennessee. U.S. Geological Survey Water Resources Investigations Report 93-4223, 1–69.

Zhu, J., Gale, M., Quarrie, S., Jackson, M., and Bryan, G. (1998) AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics* 96, 602–611.

INDEX

Accidental release, air dispersion
  dense gases 605
  release rate estimation 606–607
Accuracy *92*
Adamantane 361, *362*, 363, 364
Advection–dispersion model 539–541
Aerial photography
  geographic information system 66–67
  historical perspective 50–51
  map registration 64–66
  sources 50
  terrain visualization 67
  see also Photogrammetry; Photointerpretation
AERMOD 594–595
AFHRA, *see* Air Force Historical Research Agency
AFLP, *see* Amplified fragment length polymorphism
Air dispersion modeling
  accidental releases
    dense gases 605
    release rate estimation 606–607
  case study 599, *600*, *601*, 602
  computational fluid dynamic simulations 590
  concentration ratios 598–599
  deposition processes 580
  Gaussian plume equation
    downwind concentration 590–591
    maximum downwind concentration 591–592, *593*
    models
      AERMOD 594–595
      CALPUFF 595
      CTDMPLUS 596
      CTSCREEN 596
      ISC3 594
      Offshore and Coastal Dispersion Model 596
      SCREEN3 594
  overview 578–580
  plume rise 588–589
  redistribution after deposition
    drip line enhancement 602–603
    resuspension 604
    sediment enrichment 603–604
    soil permeability effects 602
  soil concentration estimating
    equation 597
    meteorology and terrain influences 582–583, *584*, 585–588
    source emission parameters 581–582
  soil contamination patterns 597, *598*
Air Force Historical Research Agency (AFHRA) 41–42
Air sampling, microscopy samples 612, *613*
Air–water partition coefficient, *see* Henry's law constant
ALAS, *see* Anthropogenic Lead Archeostratigraphy Model
Alliance of Automobile Manufacturers v. Carol Brunner 115
Allocation, *see* Cost allocation
ALS, *see* Alternating least squares
Alternating least squares (ALS)
  applications 253
  constraints 245–246
  convergence 246
  initial estimates 245
  least squares 245
  overview 244–245
Alumina solid phase adsorbent
  polychlorinated biphenyl sample cleanup 431–432
  semivolatile hydrocarbon sample cleanup 376–377
Ambient air, measurement 104–105
American Society for Photogrammetry and Remote Sensing (ASPRS) 50
American Society for Testing and Materials (ASTM) 7–8, 90, 100
Amplified fragment length polymorphism (AFLP) *702*, *703*, 713
Amplified ribosomal DNA restriction analysis (ARDRA) *704*, *709*
Analysis of variance (ANOVA), sample means comparisons 154–155, *156*, 157
ANOVA, *see* Analysis of variance
Anthropogenic Lead Archeostratigraphy Model (ALAS), tetraethyl lead isotope analysis *480*, 481–483
Apparent water saturation 554

Archives, resources for site history 25, 34–35, 40–42
ARDRA, *see* Amplified ribosomal DNA restriction analysis
Areal averages, estimation 177–179
Aroclors
  applications 314–315, 421, *428*
  calibration standards 429
  chlorine content 421, 426
  congener fingerprinting 426, *429*
  see also Polychlorinated biphenyls
Asbestos
  product types *617*
  vermiculite analysis 619–620
Asphaltenes, gas chromatography-isotope ratio mass spectrometry 493, *494*, 495
Association, *see* Epidemiology
ASTM, *see* American Society for Testing and Materials

Bacterial source tracking (BST) 714–716
Benz(a)anthracenes
  coal tar analysis by GC/MS/SIM *393*
  crude oil analysis by GC/MS/SIM *392*
Benzene, potential gas chromatography/photoionization detector interferences *109*
BFB, *see* p-Bromofluorobenzene
Bias 85, *92*, 143
Biodegradation, examples 521–522
Biomarkers
  coal tar signatures *408*
  crude oil signatures *407*
  GC/MS/SIM analysis 402, *403–405*, 406, 409
  stable isotope study combination *495*, *496*
BIOSLURP 552, 555
Blanks, *see* Decontamination check blank; Field blank; Trip blank
Block kriging, areal average estimation 178
Blunder 85, *92*
Bonferroni correction 204
Bootstrap resampling 203, *204*
p-Bromofluorobenzene (BFB), mass spectrometer calibration 342
BST, *see* Bacterial source tracking
BTEX chemicals
  biodegradation studies with isotopes 486–487
  linear regression analysis 166–168
tert-Butyl alcohol, MTBE degradation product 487–488

CAD, *see* Computer aided design
Calibration 114–115
CALPUFF 595
Carbon isotopes
  applications 456–457
  biodegradation studies 485–491
  bulk isotope values 463, *464*, *465*, *466*, *467*
  dissolved inorganic carbon 489
  gas chromatography-isotope ratio mass spectrometry of hydrocarbons 467–468, *469*, *470*, *471*, *472*, *473*, 474, 475, *476*, *477*
  hydrocarbon fingerprinting 415–416, *417*
  methane and biodegradation 489–490
  methyl tert-butyl ether analysis 468, 477, *478*, *479*
  origin of differences 461–462
  polychlorinated biphenyls 483–484
  reporting of measurements 462–463
  variations in organic matter 463, *464*
  volatile hydrocarbons 476–477
Carbon Preference Index (CPI) 288
CDF, *see* Cumulative distribution function
Central Limit Theorem 134, 142
Ceramic whiskers, microscopy 621–622
CERCLA, *see* Comprehensive Environmental Response, Compensation and Liability Act
CFCs, *see* Chlorofluorocarbons
Chain-of-custody, documentation 112
Chain-of-title research 33
Chemical fingerprinting
  EPA reference methods 317–319
  historical perspective 316–317
  method specialization 320–322
  overview 312–317
  polychlorinated biphenyls, *see* Polychlorinated biphenyls
  quality assurance 325–326
  quality control
    analytical procedures 331, *332–333*, 334–336
    sample preparation 326–328, *329–330*
  semivolatile hydrocarbons, *see* Semivolatile hydrocarbons
  technical challenges 319–320
  tiered approach 322, *323*, *324*
  volatile hydrocarbons, *see* Volatile hydrocarbons
Chemical mass balance (CMB) receptor models
  examples
    haze in Mount Zirkel Wilderness Area 294, *295*
    wintertime $PM_{2.5}$ sources in Denver 293, *294*
  indices and variables 275–276
  multiple sample solutions 278–280
  single sample least squares minimization solution 277–278
  tracer solution and enrichment factors 276–277
Chemical Process Safety of the American Institute of Chemical Engineers 605
Chi-square test 147, *148*, *149*
Chlorine isotopes
  gas chromatography-isotope ratio mass spectrometry evaluation 457, 485
  polychlorinated biphenyl analysis 483–484

Chlorofluorocarbons (CFCs), age-dating of contaminants in groundwater 691, *692*, 693–694
Clean Water Act (CWA) 18
Clinker particles *288*
CMB receptor models, *see* Chemical mass balance receptor models
Coal power plant, thermal desorption GC/MS of alkanes *291*
Coal tar
  high resolution hydrocarbon fingerprinting *382*
  polycyclic aromatic hydrocarbon
    GC/MS/SIM analysis *393*
    histogram *399*
  saturated hydrocarbon and biomarker signatures *408*
  triterpene fingerprints *412*
  volatile hydrocarbon analysis with gas chromatography/mass spectrometry *343*
Coefficient of determination 162
Coefficient of variation test 146–147
Colocated sample *92*
Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) 2–4, 18, 24, 36, 61, 313
Computational fluid dynamic simulations 590, 596
Computer aided design (CAD) 51, 56, 58
Cone penetrometer rig (CPT) 518
Confidence *92*–94
Confidence limits 138–140, *141*
Constitutional heterogeneity *92*
Contaminant chemistry
  Henry's law constant 514–515
  liquid density 515, *516*
  processes affecting
    biodegradation 521–522
    hydrolysis 520
    retardation 520–521
    sorption 520–521
  solubility 517–518
  vapor density *517*
  vapor pressure *517*
  viscosity 515, *516*
Contaminant transport, *see* Forensic transport models
Continuing calibration standard *332*, 334–335
Copper granules, sulfur removal 378, 433
Corporate succession research 30–32
Correction factor, Kruskal-Wallis test 158
Correlation coefficient, *see* Pearson's product moment correlation coefficient; Spearman Rank Correlation Coefficient
Cost allocation
  forensic issues 6
  risk-based methods 5–6
  stand-alone costs 5
  time considerations 6
CPI, *see* Carbon Preference Index
CPT, *see* Cone penetrometer rig
Cross-validation, principal components analysis 225
Crude oil
  gas chromatography-isotope ratio mass spectrometry 470, *471*, *472*, *473*
  high resolution hydrocarbon fingerprinting *380*
  polycyclic aromatic hydrocarbons
    GC/MS/SIM analysis *392*
    histogram *398*
  saturated hydrocarbon and biomarker signatures *407*, *410*
CTDMPLUS 596
CTSCREEN 596
Cumulative distribution function (CDF) 135
Cumulative percentage variance, principal components analysis 225
CWA, *see* Clean Water Act

Darcy velocity 548
Data quality objectives (DQOs) *90*
  analytical procedures *332–333*
  mass spectrometer tuning 331, *332–333*
  sample preparation *329–330*
DCE, *see* Dichloroethene
DCM, *see* Dichloromethane
DDE, *see* Dichlorodiphenyl dichloroethylene
DDT, *see* Dichlorodiphenyl trichlor ethane
Decafluorotriphenylphosphine (DFTPP), mass spectrometer tuning 331
Decision unit 91, 93
Decontamination check blank 98, 113
Degrees of freedom 146, 162
DEM, *see* Digital elevation model
Denaturing gradient gel electrophoresis (DGGE) *703*, *709*, 714
Dendroecology
  applications
    site assessment 683–684
    sulfur as biomarker for heating oil *684*, *685*
  calibration 682
  chemical analysis 680–681
  definition 672
  interpretation 682, *683*
  isotopic and organic compound analysis 681, *682*
  limitations 675
  overview 672, *673*, 674
  pattern of tree ring series 674
  prospects 685
  ring-width measurement 676, *677*, *678*, 679
  sample preparation 675, *676*
  sampling 675, *676*
DENEG algorithm, polytopic vector analysis 234, 241, *242*
Dense gas, accidental release 605
Dense nonaqueous phase lipids (DNAPLs) 101, 551, 694
Density, liquids 515, *516*

Deoxyribonucleic acid, *see* DNA fingerprinting
DFTPP, *see* Decafluorotriphenylphosphine
DGGE, *see* Denaturing gradient gel electrophoresis
Diagnostic ratio 109–110
Diamantane *362*, *364*, *365*
Diamond inclusions, laser ablation inductively coupled plasma mass spectrometry *660*, *661*
Diamondoids, direct injection GC/MS analysis 360, *361*, *362*, *363*, *364*, *365*
Dibenzothiophenes
  coal tar analysis by GC/MS/SIM *393*
  crude oil analysis by GC/MS/SIM *392*
DIC, *see* Dissolved inorganic carbon
Dichlorodiphenyl dichloroethylene (DDE) 95
Dichlorodiphenyl trichlor ethane (DDT) 85
Dichloroethene (DCE)
  biodegradation 487–488, 521
  biodegradation 521–522
  dendroecology 684
  transport modeling 518–519
Dichloromethane (DCM), semivolatile hydrocarbon extraction 372, 374–375, 377
Diesel
  high resolution hydrocarbon fingerprinting *381*, *387*
  polycyclic aromatic hydrocarbon histogram *398*
  thermal desorption GC/MS of alkanes *291*
  triterpene fingerprints *411*
  volatile hydrocarbon analysis with gas chromatography/mass spectrometry *343*
Differential diagnosis 16–17
Digital elevation model (DEM) 67
Digital image processing 54–56, 58
Digital raster graphics (DRGs) 64–65
Dioxin, air dispersion toxic tort case study 599, *600*, *601*, 602
Direct injection GC/MS, *see* Gas chromatography/mass spectrometry
Dissolved inorganic carbon (DIC), isotope studies of biodegradation 489
Distillation, *see* Simulated distillation
Distributional heterogeneity *92*
DNA fingerprinting
  DNA properties 686–697
  envoironmental forensics applications
    amplified fragment length polymorphism *702*, *703*, 713
    amplified ribosomal DNA restriction analysis *704*, *709*
    denaturing gradient gel electrophoresis *703*, *709*, 714
    general and specific applications *707–711*
    overview 695–696, 700
    potential applications *701–706*
    randomly amplified polymorphic DNA *705*, *710*, 714
    restriction fragment length polymorphism *705–706*, 711
    ribotyping *706*, *711*, 714–716
    single-stranded conformation polymorphism *702–703*, *708*, 713
    terminal restriction fragment length polymorphism 700, *701*, *707*, 712–713
    thermal gradient gel electrophoresis *703*, *709*, 714
  polymerase chain reaction 697–699
  principles 699–700
DNAPLs, *see* Dense nonaqueous phase lipids
Dose 12
DQOs, *see* Data quality objectives
DRGs, *see* Digital raster graphics
Drip line 602–603
Duplicates, *see* Replicates

EDA, *see* Exploratory data analysis
Edge enhancement, digital image processing 56
EDXRF, *see* Energy dispersive X-ray fluorescence
EF model, *see* Enrichment factor model
EF, *see* Elemental enrichment factor
Effective solubility 517–518
Effective water saturation 553–554
Eigenvector decomposition
  polytopic vector analysis 237–238
  principal components analysis 220–222
EIT, *see* Electrical impedance tomography
Electrical impedance tomography (EIT) 520
Electron microscopy, *see* Environmental forensic microscopy
Electrospray ionization-mass spectrometry (ESI-MS) 420
Elemental enrichment factor (EF) 603
Empirical orthogonal function (EOF) 280
En Core® sampler *347*
Energy dispersive X-ray fluorescence (EDXRF), dendroecology 680
Enrichment factor (EF) model 277
Environmental audit
  American Society for Testing and Materials standards 7–8
  purposes 7
Environmental forensic microscopy
  acid soluble and insoluble particles *617*
  asbestos
    product types *617*
    vermiculite analysis 619–620
  case studies
    automobiles with sooty deposition *626*, *627*, *628*
    carpet stain 625–626
    lead elevation in child *624*, *625*
    World Trade Center signature search 628, *629*, 630
  ceramic whiskers 621–622
  glass fibers 621

instrumentation 614
ISO Method 14966 620–621
nonfibrous particulate 622, *623*, 624
overview 611–612
particle analysis *615*, 616
polarized light microscopy 617
sampling
air 612, *613*
guidelines 616–617
surface dust 613–614
scanning electron microscopy 617–618
transmission electron microscopy 618
Environmental Protection Agency (EPA) 3, 17, 26, 59–60, 84, 100, 104
air dispersion modeling 579, 586–587
chemical fingerprinting reference methods 317–319
EOF, *see* Empirical orthogonal function
EPA, *see* Environmental Protection Agency
Epidemiology
association versus causation 13–14
odds ratio 13
statistical significance 15–16
Texas sharpshooter effect 14–15
Errors
sources in measurement process
blunders 85
bias 85
random error 85
planning 87–88
implementation 88–89
data handling 89–90
Type I 142
Type II 142
Errors in variable problem 190–191
Escherichia coli, bacterial source tracking 715
ESI-MS, *see* Electrospray ionization-mass spectrometry
ESMR, *see* Extended self-modeling curve resolution
Exploratory data analysis (EDA), rationale 208–210
Exposure 12
EXTENDED CABFAC algorithm, polytopic vector analysis 236
Extended self-modeling curve resolution (ESMR) *248*, *249*, 250
Exxon Valdez 18

FA, *see* Factor analysis
FAA, *see* Federal Aviation Administration
Factor analysis (FA) 280
FDA, *see* Food and Drug Administration
Federal Aviation Administration (FAA) 586
Federal Record Center (FRC) 40
Field blank 98, 113
Finger flow 534
Finished cement particles *288*
Fisher's *z* 165–166
Flame ionization detection, *see* Gas chromatography/flame ionization detector
Fluoranthenes
coal tar analysis by GC/MS/SIM *393*
crude oil analysis by GC/MS/SIM *392*
Fluorescing particulate matter (FPM) *189*, *190*
Food and Drug Administration (FDA) 85
Forensic transport models
air dispersion, *see* Air dispersion modeling
groundwater contaminant transport
conservative solute flow 537–538
inverse models 538–551, *552*, *553*, 554–558
model classification 536–537
nonaqueous phase liquids 518–520
paved surface contaminant transport
fluid transport overview 522–523
liquid 526–529
vapor 523–526
soil contaminant transport
assumptions 536
colloidal transport 534
finger flow 534
liquid 530–531
preferential flow 533
Seasonal Soil Compartment Model 531, *532*
vapor 529–530
VLEACH 532–533
Fourier transform microspectroscopy, *see* Environmental forensic microscopy
FPM, *see* Fluorescing particulate matter
FPM, *see* Fluorescing particulate matter
FRC, *see* Federal Record Center
Frobenius norm 250
F statistic 145, 149, 156
Fuel oil
high resolution hydrocarbon fingerprinting *380*
polycyclic aromatic hydrocarbon histogram *398*
triterpene fingerprints *410*
Full range whole oil analysis 369, *370*, 371
Furan, air dispersion toxic tort case study 599, *600*, *601*, 602
FUZZZY QMODEL algorithm, polytopic vector analysis 236, 239

Gas chromatography (GC)
analytical parameters 334–335
detectors 108
documentation 112–113
flame ionization detection, *see* Gas chromatography/flame ionization detector
mass spectrometry detection, *see* Gas chromatography-isotope ratio mass spectrometry; Gas chromatography/mass spectrometry
tips *109*
two-dimensional gas chromatography 419–420

Gas chromatography-isotope ratio mass spectrometry (GCIRMS)
asphaltenes 493, *494*, 495
biodegradation studies 485–491
chlorine isotope evaluation 457, 485
combination with other studies *491*, *492*, 493, *494*, *495*, *496*, 497
hydrocarbon fingerprinting 418–419
hydrocarbons 467–468, *469*, *470*, *471*, *472*, *473*, 474, 475, *476*, *477*
instrumentation *469*
methyl tert-butyl ether 468, 477, *478*, *479*
volatile hydrocarbons 476–477
see also Stable isotopes
Gas chromatography/electrochemical detection (GC/ECD), polychlorinated biphenyls 433, *434*, 435
Gas chromatography/flame ionization detector (GC/FID)
high resolution hydrocarbon fingerprinting 379, *380–382*, 383, *384*, *385*, *386*, *387*, 388
mass discrimination 385, *386*
Gas chromatography/high-resolution mass spectrometry, polychlorinated biphenyls *432*, 436–437, *438*, 439
Gas chromatography/mass spectrometry (GC/MS)
direct injection GC/MS
gasoline
alkyl lead additives 357, *358*, *359*, *360*
oxygenate analysis in nonaqueous phase lipids 362–365, *366*, *367*, *368*
polar fuel additive analysis 369
PIANO compounds
overview 353–355
sample analysis 355, *356*, *357*
sample preparation 355
standard preparation 355
polychlorinated biphenyls 435–436
purge-and-trap GC/MS of PIANO compounds
sample analysis 351, *352*, *353*, *354*
sample preparation 349–351
standard preparation 349
volatile hydrocarbon air and vapor analysis with gas chromatography/mass spectrometry 340, *341*, *342*, *343*, 344, *345*
see also Gas chromatography-isotope ratio mass spectrometry
Gasoline
direct injection GC/MS
alkyl lead additives 357, *358*, *359*, *360*
oxygenate analysis in nonaqueous phase lipids 362–365, *366*, *367*, *368*
polar fuel additive analysis 369
purge-and-trap GC/MS of PIANO compounds *354*
tetraethyl lead isotope analysis 479, *480*, 481–483
thermal desorption GC/MS of alkanes *291*
volatile hydrocarbon analysis with gas chromatography/mass spectrometry *343*, *346*
Gaussian plume equation
downwind concentration 590–591
maximum downwind concentration 591–592, *593*
models
AERMOD 594–595
CALPUFF 595
CTDMPLUS 596
CTSCREEN 596
ISC3 594
Offshore and Coastal Dispersion Model 596
SCREEN3 594
GC, *see* Gas chromatography
GC/ECD, *see* Gas chromatography/electrochemical detection
GC/FID, *see* Gas chromatography/flame ionization detector
GCIRMS, *see* Gas chromatography-isotope ratio mass spectrometry
GC/MS, *see* Gas chromatography/mass spectrometry
Gel permeation chromatography (GPC)
polychlorinated biphenyl sample cleanup 433–433
semivolatile hydrocarbon sample cleanup 377–378
Geographic information system (GIS) 51, 58, 66–67
GIS, *see* Geographic information system
Glass fibers, microscopy 621
Global positioning system (GPS) 51
GPC, *see* Gel permeation chromatography
GPS, *see* Global positioning system
Groundwater
age dating of contaminants
chlorofluorocarbon studies 691, *692*, 693–694
limitations 694
overview 686–687
tritium studies 687, *688*, 689–690, 694
age dating with radioisotopes 459–460
contaminant transport models
classification 536–537
conservative solute flow 537–538
inverse models 538–551, *552*, *553*, 554–558
measurement 102
Guillory v. Domtar Industries, Inc. 111

Half-time 164, 689
Harm, risk equivalence 4–5
Hazard index 17
Heartwood sapwood boundary (HSB) 683
Heating oil, sulfur as biomarker in dendroecology *684*, *685*
Helium, tritium degradation in groundwater 688–689

Henry's law constant 514–515
High-resolution mass spectrometry (HRMS), *see* Gas chromatography/high-resolution mass spectrometry
High temperature simulated distillation (HTSD) 414–415, *417*
Histogram equalization, digital image processing *57*
Holding time (HT) *330*, *333*
Hot patch
  high resolution hydrocarbon fingerprinting *382*
  polycyclic aromatic hydrocarbon histogram *399*
  triterpene fingerprints *412*
HRMS, *see* High-resolution mass spectrometry
HSB, *see* Heartwood sapwood boundary
HT, *see* Holding time
HTSD, *see* High temperature simulated distillation
Hydrolysis 520
Hypothesis testing, statistical analysis 140–142

IC, *see* Independent check
ICP, *see* Inductively-coupled plasma
IEC, *see* International Electrochemical Commission
Increment *92*
Independence, data 144
Independent check (IC) *333*, 334
Inductively coupled plasma (ICP), documentation for trace metal analysis 112, 114
Inductively coupled plasma mass spectrometry, *see* Laser ablation inductively coupled plasma mass spectrometry; Magnetic sector inductively coupled plasma mass spectrometry
Industrial behavior, photointerpretation 61–62
Initial calibration curve 331, 334
Insurance litigation
  equitable cost sharing 12
  expected and intended damage 10–11
  imminence of off-site migration 9
  sudden and accidental releases 11–12
  trigger of coverage 9–10
Internal standard (IS) *333*, 335
International Electrochemical Commission (IEC) 84
International Organization for Standardization (ISO) 84
Inverse modeling, *see* Forensic transport models
IRMS, *see* Isotope ratio mass spectrometry
Irreducible water saturation 553
IS, *see* Internal standard
ISC3 594
ISO, *see* International Organization for Standardization
Isopleth mapping
  air deposition *598*
  regional trend identification 173–174, *175*, *176*, *177*, *178*
Iso-tetramantane *362*
Isotope ratio mass spectrometry (IRMS), *see* Gas chromatography-isotope ratio mass spectrometry; Liquid chromatography/isotope ratio mass spectrometry

KD apparatus, *see* Kuderna Danish apparatus
Kerosene
  high resolution hydrocarbon fingerprinting *381*
  polycyclic aromatic hydrocarbon histogram *398*
  triterpene fingerprints *411*
Kinematic viscosity 515, *516*
Kruskal-Wallis test, sample means comparisons 157–158
Kuderna Danish (KD) apparatus, semivolatile hydrocarbon extraction 373, 375, 378

LA-ICP-MS, *see* Laser ablation inductively coupled plasma mass spectrometry
Laboratory control spike (LCS) 327, *329*
Laboratory subsample *93*
Laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS)
  applications
    base paint analysis 653, *654*, 655
    dendrochemical studies 651–652
    diamond formation studies *660*, *661*
    lead
      soil analysis 655, *656*, 657
      sourcing in environment 652–653, *654*
    metal exposure in tissue *657*, *658*, *659*, 660
  calibration for quantitative analysis 649, *650*
  historical perspective 638
  laser ablation operation *641*, 642, *643*
  overview 637–638, *639*
  prospects 662
  rastering analysis 646, *647*, *648*, 649
  sampling strategy 643, *644*, *645*, *646*
Laser induced plasma spectroscopy (LIPS), dendroecology 680
LC/IRMS, *see* Liquid chromatography/isotope ratio mass spectrometry
Lead
  environmental forensic microscopy case study *624*, *625*
  isotope analysis, *see* Tetraethyl lead
  laser ablation inductively coupled plasma mass spectrometry
    base paint analysis 653, *654*, 655
    soil analysis 655, *656*, 657
    sourcing in environment 652–653, *654*

Liability
  allocation at Superfund sites 2–6
  contexts and related forensics issues *2*
Light nonaqueous phase lipids (LNAPLs) 101, 551–552
Line equation 161
Linear regression analysis
  correlation matrix *166*, *167*
  multiple linear regression and correlation 170–171
  overview 161–164
  Pearson's product moment correlation coefficient 164–168
  scatter plots *167*, *168*
  Spearman Rank Correlation Coefficient 168–170
  table construction *162*
LIPS, *see* Laser induced plasma spectroscopy
Liquid chromatography/isotope ratio mass spectrometry (LC/IRMS) 497
Liquid transport
  paved surface 526–529
  soil 530–531
LNAPLs, *see* Light nonaqueous phase lipids
Log-log regression models 187–193
Log ratio difference (LRD)
  chemical profile evaluation 198–199, *200*, *201*
  computing probabilities for acceptable intervals 202–203
  multiple samples 204
  two-sample log ratio difference tests 203, *204*
Log-ratio-log (LRL) plot
  hypothesis log testing 196, *198*
  interesting patterns *196*, *197*
  principles 193, *194*, *195*, 196
LRL plot, *see* Log-ratio-log plot
Lubricating oil
  high resolution hydrocarbon fingerprinting *380*
  polycyclic aromatic hydrocarbon histogram *398*
  triterpene fingerprints *410*

Magnetic sector inductively coupled plasma mass spectrometry 639, *640*, 641
Malinowski indicator function, principal components analysis 225
Marine oil pollution
  causes
  regulations 19
MARPOL 73/78 19
Mass spectrometry (MS)
  electrospray ionization-mass spectrometry 420
  gas chromatography/mass spectrometry, *see* Gas chromatography/mass spectrometry
  high-resolution mass spectrometry, *see* Gas chromatography/high-resolution mass spectrometry
  inductively coupled plasma mass spectrometry, *see* Laser ablation inductively coupled plasma mass spectrometry; Magnetic sector inductively coupled plasma mass spectrometry
  isotope ratio mass spectrometry, *see* Gas chromatography-isotope ratio mass spectrometry; Liquid chromatography/isotope ratio mass spectrometry
  tuning 331
Matrix spike sample 327, *330*
Maximum downwind concentration 591–592, *593*
McCrone Particle Atlas 615
MDL, *see* Method Detection Limit
Mean, statistical analysis 133
Measurement process
  analysis 107–111
  assessment 111
  data value assessment 116, *117*, 118
  documentation 111–112
  error sources
    bias 85
    blunders 85
    confidence 93–94
    data handling 89–90
    implementation 88–89
    planning 87–88
    random error 85
    sample plan design
      concentration sampling 96–97
      identification sampling 95–96
      quality control samples 87, 97, *98*, 99
      sampling error consequences 94
  implementation of plan
    ambient air 104–105
    groundwater 102
    sediment 102
    soil
      definition 100–101
      nonaqueous phase lipids 101
      polycyclic aromatic hydrocarbons 102
      sampling 101
    soil gas 99–100
    surface contamination 104
    surface water 103–104
  planning
    considerations *86*
    Data quality objectives *90*
    decision unit 91, *93*
    questions 91
  project life cycle 85
  quality control interpretation 112–116
  reliability and relevance factors 84
  representativeness in sample design 143, 158–160
  sample handling 105–106
  sampling terms *92–93*
  statistical analysis 116
  subsampling 106–107
Meat cooking 287–288, *289*

Mercury, laser ablation inductively coupled plasma mass spectrometry in base paint analysis 653, *654*, 655
Meteorology
  data sources 586–587
  soil concentration estimating in air dispersion modeling 582–583, *584*, 585–587
Methane, isotope studies of biodegradation 489–490
Method Detection Limit (MDL) 115, *395*, 396
Methyl tert-butyl ether (MTBE)
  biodegradation studies 485–488, 522
  gas chromatography 107–109
  gas chromatography-isotope ratio mass spectrometry 468, 477, *478*, *479*
  physicochemical properties 364–365, *366*
  stable isotope studies 418–419, 468, 477–479
Microscopy, *see* Environmental forensic microscopy
Miesch coefficient of determination 227, *228*, *229*
MLR, *see* Multiple linear regression
MODFLOW 544
MOFAT 554
Mount Zirkel Wilderness Area 294, *295*
MS, *see* Mass spectrometry
MTBE, *see* Methyl tert-butyl ether
Multi-increment sample *93*
Multiple linear regression (MLR), chemical mass balance 279

NAAQS, *see* National Ambient Air Quality Standards
Naphthabenzothiophenes
  coal tar analysis by GC/MS/SIM *393*
  crude oil analysis by GC/MS/SIM *392*
Naphthalenes
  coal tar analysis by GC/MS/SIM *393*
  crude oil analysis by GC/MS/SIM *392*
NAPLs, *see* Nonaqueous phase lipids
NASA, *see* National Aeronautics and Space Administration
National Aeronautics and Space Administration (NASA) 50
National Ambient Air Quality Standards (NAAQS) 104
National Archives 25, 34
National Climatic Data Center (NCDC) 586
National Environmental Methods Index (NEMI) 107
National Weather Service (NWS) 586
Natural Resource Damage
  assessments 19
  claims 18
NCDC, *see* National Climatic Data Center
N-dimensional edge detection, extended self-modeling curve resolution 249–250
NEMI, *see* National Environmental Methods Index
Neural network, empirical receptor models 281
Neutral atmosphere 585
NEvap, *see* Nitrogen gas evaporation
Nitrogen gas evaporation (NEvap), semivolatile hydrocarbon extraction 373, 378
NOAAA National Status and Trends Program 373–374, 377
Nonaqueous phase lipids (NAPLs) 99–101, 328, 359, *360*
  direct GC/MS of oxygenates in nonaqueous phase lipids 362–365, *366*, *367*, *368*
  full range whole oil analysis 369, *370*, 371
  sample collection for volatile hydrocarbon analysis 348, 350
  semivolatile hydrocarbon extraction 375
  transport model selection 518–520
Nonfibrous particulate, microscopy 622, *623*, 624
Normalcy, statistical analysis 142–143
Normalized varimax loading, principal components analysis 225
Null hypothesis 140–142
NWS, *see* National Weather Service

OCD, *see* Offshore And Coastal Dispersion Model
Octanol-water partition coefficient, measurement error 85–86
Odds ratio (OR) 13
Offshore And Coastal Dispersion Model (OCD) 596
Oil pollution, *see* Marine oil pollution
Oil Pollution Act (OPA)
OPA, *see* Oil Pollution Act
Operations research 33–34
OR, *see* Odds ratio
Organization for Economic Co-operation and Development (OECD) 84
Orthophotos 65

PAHs, *see* Polycyclic aromatic hydrocarbons
Paint, base paint analysis with laser ablation inductively coupled plasma mass spectrometry 653, *654*, 655
Paired *t*-test, sample means comparisons 152–154
Particle mass (PM) 104
Particle size, particle mass quantities 282, *283*, 284, *285–286*
Paved surface, contaminant transport models
  fluid transport overview 522–523
  liquid 526–529
  vapor 523–526
PB, *see* Procedure blank
PBMS, *see* Performance-based measurement systems
PCA, *see* Principal components analysis
PCBs, *see* Polychlorinated biphenyls
PCRAMMET 587
PDF, *see* Probability distribution function

Pearson's product moment correlation coefficient 164–168
People v. Hale 105
People v. Sangani 106
Perflorotributylamine (PFTBA), mass spectrometer tuning 331
Performance-based measurement systems (PBMS) 320
PFTBA, *see* Perflorotributylamine
Phenanthrenes
  coal tar analysis by GC/MS/SIM *393*
  crude oil analysis by GC/MS/SIM *392*
Photogrammetry
  definition 50–51
  historical perspective 50–51
  principles
    digital image processing 54–56, 58
    photographic scale 51–52
    radial displacement 52
    resolution and media 53–54
    stereoscopic viewing 52–53
Photographic scale 51–52
Photointerpretation
  case studies
    chemical drainage from treated wood at creosote plant 72, *73*, *74*
    dark staining at maintenance yard 72–74, *75*, 76, *77*, *78*, *79*
    informal dumping area covered by plant addition 67, *68*, *69*
    transfer and storage spillage at circuit board fabrication plants 69–70, *71*, *72*
  historical perspective 58
  industrial behavior 61–62
  photoreading comparison 58–60
  process delineation 60
  staining 63–64
  transfer and storage spillage 62–63
  waste disposal 61
Photoionization detector (PID) 108
PIANO compounds
  chemical fingerprinting 337, *338–339*
  direct injection GC/MS analysis
    overview 353–355
    sample analysis 355, *356*, *357*
    sample preparation 355
    standard preparation 355
  purge-and-trap GC/MS analysis
    sample analysis 351, *352*, *353*, *354*
    sample preparation 349–351
    standard preparation 349
PID, *see* Photoionization detector
Plume rise 588–589
PM, *see* Particle mass
PMF, *see* Positive matrix factorization
Polarized light microscopy, *see* Environmental forensic microscopy
Polychlorinated biphenyls (PCBs)
  Aroclors 314–315
  congeners *422–426*
  gas chromatography fingerprinting
    calibration standards 429
    detection techniques 427–430
    extract cleanup
      alumina 431–432
      gel permeation chromatography 432–433
      sulfur removal with copper granules 433
      sulfuric acid cleanup 433
    gas chromatography/electrochemical detection 433, *434*, 435
    gas chromatography/high-resolution mass spectrometry *432*, 436–437, *438*, 439
    gas chromatography/mass spectrometry 435–436
    internal standards 433
    laboratory differences 428
    sample preparation 431
  homolog analytes *426*
  isotope analysis 483–484
  level of chlorination 421
  nomenclature 421, *427*
  octanol-water partition coefficients 85
  receptor modeling demonstration data sets 212, *213*, 214
  soil 101
  see also Aroclors
Polycyclic aromatic hydrocarbons (PAHs)
  analysis of variance in sediments 157
  gas chromatography-isotope ratio mass spectrometry 475, *476*, *477*
  GC/MS/SIM analysis
    coal tar signatures *393*
    crude oil signatures *392*, *407–408*
    histogram features 396–397, *398–400*
    method detection limit *395*, 396
    operating conditions 394
    polycyclic aromatic hydrocarbon ratio stability *401*, 402
    saturate and biomarkers 402, *403–405*, 406, 409, 413
    target analytes 389, *390–391*
  ratios 109–110
  soil sampling 102
Polymerase chain reaction (PCR) 697–699
Polytopic vector analysis (PVA)
  applications 253
  DENEG algorithm 234, 241, *242*
  eigenvector decomposition 237–238
  end-members
    composition and mixing proportion determination 238–241, *242*, 243
    number determination 237–238
  example 243, *244*
  EXTENDED CABFAC algorithm 236
  FUZZZY QMODEL algorithm 236, 239
  historical perspective 235–236
  scaling functions 236–237
Population 132
Population distributions 134, *135*, *136*, 137
Population parameter 132
Positive matrix factorization (PMF) 250–251, 279

Potentially responsible parties (PRPs) 3–6, 28, 36, 45
PQL, *see* Practical Quantification Limit
Practical Quantification Limit (PQL) 115
Precision representativeness *93*
Preferential flow 533
Principal components analysis (PCA)
  applications 116, 204, 253
  data transformations 217–219, *220*
  demonstration data sets 212, *213*, 214
  eigenvector decomposition 220–222
  number of significant principal components determination
    single index methods *223*, *224*, 225–226
    variable-by-variable goodness of fit 226–227, *228*, *229*, 230
  output *231*, *232*, 233
  overview 214, *215*, *216*, 217
  receptor modeling applications, *see* Receptor modeling
Probability distribution function (PDF) 134, 137
Procedure blank (PB) 326–328, *329*
Protein synthesis 697
PRPs, *see* Potentially responsible parties
Purge-and-trap GC/MS, PIANO compounds
  sample analysis 351, *352*, *353*, *354*
  sample preparation 349–351
  standard preparation 349
PVA, *see* Polytopic Vector Analysis

QAPP, *see* Quality assurance project plan
QMP, *see* Quality Management Plan
Quality assurance
  chemical fingerprinting 325–326
  error control 87
Quality assurance project plan (QAPP) 111
Quality control
  chemical fingerprinting
    analytical procedures 331, *332–333*, 334–336
    sample preparation 326–328, *329–330*
  data usage 116, *117*, 118
  interpretation 112, *113*, *114*, 115–116
  samples 87, 97, *98*, 99, *115*
Quality Management Plan (QMP) 325

Radial displacement 52
Radioisotopes
  age dating
    groundwater 459–460
    sediment 458–459
  applications 457
Radon 519–520
Raman spectroscopy, diamond formation studies *660*, *661*
Randomly amplified polymorphic DNA (RAPD) *705*, *710*, 714
RAPD, *see* Randomly amplified polymorphic DNA
Rastering, laser ablation inductively coupled plasma mass spectrometry 646, *647*, *648*, 649
Ratio data
  chemical indicator concentrations 186–190
  log-log regression models 187–193
  log-ratio-log plot
    hypothesis log testing 196, *198*
    interesting patterns *196*, *197*
    principles 193, *194*, *195*, 196
  value 185–186
RCRA, *see* Resource Conservation and Recovery Act
Receptor modeling
  alteration processes and influence on mixing model 251–252
  demonstration data sets 212, *213*, 214, *261–265*, *267–272*
  formal description of problem 210–211
  self-training methods, *see* Alternating least squares; Extended self-modeling curve resolution; Polytopic vector analysis; Positive matrix factorization; Target transformation factor analysis
  source apportionment of suspended particles
    chemical composition 284–287
    chemical mass balance receptor models 275–280
    empirical receptor models 280–282
    examples
      haze in Mt. Zirkel Wilderness Area 294, *295*
      product deposition near cement production factory 296, *297*
      wintertime $PM_{2.5}$ sources in Denver 293, *294*
    organic compounds 287–290, *291*, 292–293
    overview 273–275
    particle size 282, *283*, 284, *285–286*
    temporal and spatial variability 290, *292*, 293
Reference sample (RS) *333*, 334
Relative response factor (RRF) 342
Relative retention time (RRT) 342
Replicates, field samples 99, 114–116, 327, *329*
Representativeness, sample design 143, 158–160
Resolution, photographs 53–54
Resource Conservation and Recovery Act (RCRA) 313
Response factor (RF) 336
Restriction fragment length polymorphism (RFLP) *705–706*, 711
Resuspension 604
Retardation 520–521, 524
RETC 554
Reverse modeling, *see* Forensic transport models
RF, *see* Response factor
RFLP, *see* Restriction fragment length polymorphism
Ribbon NAPL Sample 518

Ribonucleic acid (RNA) 697
Ribotyping *706*, *711*, 714–716
Risk
  assessment in toxic tort 17
  harm equivalence 4–5
RNA, *see* Ribonucleic acid
Roadside dust, thermal desorption GC/MS of alkanes *291*
Roadway pavement
  high resolution hydrocarbon fingerprinting *382*
  polycyclic aromatic hydrocarbon histogram *399*
  triterpene fingerprints *412*
Root mean square deviation 133
RRF, *see* Relative response factor
RRT, *see* Relative retention time
RS, *see* Reference sample
Rubber sheeting, digital image processing 56

SAFER method, extended self-modeling curve resolution 248, *249*
Saltation 604
Sample handling 105–106
Sample mean
  comparisons
    Student's *t*-test 144–150
    Wilcoxon's Rank Sum test 150–152
    paired *t*-test 152–154
    analysis of variance 154–157
    Kruskal-Wallis test 157–158
    sampling design issues 158–160
  definition 133
Sample standard deviation 133–134
Sample statistics 132
Sanborn fire insurance maps 26
SARA, *see* Superfund Amendments and Reauthorization Act
Scanning electron microscopy, *see* Environmental forensic microscopy
SCRAM, *see* Support Center for Regulatory Air Modeling
SCREEN3 594
Screen test, principal components analysis 225
Seasonal Soil Compartment Model (SESOIL) 531, *532*
SEC, *see* Security and Exchange Commission
Security and Exchange Commission (SEC), corporate succession research 31
Sediment
  age dating with radioisotopes 458–459
  enrichment after air dispersion 603–604
  measurement 102
  semivolatile hydrocarbon extraction 373–374
Selected ion monitoring (SIM), *see* Gas chromatography/mass spectrometry
SEM, *see* Standard error of the mean
Semivolatile hydrocarbons
  definition 371
  gas chromatography/flame ionization detector and high resolution hydrocarbon fingerprinting 379, *380–382*, 383, *384*, *385*, *386*, *387*, 388
  polycyclic aromatic hydrocarbon analysis by GC/MS/SIM
    coal tar signatures *393*
    crude oil signatures *392*, *407–408*
    histogram features 396–397, *398–400*
    method detection limit *395*, 396
    operating conditions 394
    polycyclic aromatic hydrocarbon ratio stability *401*, 402
    saturate and biomarkers 402, *403–405*, 406, 409, 413
    target analytes 389, *390–391*
  sample preparation
    extraction
      nonaqueous phase liquid samples 375
      soil and sediment samples 373–374
      tissue samples 374–375
      water samples 372–373
      wipe samples 375–376
    extract cleanup and purification
      alumina solid phase adsorbent 376–377
      gel permeation chromatography 377–378
      silica gel solid phase adsorbent 378–379
SESOIL, *see* Seasonal Soil Compartment Model
Sesquiterpanes
  biomarker target analytes *403*
  coal tar signatures *408*
  crude oil signatures *407*
Signal-to-noise ratio, principal components analysis 225
Silica gel solid phase adsorbent, semivolatile hydrocarbon sample cleanup 378–379
Silver, laser ablation inductively coupled plasma mass spectrometry analysis of tissue *657*, *658*, *659*, 660
SIM, *see* Selected ion monitoring
Simulated distillation
  high temperature simulated distillation 414–415, *417*
  hydrocarbon fingerprinting 413, *414*, *416*
Single-stranded conformation polymorphism (SSCP) *702–703*, *708*, 713
Site history
  hypothetical cases and workplans
    examples *29*
    multiparty site with *in situ* releases
      federal archive research 40–42
      industry standards research 42–43
      past occupants and locations of operations 37–38
      state and local agency research 38–40
      target entity list 38
    multiparty site with transported and comingled waste

generator identification 44–45
landfill history 43–44
narrative completion 45
overview 28–30
single-party sites with history of potential problems
corporate succession research 30–32
facilities research 32
ownership and operations research 32–35
work product 35–36
Internet resources 26–27
narrative history 24–25
work plan development 25–28
Soil
air dispersion contamination patterns 597, *598*
contaminant concentration estimating in air dispersion modeling
equation 597
meteorology and terrain influences 582–583, *584*, 585–588
source emission parameters 581–582
contaminant transport models
assumptions 536
colloidal transport 534
finger flow 534
liquid 530–531
preferential flow 533
Seasonal Soil Compartment Model 531, *532*
vapor 529–530
VLEACH 532–533
definition 100–101
gas measurement 99–100
nonaqueous phase lipids 101
polycyclic aromatic hydrocarbons 102
sample collection for volatile hydrocarbon analysis *347*, 350–351
sampling 101
semivolatile hydrocarbon extraction 373–374
SOILPARA 554
Solubility 517–518
SOPs, *see* Standard operating procedures
Sorption 520–521
Spearman Rank Correlation Coefficient 168–170
Specific gravity 515, *516*
Splits, field samples 98–99, 113–114
SRM, *see* Standard reference material
SSCP, *see* Single-stranded conformation polymorphism
Stable atmosphere 584–585
Stable isotopes
applications 456–457
biodegradation studies 485–491
bulk isotope values 463, *464*, *465*, *466*, *467*
carbon isotopes
dissolved inorganic carbon 489
gas chromatography-isotope ratio mass spectrometry of hydrocarbons 467–468, *469*, *470*, *471*, *472*, *473*, 474, 475, *476*, *477*
methane 489–490
methyl tert-butyl ether analysis 468, 477, *478*, *479*
origin of differences 461–462
polychlorinated biphenyls 483–484
variations in organic matter 463, *464*
volatile hydrocarbons 476–477
combination with other studies *491*, *492*, 493, *494*, *495*, *496*, 497
compound-specific stable isotope measurement 418–419
dendroecology 681, *682*
hydrocarbon fingerprinting 415–416, *417*
reporting of measurements 462–463
tetraethyl lead isotope analysis 479, *480*, 481–483
Staining, photointerpretation 63–64, 72–78
Standard error of the mean (SEM) 134
Standard error of the slope 163
Standard operating procedures (SOPs) 87, 111
Standard reference material (SRM) *333*
Statistical significance, epidemiology 15–16
Steranes
biomarker target analytes *404–405*
coal tar signatures *408*
crude oil signatures *407*
Stereoscopic viewing, photographs 52–53, 55
Student's *t*-test
distribution 135, *136*
sample means comparisons 144–145, *146*, *147*, *148*, *149*, 150
Subsampling 106–107
Sulfur
biomarker for heating oil in dendroecology *684*, *685*
removal with copper granules 378, 433
Sulfuric acid, polychlorinated biphenyl sample cleanup 433
Sum of squares 156, 162
Superfund Amendments and Reauthorization Act (SARA) 24
Superfund site, liability allocation 2–6
Support Center for Regulatory Air Modeling (SCRAM) 586
Surface contamination, measurement 104
Surface dust sampling, microscopy samples 613–614
Surface water, measurement 103–104
Surrogate *330*

Target transformation factor analysis (TTFA) 246, *247*, 248
Tax research 33
TCE, *see* Trichloroethylene
TCFE, *see* Trichlorofluoroethene
TEF, *see* Toxicity equivalence factor
TEL, *see* Tetraethyl lead
TEM, *see* Total extractable material
Temperature inversion 585

TEQ, *see* Toxic equivalent quantity
Terminal restriction fragment length polymorphism (TRFLP) 700, *701*, *707*, 712–713
Terrain
  data sources 588
  soil concentration estimating effects in air dispersion modeling 587–588
    Tetraethyl lead (TEL)
  direct injection GC/MS analysis 357, *358*, *359*, *360*
  isotope analysis 479, *480*, 481–483
Texas sharpshooter effect 14–15
TGGE, *see* Thermal gradient gel electrophoresis
Thermal gradient gel electrophoresis (TGGE) *703*, *709*, 714
Thiessen polygon, areal average estimation 178, *179*
Time series, empirical receptor models 281–282
Tissue
  laser ablation inductively coupled plasma mass spectrometry of metal exposure *657*, *658*, *659*, 660
  semivolatile hydrocarbon extraction 374–375
TNT, isotope analysis for source tracing 482
TO-15 method 340, *341*, 342, *343*, 344, *345*
TOC, *see* Total organic carbon
Total extractable material (TEM) 383, 386–387
Total organic carbon (TOC) 100
Total petroleum hydrocarbons (TPH) 319, 379, 383, 386–387
Total suspended particulate (TSP) 152, *153*, 157
TOUGH2 556
Toxic equivalent quantity (TEQ) 599
Toxic tort
  differential diagnosis 16–17
  dose 12
  epidemiology
    association versus causation 13–14
    odds ratio 13
    statistical significance 15–16
    Texas sharpshooter effect 14–15
  exposure 12
  risk assessment 17
Toxicity equivalence factor (TEF), polychlorinated biphenyl congeners *422–426*
TPH, *see* Total petroleum hydrocarbons
Transmission electron microscopy, *see* Environmental forensic microscopy
Tree rings, *see* Dendroecology
Trend
  definition 171–172
  local deviation 172
  regional trend identification
    applications 172–173
    isopleth mapping 173–174, *175*, *176*, *177*, *178*
  surface model equations *173*
TRFLP, *see* Terminal restriction fragment length polymorphism
Triamantane *362*
Trichloroethylene (TCE)
  biodegradation 521–522
  dendroecology 684
  soil sampling 101
  transport modeling 518–519, 527–528
Trichlorofluoroethene (TCFE), biodegradation 522
Trip blank 97, 113
Triplicate, field sample *329*
Triterpanes
  biomarker target analytes *403–404*
  coal tar signatures *408*
  crude oil signatures *407*, *410*
  fuel signatures *410–411*
Tritium, age-dating of contaminants in groundwater 687, *688*, 689–690, 694
TSP, *see* Total suspended particulate
t-test, *see* Paired *t*-test; Student's *t*-test
TTFA, *see* Target transformation factor analysis
Two-dimensional gas chromatography 419–420
Type I error 142
Type II error 142

UCL, *see* Upper confidence limit
Ultraviolet fluorescing particulate matter (UVPM) *188*, *189*, *190*
United States Geological Survey (USGS) 50, 64, 85
United States v. Hess Environmental Laboratories 110
Unstable atmosphere 585
Upper confidence limit (UCL) 139–140, *141*
USGS, *see* United States Geological Survey
UVPM, *see* Ultraviolet fluorescing particulate matter

Vapor cloud 523
Vapor density *517*, 523
Vapor pressure *517*
Vapor transport
  paved surface 523–526
  soil 529–530
Variance 133
Variance ratio test 145
Vermiculite analysis, asbestos 619–620
Viscosity 515, *516*
VLEACH 532–533
Volatile hydrocarbons
  air and vapor analysis with gas chromatography/mass spectrometry 340, *341*, *342*, *343*, 344, *345*
  direct injection GC/MS
    gasoline alkyl lead additives 357, *358*, *359*, *360*

diamondoid analysis 360, *361*, *362*, *363*, *364*, *365*
full range whole oil analysis 369, *370*, 371
oxygenate analysis in nonaqueous phase lipids 362–365, *366*, *367*, *368*
PIANO compounds 337, *338–339*
  overview 353–355
  sample analysis 355, *356*, *357*
  sample preparation 355
  standard preparation 355
purge-and-trap GC/MS of PIANO compounds
  sample analysis 351, *352*, *353*, *354*
  sample preparation 349–351
  standard preparation 349
sample collection
  nonaqueous phase lipid samples 348
  soil samples *347*
  water samples 348
Waste disposal, photointerpretation 61
Water
  sample collection for volatile hydrocarbon analysis 348, 350
  semivolatile hydrocarbon extraction 372–373
  see also Groundwater; Surface water
Wilcoxon's Rank Sum test, sample means comparisons 150–152
Wipe samples, semivolatile hydrocarbon extraction 375–376
Witness interview, potentially responsible party identification 45
Workplan, *see* Site history
World Trade Center (WTC), signature search with environmental forensic microscopy 628, *629*, 630
WTC, *see* World Trade Center

z-score 134–135, 137–138