scientific data

OPEN

DATA DESCRIPTOR

Carbon Monitor Cities near-real-time daily estimates of CO_2 emissions from 1500 cities worldwide

Da Huo[1]✉, Xiaoting Huang[1], Xinyu Dou[1], Philippe Ciais[2], Yun Li[1], Zhu Deng[1], Yilong Wang[3], Duo Cui[1], Fouzi Benkhelifa[4], Taochun Sun[1], Biqing Zhu[1,2], Geoffrey Roest[5], Kevin R. Gurney[5], Piyu Ke[1], Rui Guo[1], Chenxi Lu[1], Xiaojuan Lin[1], Arminel Lovell[6], Kyra Appleby[6], Philip L. DeCola[7], Steven J. Davis[8] & Zhu Liu[1]✉

Building on near-real-time and spatially explicit estimates of daily carbon dioxide (CO_2) emissions, here we present and analyze a new city-level dataset of fossil fuel and cement emissions, Carbon Monitor Cities, which provides daily estimates of emissions from January 2019 through December 2021 for 1500 cities in 46 countries, and disaggregates five sectors: power generation, residential (buildings), industry, ground transportation, and aviation. The goal of this dataset is to improve the timeliness and temporal resolution of city-level emission inventories and includes estimates for both functional urban areas and city administrative areas that are consistent with global and regional totals. Comparisons with other datasets (i.e. CEADs, MEIC, Vulcan, and CDP-ICLEI Track) were performed, and we estimate the overall annual uncertainty range to be ±21.7%. Carbon Monitor Cities is a near-real-time, city-level emission dataset that includes cities around the world, including the first estimates for many cities in low-income countries.

Background & Summary

More than 60% of global fossil-fuel CO_2 emissions are produced in cities[1,2], and high-quality city-level emissions inventories are urgently needed to support international climate mitigation efforts[3–5]. For example, many cities have adopted goals of reaching net-zero emissions by 2030 or 2050, which require them to monitor and report emissions on a timely basis[6]. Unfortunately, a global, open, and harmonized dataset of city-level emission inventories is yet lacking[7,8]. Instead, most CO_2 emission inventories are conducted at the country level, as city-level fossil fuel consumption data are more difficult to acquire[9]. Furthermore, many inventories–including national inventories reported to the United Nations Framework Convention on Climate Change (UNFCCC) often lag reality by one years or more[10,11]. Thus, many city-level mitigation efforts are hampered by a lack of timely and high-quality emissions data with which to set benchmarks and monitor progress[12–15].

City-level CO_2 emissions may refer to either the CO_2 emissions produced within the territory of a city or emissions related to all the goods and services consumed in a city, which often include substantial emissions produced outside the city boundary[8,16,17]. The in-boundary emissions are typically referred to as scope 1, emissions from imported electricity as scope 2, and all other trans-boundary emissions associated with other city activities are referred to as scope 3[17,18]. Three conventional approaches have been used to attribute CO_2 emissions to cities: purely geographic production-based accounting, community infrastructure-based accounting

[1]Department of Earth System Science, Tsinghua University, Beijing, 100084, China. [2]Laboratoire des Sciences du Climate et de l'Environnement LSCE, Orme de Merisiers 91191, Gif-sur-Yvette, France. [3]Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101, China. [4]Nexqt, City Climate Intelligence, 9 rue des colonnes, Paris, 75002, France. [5]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, 86011, USA. [6]CDP Worldwide, London, UK. [7]Department of Atmospheric and Oceanic Sciences, University of Maryland, College Park, MD, 20742, USA. [8]Department of Earth System Science, University of California, Irvine, 3232 Croul Hall, Irvine, CA, 92697-3100, USA. ✉e-mail: dh2107@tsinghua.edu.cn; zhuliu@tsinghua.edu.cn

Content courtesy of Springer Nature, terms of use apply. Rights reserved

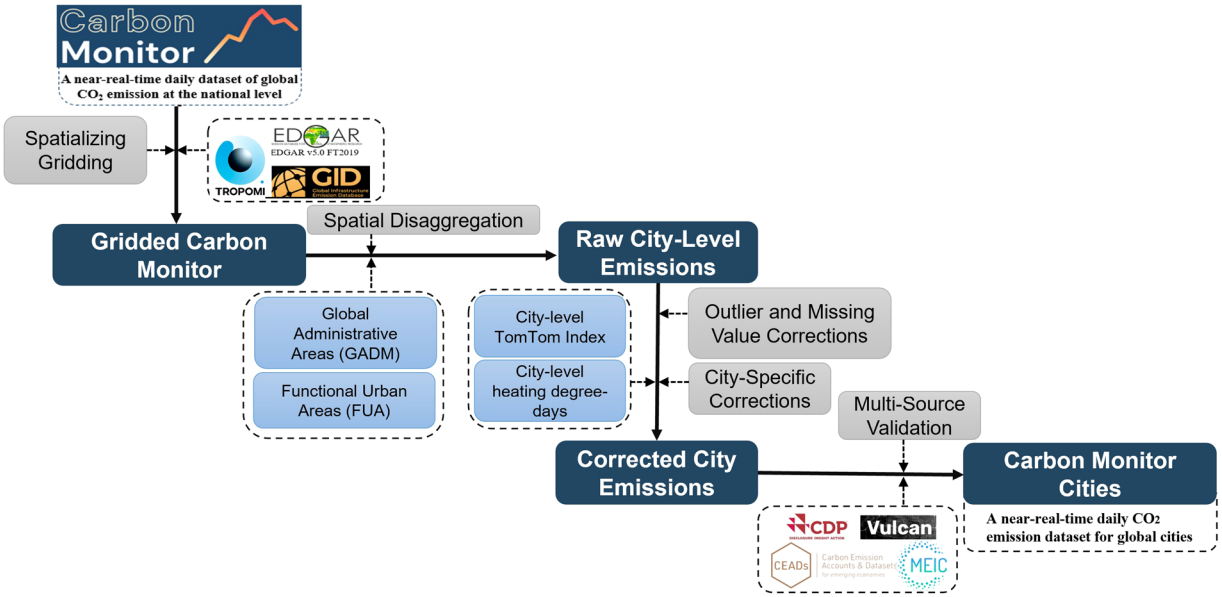

Fig. 1 Flowchart illustrates the main workflow and data used in each stage.

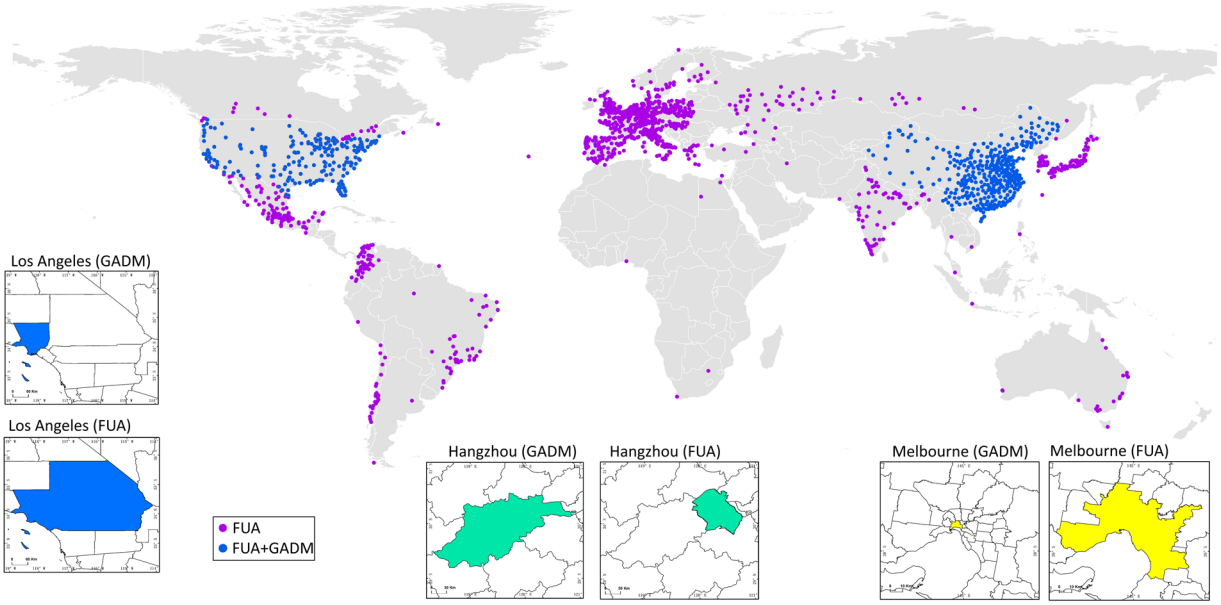

Fig. 2 Map showing all the cities covered in this dataset. Purple dots indicate cities with emissions estimated based on functional urban areas (FUA), and blue dots indicate cities with emissions estimated based on both FUA and administrative areas (GADM). Subplots depict examples of the comparison between the administrative city area versus the functional urban areas for cities in different regions.

(geographic-plus), and consumption-based accounting[16,17]. These approaches can provide good estimates for major cities that disclose high-quality energy consumption data[16]. However, they can not be readily applied to a larger scale, where city-specific data are largely absent, especially for smaller cities[19]. Downscaling represents a solution to the scalability issue. Some recent studies use economic input-output (IO) tables down-scaled from national statistics to attribute emissions to cities[18,20], and other studies use spatial proxies to disaggregate national or sub-national emissions to finer scales. Popular spatial proxies include night-time light imagery and existing gridded emission maps, such as the Emission Database for Global Atmospheric Research (EDGAR)[21–23], which also relies on other spatial data like population density and road networks. Downscaling has been used to construct multiple city-level datasets that cover a large number of cities[17,21], and we adopt a similar approach in this study.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Country/Region	Data source	Description
China	National Grid Daily Electric Load (https://www.sgcc.com.cn/)	Daily thermal production
India	Power System Operation Corporation Limited (https://posoco.in/reports/daily-reports/)	Daily thermal production (from coal, lignite, gas, naphtha and diesel)
United States	Energy Information Administration's (EIA) Hourly Electric Grid Monitor (https://www.eia.gov/beta/electricity/gridmonitor/)	Hourly thermal production (from coal, petroleum, and natural gas)
EU27	ENTSO-E Transparency platform (https://transparency.entsoe.eu/dashboard/show)	Hourly thermal production
United Kingdom	Balancing Mechanism Reporting Service (BMRS) (https://www.bmreports.com/)	Hourly power generation
Russia	United Power System of Russia (http://www.so-ups.ru/index.php)	Total hourly generation
Japan	Organization for Cross-regional Coordination of Transmission Operators (OCCTO) (https://www.occto.or.jp/en/)	Hourly thermal generation
Brazil	Operator of the National Electricity System (http://www.ons.org.br/Paginas/)	Hourly thermal production

Table 1. Data sources for the power sector.

Cities that disclose their emissions typically follow protocols or standards such as the Global Protocol for Community-Scale Greenhouse Gas Emission Inventories (GPC)[24], the International Council for Local Environmental Initiatives (ICLEI)[25] or the ICLEI-USA[16]. However, the reliability of these self-reported inventories is difficult to assess due to the lack of peer-review[5,8,26]. Inter-dataset comparison is also difficult due to inconsistent definitions of spatial and temporal scales, protocols, sector coverage, activity data sources, and accounting methods[5,8,25,27,28], and many original input data are untraceable[8]. Therefore, a methodological framework that supports inter-dataset comparisons and calibration is yet to be developed for cities.

Where city-level emissions inventories exist, they often rely on data provided by organizations such as the China Emission Accounts and Datasets (CEADs: https://www.ceads.net/[29]), the Multi-resolution Emission Inventory (MEIC, http://meicmodel.org/[30,31]), or inventory warehouses like the CDP-ICLEI Track (https://www.cdp.net/). But the coverage, timeliness, and temporal resolution of these data are not always sufficient to support agile and informed decision making. For example, although several high-quality datasets are available for high-income countries, such as the Covenant of Mayors database (https://www.globalcovenantofmayors.org/[25,32]) and OpenGHGmap (https://openghgmap.net/[19]) for EU cities, and the Vulcan and Hestia datasets for U.S. cities (https://vulcan.rc.nau.edu/, https://hestia.rc.nau.edu/[4,5,33]), the tension between human development and decarbonization requires an increasing focus on rapidly expanding cities in low-income and emerging regions in South America, South and Southeast Asia, Africa and the Middle East where high-quality emission inventories are lacking[34,35]. Moreover, most existing city-level inventories have the issues of long time lag and low temporal resolution. Recently, the methodological frameworks for estimating near-real-time (NRT) daily CO_2 emissions have been developed and successfully used for studying the impacts of COVID-19 on global CO_2 emissions[36–39]. Here, we build on these approaches to provide NRT daily emission estimates for hundreds of cities worldwide, including many in low-income regions.

The Carbon Monitor Cities (CM-Cities) dataset presented in this paper also provides a possible solution to address the inconsistency between administrative emissions versus community-wide emissions. Differences in spatial scope and accounting methods inevitably complicate comparisons among cities[8,17], and one possible solution is to compile the inventories based on different functional zones of the city, such as differentiating the core and commuting zones[25]. This dataset is constructed based on a global harmonized workflow that consistently quantifies production-based emissions from core administrative areas in top emitting countries and, separately, emissions from the world's major metropolitan/functional urban areas.

Methods

Workflow. Carbon Monitor Cities is downscaled from the Carbon Monitor, which is a NRT national level emission dataset at a global scale[10]. Specifically, CM-Cities is produced following a four-stage workflow (Fig. 1). The first stage mainly involves the construction of Global Gridded Daily CO_2 Emission Datasets (GRACED)[40], which are daily emission maps generated by spatializing Carbon Monitor daily emissions using the Global Carbon Grid (GID), the Emissions Database for Global Atmospheric Research (EDGAR) and TROPOspheric Monitoring Instrument (TROPOMI). GRACED covers seven sectors (power, industry, residential and commercial buildings, ground transportation, domestic aviation, international aviation, and international shipping) and provides NRT emission maps for fossil fuel combustion and cement production with a global spatial resolution of 0.1° by 0.1° and a temporal resolution of one day. GRACED is an intermediate gridded dataset between the Carbon Monitor and CM-Cities, and the methods for generating this gridded dataset is described in a later section.

In the second stage, we disaggregated the gridded daily emissions into cities based on two types of city areas: Global Administrative Areas (GADM) and Functional Urban Areas (FUA) to address the definition differences of "a city" in different countries. The FUA is defined by the Organisation for Economic Co-operation and Development (OECD) and the European Union as the high-density urban centres plus their surrounding commuting zones[41]. For OECD countries, we used the OECD FUA, which provides higher quality FUA for OECD countries (https://www.oecd.org/regional/regional-statistics/functional-urban-areas.htm[42]). For other countries, the Global Human Settlement FUA is used (https://ghsl.jrc.ec.europa.eu/ghs_fua.php[43]). The GADM level-2 administrative areas are used for prefecture-level cities in China and counties in the United States. The details of the features and the usage of FUA and GADM datasets are described in later sections. The spatial downscaling/disaggregation is performed by first converting the FUA and GADM shapefiles into raster datasets

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Country/Region	Sector	Data	Data Source
China	Steel industry	Crude steel production	World Steel Association (https://www.worldsteel.org/)
	Cement industry	Cement and clinker production	National Bureau of Statistics (http://www.stats.gov.cn/english/)
	Chemical industry	Sulfuric acid, caustic soda, soda ash, ethylene, chemical fertilizer, chemical pesticide, primary plastic and synthetic rubber	National Bureau of Statistics (http://www.stats.gov.cn/english/)
	Other industry	Crude iron ore, phosphate ore, salt, feed, refined edible vegetable oil, fresh and frozen meat, milk products, liquor, soft drinks, wine, beer, tobaccos, yarn, cloth, silk and woven fabric, machine-made paper and paperboards, plain glass, ten kinds of nonferrous metals, refined copper, lead, zinc, electrolyzed aluminum, industrial boilers, metal smelting equipment, and cement equipment	National Bureau of Statistics (http://www.stats.gov.cn/english/)
India	—	Industrial Production Index (IPI)	Ministry of Statistics and Programme Implementation (http://www.mospi.nic.in) Trading Economics (https://tradingeconomics.com)
United States	—	Industrial Production Index (IPI)	Federal Reserve Board (https://www.federalreserve.gov)
EU27 and UK	—	Industrial Production Index (IPI)	Eurostat (https://ec.europa.eu/eurostat/home) Trading Economics (https://tradingeconomics.com)
Russia	—	Industrial Production Index (IPI)	Federal State Statistics Service (https://eng.gks.ru)
Japan	—	Industrial Production Index (IPI)	Ministry of Economy, Trade and Industry (https://www.meti.go.jp)
Brazil	—	Industrial Production Index (IPI)	Brazilian Institute of Geography and Statistics (https://www.ibge.gov.br/en/institutional/the-ibge.htm)

Table 2. Data sources for industrial production.

with a unique ID assigned to each city. Then the raster city area maps are used as masks to extract the matching grid cells in the GRACED emission maps. We then aggregate emission values for grid cells that correspond to the same city mask to yield the total sectoral emission value for a given city.

In the third stage, we use city-level data to correct for the residential and ground transport sectors to address the bias in raw city-level inventories from the second stage. We use city-specific TomTom daily transport congestion data and daily heating degree days (HDD) for the corrections.

The fourth stage involves error correction and data validation. We first identify and remove outliers (which are mostly errors introduced by previous processing steps and/or from the source data) using statistical approaches. We then collected city-level inventories from other datasets (mostly annual data) and compared them to our results for validation. The detailed procedures for these processes are described in later sections.

CM-Cities currently includes city-level emission inventories from 01/01/2019 to 31/12/2021 for five main sectors: 1. power generation, 2. residential and commercial buildings, 3. industrial production, 4. ground transportation, and 5. aviation. These five sectors combined account for over 70% of fossil fuel CO_2 emissions from a city[33]. Custom code used in this work is described in the Code Availability section.

Coverage. CM-Cities covers 1500 cities in 46 counties (Fig. 2). Most of the cities are clustered in Europe, Asia, North and South America. Major cities in Oceania and Africa are also included. Figure 2 also shows comparisons between the FUA and the GADM for Los Angeles (US), Hangzhou (China), and Melbourne (Australia). The FUA typically covers a larger area than the administrative area, but for cities in some countries, such as China, the FUA is typically smaller than the administrative city area. The use of both area definitions facilitates dataset comparisons, which is highlighted for cities in China and the United States. These two different spatial scopes also provide critical information for differentiating administrative emissions versus community-wide emissions.

Near-real-time daily emissions by sector. CM-Cities is downscaled from the GRACED dataset, in which spatial distribution and daily variations of emissions are combined. This section describes the methods for estimating NRT daily emissions from a temporal perspective, and the next section describes the spatial gridding procedure. The estimation of daily emission variations follows the Carbon Monitor national dataset[10,36,44], which provides daily fossil fuel CO_2 emissions since January 1st, 2019 on the global and national levels, with detailed estimates in 7 main sectors, i.e., power, industry, ground transport, residential (including commercial), domestic aviation, international aviation, and international shipping. Emissions from international bunkers (including the international aviation sector and international shipping sector) are only accounted for at the global level and usually excluded from the national territorial emissions according to the IPCC guidelines. Therefore, CM-Cites considers the other 5 sectors, i.e., power, industry, ground transport, residential, and domestic aviation.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

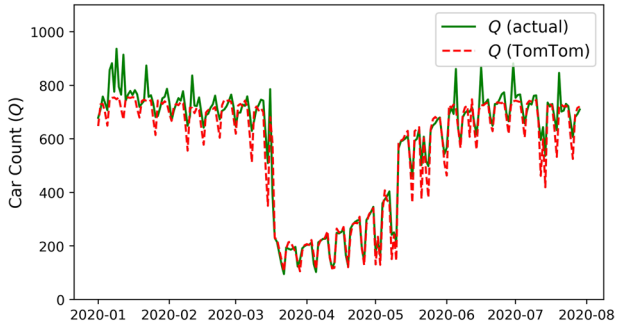

Fig. 3 Comparison between the actual and TomTom estimated hourly car counts on the measured roads in Paris. TomTom-based estimates accurately depicted the drop in traffic during the lock down period in 2020.

Power sector. Daily power generation data are acquired from multiple open data sources depending on the country (Table 1), which provides live power generation data with a daily or hourly resolution, and accounts for more than 70% of the total CO_2 emissions in the power sector[10]. The emission factors are estimated using EDGAR's electricity emissions, divided by our collection of coal-fired electricity data in various countries. The daily emissions are estimated as:

$$Emis_{power,daily} = Emis_{power,yearly}(AD_{power,daily}/AD_{power,yearly}), \tag{1}$$

where AD is the power generation. For emissions from other countries (countries not listed in Table 1), we assumed a linear relationship between daily global emission and daily total emissions from these countries, and then adjusted the emissions for countries that adopted lock-down measures during the COVID-19 following the method used by the Carbon Monitor national dataset[10].

Industry sector. For the industry sector, the daily emissions are calculated from the monthly industrial production index and the daily power generation data. Monthly industrial production data are acquired from several datasets (Table 2). The monthly CO_2 emissions estimated from the Industrial Production Index (IPI) are then disaggregated into a daily scale using daily power generation data. This approach is based on two assumptions: 1. A linear relationship exists between daily industrial production and industrial fossil fuel use. 2. A linear relationship exists between daily industry activity and daily electricity production[10]. The monthly and daily industry emissions are estimated following:

$$Emis_{ind,monthly,currentyear,c} = Emis_{ind,yearly,2019,c}(IPI_{monthly,currentyear,c}/IPI_{yearly,2019,c}), \tag{2}$$

$$Emis_{ind,daily} = Emis_{ind,monthly}(Elec_{daily}/Elec_{monthly}), \tag{3}$$

where $Emis_{ind,\ monthly,\ currentyear,\ c}$ is the monthly industry emissions for country c in current year, $Emis_{ind,\ yearly,\ 2019,\ c}$ is the yearly industry emissions for country c in 2019 (year of the latest update of baseline emissions), IPI is the corresponding Industrial Production Index. $Emis_{ind,\ daily}$ and $Emis_{ind,\ monthly}$ are the daily and monthly industry emissions, respectively. $Elec_{daily}$ and $Elec_{monthly}$ are the daily and monthly electricity production, respectively. For countries not listed in Table 2, the industry sector emissions are estimated in the same way as for the power sector.

Ground transport sector. Daily emissions from ground transportation are estimated using TomTom live congestion index and EDGAR road transportation emissions. The TomTom traffic congestion level represents the extra time spent on a trip in congested conditions, as a percentage, compared to uncongested conditions. TomTom congestion level data were obtained for more than 400 cities around the world at a temporal resolution of one hour (https://www.tomtom.com/traffic-index/). This approach permits the estimation of NRT emissions from ground transportation with a temporal resolution up to one hour, and TomTom grants users permission for non-commercial usage. The TomTom live congestion level data was proven to be highly accurate for most cities, and Carbon Monitor has successfully adopted this approach[10]. Note that a zero-congestion level means the traffic is fluid or "normal" but does not mean there are no vehicles and zero emissions. The lower threshold of emissions when the congestion level is zero was estimated using real-time data from an average of 60 roads in the city of Paris. TomTom data accurately depicts the traffic volume using a sigmoid function-based regression (Eq. 4), and Fig. 3 is a comparison between the actual and TomTom estimated hourly car counts on the measured roads in Paris. The estimated traffic volume is then used to allocate the EDGAR on-road emissions to each day (Eq. 5).

$$Q_d = a + \frac{\beta X^{\gamma}}{\lambda^{\gamma} + X^{\gamma}}, \tag{4}$$

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Data	Data Description	Resolution
Global Carbon Grid (GID) v1.0	Global Grid in the Global Infrastructure Emission Database. Including power, industry, residential, transport, shipping, and aviation sectors with high data quality in spatial fine-grained maps (http://gidmodel.org)	0.1°×0.1° global, Annual
Emission Database for Global Atmospheric Research (EDGARv5.0)	EDGAR v5.0 FT2019, covers major fossil CO_2 sources globally, with monthly emissions provided per main source category (https://edgar.jrc.ec.europa.eu)	0.1°×0.1° global, Monthly
TROPOMI NO_2 Retrievals	NO_2 thermal chemical vapor deposition retrievals acquired by the TROPOspheric Monitoring Instrument (TROPOMI) on board the Sentinel-5 Precursor satellite, launched in 2017	0.1°×0.1° global, Daily

Table 3. List of gridded data used for producing Global Gridded Daily CO_2 Emissions Dataset (GRACED).

$$Emis_{trans,d} = Q_d \frac{Emis_{onroad}}{\sum_{d=1}^{n} Q_d}, \tag{5}$$

where Q_d is the mean vehicle number per hour in day d, X is the daily mean TomTom congestion level data, and a, β, γ, λ are regression parameters, $Emis_{trans,\,c,\,d}$ is the ground transport emissions in day d, $Emis_{onroad}$ is the annual EDGAR road transportation emissions, n is the number of days in a year. For cities not covered by TomTom, we assumed that the emission changes follow the mean changes of other cities in the country. If no city in the country has TomTom data, then the relative emission changes are assumed to follow the same pattern of the total emissions from all TomTom-covered countries.

Residential sector. Carbon Monitor uses the fluctuation of air temperature to capture the daily variations in the energy consumption of residential and commercial buildings. The assumption associated with this method is that the heating demand, which is the largest contribution to the daily variability in emissions for this sector, is strongly governed by air temperature[45], which determines the HDD (cooling in summer mainly consumes electricity that is covered in the power sector). This approach uses population-weighted HDD for different geographic locations for each day based on the ERA-5 reanalysis of air temperature[46] and also accounts for temperature-independent cooking emissions following EDGAR. The EDGAR residential emissions are then downscaled to daily values based on daily variations in population-weighted heating degree days.

$$Emis_{res,c,d} = Emis_{res,c,m} R_{heating,c,m} \frac{HDD_{c,d}}{\sum_m HDD_{c,d}} + Emis_{res,c,m} \frac{1 - R_{heating,c,m}}{N_m}, \tag{6}$$

$$HDD_{c,d} = \sum_g R_{pop,g} He\,(18 - T_{g,d}), \tag{7}$$

where $Emis_{res,\,c,\,d}$ and $Emis_{res,\,c,\,m}$ are the residential emissions for country c in day d and month m respectively, $R_{heating,\,c,\,m}$ is the percentage of residential emissions from heating demand in country c in month m, $HDD_{c,d}$ is the population-weighted heating degree day for country c in day d, N_m is the number of days in month m, $R_{pop,\,g}$ is the ratio of the population in grid g to the total national population, which is acquired from the Gridded Population of the World, version 4[47], He is a Heaviside step function that converts any negative values to zero, $T_{g,\,d}$ is the average air temperature in Celsius for grid g in day d at 2 meters derived from ERA5[46], and 18 is a HDD reference temperature of 18 °C.

Aviation sector. Emissions in the aviation sector are computed from individual commercial flights data from the Flightradar24 database (https://www.flightradar24.com). This sector covers domestic flights, and all airports around a city were selected even if they are not part of the FUA, but some airports are not covered by the GADM, since we follow a territorial approach for emission allocation, if a city does not have an airport but emissions are present within the FUA boundary, (e.g., the city of Dongguan does not have its own airport but has two nearby airports in Guangzhou and Shenzhen). In this case, we have attributed the daily patterns of the airport that is closest to the city. The daily CO_2 emissions were estimated as the product of distance flown and a constant emission factor (EF_{avi}).

$$Emis_{avi} = DF \cdot EF_{avi}, \tag{8}$$

where DF is distance flown, which is computed using great circle distance between the take-off, cruising, descent, and landing points for each flight and are cumulated over all flights. The emission factor per kilometer flown is assumed to be a constant for the mix of all aircraft from an airport (including regional, narrowbody passenger, widebody passenger, and freight operations) as the share of flight types has not significantly changed since 2019.

Gridded daily CO_2 emissions. Carbon Monitor Cities disaggregates the Carbon Monitor national emissions to cities using the GRACED dataset developed by the Carbon Monitor team[40], which consists of emission maps generated by spatializing and gridding the daily national emission inventories from Carbon Monitor into grid cells. This was achieved by estimating spatial distribution proxies from satellite data and existing gridded products while maintaining consistency between bottom-up accounting results and the spatial sum of the

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Column	Description
City	Name of the city
Country	Country where the city is located. The following countries are covered in this dataset: Argentina, Australia, Austria, Bangladesh, Belgium, Brazil, Canada, Chile,China, Colombia, Denmark, Egypt, Finland, France, Germany, Greece, Hungary, India, Indonesia, Iran, Italy, Japan, Korea, Malaysia, Mexico, Myanmar, Netherlands, Nigeria, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Russia, Singapore, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, United Arab Emirates, United Kingdom, United States, Vietnam
Date	Date (YYYY-MM-DD) on which the emissions were estimated. Currently, the dataset provides emissions from 2019-01-01 to 2021-12-31
Sector	Sector for which the emissions were estimated, including power, industry, residential, ground transport, aviation
Value	Magnitude of daily emissions with a unit of $ktCO_2$
Timestamp	Unix timestamp at 00:00:00 (GMT + 0000) on each day for scientific visualization

Table 4. Data attributes.

gridded results. Three datasets were used in producing GRACED: 1. The Global Carbon Grid (GID), which provides global CO_2 emissions data from major industry and power plant point sources with a resolution of 0.1° in 2019, 2. The Emissions Database for Global Atmospheric Research (EDGAR), which provides sectoral emissions as specified by the IPCC guidelines. 3. The NO_2 thermal chemical vapor deposition retrieval product from the TROPOspheric Monitoring Instrument (TROPOMI) onboard the Sentinel-5 Precursor satellite. Given that GID has higher data quality in fine-grained spatial scales and point sources of industries and power plants, the GID-based point sources and the EDGAR emission maps were combined for constructing GRACED (Eq. 9). While the spatial emission patterns derived from GID and EDGAR (with latest updates in 2019) cannot accurately reflect the situation in 2020 and 2021, the NRT TROPOMI NO_2 retrievals were used as a proxy for CO_2 to capture the daily variability in CO_2 emission following GRACED[40]. After several data processing steps, such as rolling-average and thresholding, the NO_2 data can reasonably indicate the spatial distribution of CO_2 sources[48]. Table 3 lists the gridded data used for producing GRACED. For the aviation sector, EDGAR's monthly data are used for spatial distribution (Eq. 10). Thus, the gridded emissions $EmiGrid_{g,d,s}$ for grid g, date d and sector s were estimated as:

$$EmiGrid_{g,d,s} = CM_{c,d,s}\frac{GID_{g,s}}{\sum_{i=1}^{n} GID_{i,s}}\frac{EDGAR_{g,m,s}}{\sum_{j=1}^{12} EDGAR_{g,j,s}} \cdot 12, \tag{9}$$

$$EmiGrid_{g,d,avi} = CM_{c,d,avi}\frac{EDGAR_{g,m,avi}}{\sum_{i=1}^{n} EDGAR_{i,m,avi}}, \tag{10}$$

where $CM_{c,d,s}$ represents the value of Carbon Monitor national emission for country c, day d and sector s. s includes the power, industry, residential, and ground transport sectors. avi is the aviation sector. $GID_{g,s}$ is the value of GID gridded CO_2 emissions for grid g and sector s. n is the total number of grids within this country and j is the total number of month. $EDGAR_{g,m,s}$ represents the EDGAR gridded CO_2 emissions for grid g, sector s, and month m which date d belongs to.

City-level spatial disaggregation. The spatial disaggregation is performed by first converting the city area shapefiles (FUA or GADM) into raster datasets with a unique ID assigned to each city. Then the raster city area maps are used as masks to extract the matching grid cells in GRACED emission maps. We then aggregate emission values for grid cells that correspond to the same city mask to yield the total sectoral emission value for a given city. For the aviation sector, emissions from all planes within the city's territory are included. The international shipping sector is not included in this dataset because most of the emissions from this sector occurred in the open ocean that cannot be allocated to specific cities. The jurisdiction issue also applies to the aviation sector, but we keep the territorial-based allocation approach in the dataset for completeness.

We use both the administrative areas and FUA because boundary definition has always been a problem in city-level inventory completion[17], as the administrative city areas in most countries do not reflect emissions from the larger commuting zones of a city, which may constitute a large part of the emissions, meanwhile, FUA represents the most viable spatial dataset for covering the more complete urban areas. In addition, FUA is clearly-defined and produced using a consistent method for cities worldwide, while the definition of administrative city areas may vary significantly across different countries. Therefore, the use of both spatial scopes represents a potential solution to differentiate administrative emissions versus community-wide emissions and makes inter-dataset comparisons easier as demonstrated in the validation section.

City-level corrections. The disaggregation from EDGAR spatial distributions is insufficient especially for the residential and ground transport sectors, because EDGAR uses a disaggregation of national sectoral totals per population for residential, and per road network for ground transport, which introduces bias to cities. Therefore, we correct these two sectors at individual city level. The ground transport sector emission is corrected using city-specific TomTom data (by applying Eqs. 4,5 at city scale) for 416 cities worldwide that have their own NRT TomTom indices (list of these cities can be found in the documentation on Carbon Monitor website), which

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Dataset	CM-Cities	CEADs	MEIC	CDP-ICLEI Track	Vulcan
Spatial coverage	Global cities	China national, provincial, prefectural	China national, provincial	Global cities	U.S. counties
Temporal coverage	2019–2021	1997–2019	2000–2017	2010–2021	2010–2015
Temporal resolution	Daily	Monthly	Annual	Annual	Annual, hourly
Protocol	—	—	—	Various	—
Overall uncertainty	±21.7%	−15% to 30%	±15%	All data is self-reported, CDP-ICLEI Track does not assess the uncertainty	Sectoral uncertainties provided below
Area definition	GADM, FUA	—	Population density, GDP	Mostly city administrative, some include adjacent areas	Administrative county area
Total emissions comparison (with CM-Cities)	—	$R^2=0.96$, $Rd=11\%$, $n=30$	—	$R^2=0.74$, $Rd=31\%$, $n=24$	$R^2=0.82$, $Rd=26\%$, $n=50$
Power sector method	Daily power generation downscaling. $\Delta=\pm10\%$	Energy consumption for production and supply of electric power, steam and hot water	Unit-level power generation. $\Delta=-15\%$ to 16%	City report (scope 1–3 for relevant GPC stationary energy subsectors, including residential and commercial buildings, industry, agriculture, forestry and fishing)	CAMD, DOE/ EIA fuel, EPA NEI point electricity production. $\Delta=\pm13\%$
Power comparison (with CM-Cities)	—	$R^2=0.76$, $Rd=30\%$, $n=30$	$R^2=0.93$, $Rd=21\%$, $n=30$	—	$R^2=0.60$, $Rd=114\%$, $n=50$
Industry sector method	Industrial production index downscaling. $\Delta=\pm36\%$	Energy consumption for individual manufacturing sectors	—	City report (direct scope 1 emissions from industrial processes and product use)	EPA NEI industrial point sources. $\Delta=\pm12.8\%$
Industry comparison (with CM-Cities)	—	$R^2=0.92$, $Rd=28\%$, $n=30$	—	—	$R^2=0.58$, $Rd=67\%$, $n=50$
Residential sector method	HDD. $\Delta=\pm40\%$	—	—	City report (scope 1–3 for relevant GPC stationary energy subsectors, including residential and commercial buildings)	EPA NEI residential and commercial nonpoint buildings. $\Delta=\pm12.8\%$
Residential comparison (with CM-Cities)	—	—	—	—	$R^2=0.82$, $Rd=35\%$, $n=50$
Ground transport sector method	TomTom congestion index. $\Delta=\pm9.3\%$	—	Vehicle ownership statistics and digital road map	City report (scope 1–3 for GPC transportation subsectors, including on-road, railways, waterborne navigation, aviation, and off-road)	EMFAC, EPA NEI onroad. $\Delta=\pm14.2\%$
Ground transport comparison (with CM-Cities)	—	—	$R^2=0.62$, $Rd=31\%$, $n=30$	—	$R^2=0.90$, $Rd=41\%$, $n=50$
Aviation sector method	Flightradar24 flight data. $\Delta=\pm10.2\%$	—	—	City report aviation under transportation sector	EPA NEI point airport. $\Delta=\pm7.8\%$
Aviation comparison (with CM-Cities)	—	—	—	—	$R^2=0.69$, $Rd=58\%$, $n=50$
References	—	[29]	[30,31]	—	[4]

Table 5. Summary of city emission datasets (Δ is uncertainty) and comparison statistics including coefficient of determination (R^2), mean relative difference (Rd), and sample size (n) when compared with CM-Cities.

represent more accurate ground transport emission estimates for these cities. For cities that do not have their own TomTom data, we spatially disaggregate the national mean estimates following Eqs. 4, 5, 9.

The residential sector is corrected using city-level HDD to overcome the bias of downscaling from national inventory. Specifically, we first calculate the daily mean HDD for each city from the population-weighted HDD grid (Eq. 7), and then use it as the baseline to compute a correction factor for each city by comparing it with the mean national HDD to update the emissions for the residential sector:

$$Emis_{res,i} = Emis0_{res,i}\left(1 - \frac{HDD_c - HDD_i}{max(HDD_c)}\right), \tag{11}$$

where $Emis_{res,i}$ and $Emis0_{res,i}$ represent the corrected and uncorrected residential emissions, respectively for city i, HDD_c is the mean daily HDD for the country, and HDD_i is the mean daily HDD for city i.

Outlier correction. Outliers exist in the data mainly due to errors in the source datasets, such as mistakes in unit conversions or data entry, etc. To correct these outliers, we apply a statistical method based on intrinsic properties of the distribution of the emissions in the database. This allows more accurate identification of outliers

Content courtesy of Springer Nature, terms of use apply. Rights reserved

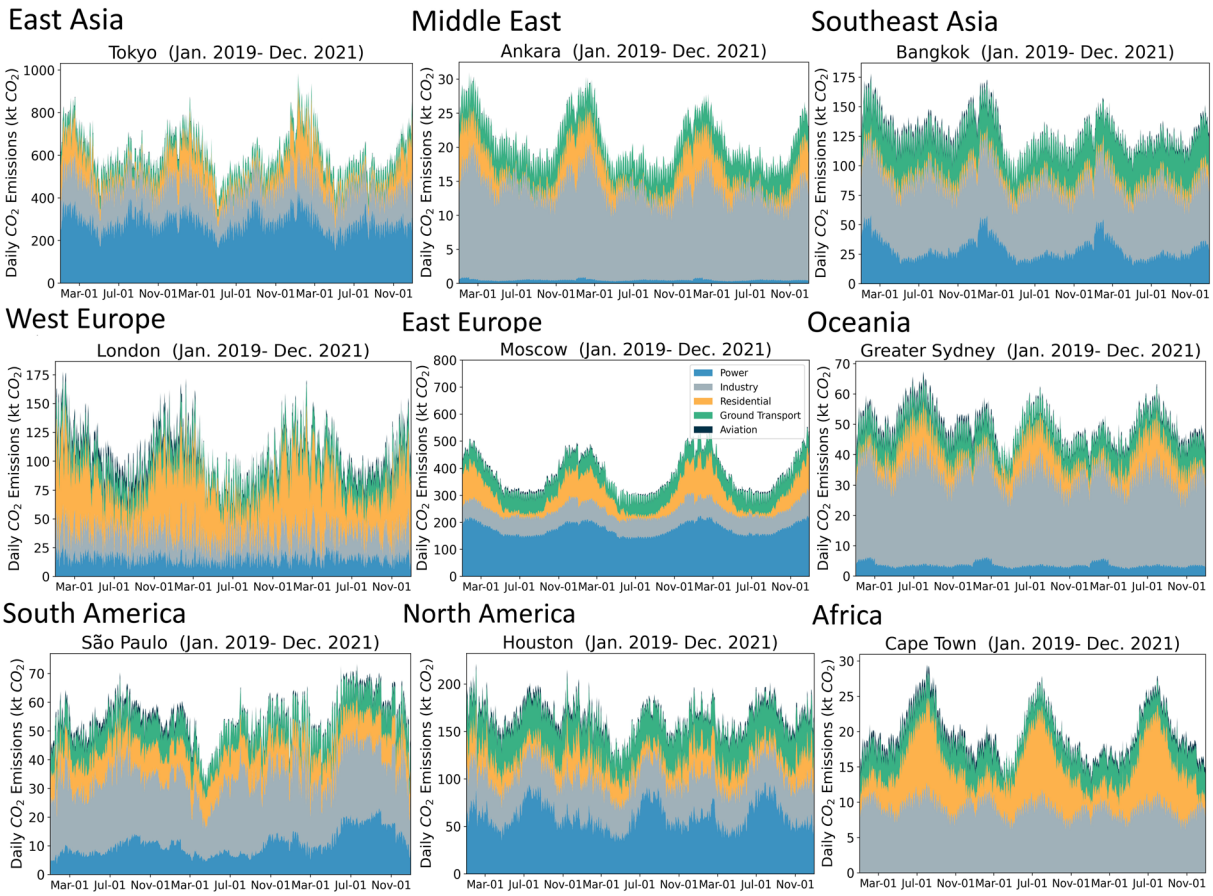

Fig. 4 Daily by sector CO_2 emissions for cities (FUA) in different regions of the world. Including Tokyo in East Asia, Ankara in the Middle East, Bangkok in Southeast Asia, London in West Europe, Moscow in East Europe, Greater Sydney in Oceania, SÃ£o Paulo in South America, Houston in North America, and Cape Town in Africa.

that are likely to be the results of incorrect data entry. Similar statistical approaches have been successfully applied to correct for outliers in emission datasets[25]. The outlier identification method is based on standard deviation (STD). Specifically, we consider an emission value as an outlier if the differences between the current value and its daily neighbours are both greater than 3 times the yearly STD for that sector (Eq. 12). This threshold is determined by experimenting with data with known error and data with periodical high variation (e.g., weekday versus weekends for the ground transport emissions). These experiments determined the lower and upper bounds of the threshold such that it correctly identifies outliers and keeps the inherent variance within the data.

$$|x - x_{n+1}| > 3 \cdot STD \; and \; |x - x_{n-1}| > 3 \cdot STD. \tag{12}$$

Limitations and future work. This dataset focuses on improving the timeliness, temporal resolution, and coverage of city-level inventories for studying NRT emission dynamics and also providing emission estimates for many cities in low-income regions. This dataset does not account for emissions related to land use, land use change, waste, and forestry, therefore, some emissions caused by long-term urban expansion are not captured. This dataset is constructed based on daily activity and models that can cover a majority rather than the entire daily emission-related activities due to data availability. Therefore, we acknowledge that a small portion of daily variations in city emissions are not reflected in this dataset. This dataset is derived from the gridded Carbon Monitor which is based on downscaled national inventories, combined with point sources and spatial distributions from GID and EDGAR, therefore, one limitation is the lack of using city-specific bottom-up activity data except for the ground transport sector, which may introduce additional uncertainties. We also noted that some input data may contain inherent errors and missing values (other than the above-mentioned outliers), especially for cities in less developed nations, we do not intend to fix this kind of errors in the source data without enough background information of the specific city, but we consider our results represent a meaningful first-order estimate for many of these cities that are lacking any emission inventories.

Estimating NRT daily emissions for cities is a relatively new research direction and requires ongoing efforts to calibrate and update the workflow to improve data quality in the future. Further validation of the data is a crucial next step, and we plan to conduct more data validation and quality improvements in the future through

Content courtesy of Springer Nature, terms of use apply. Rights reserved

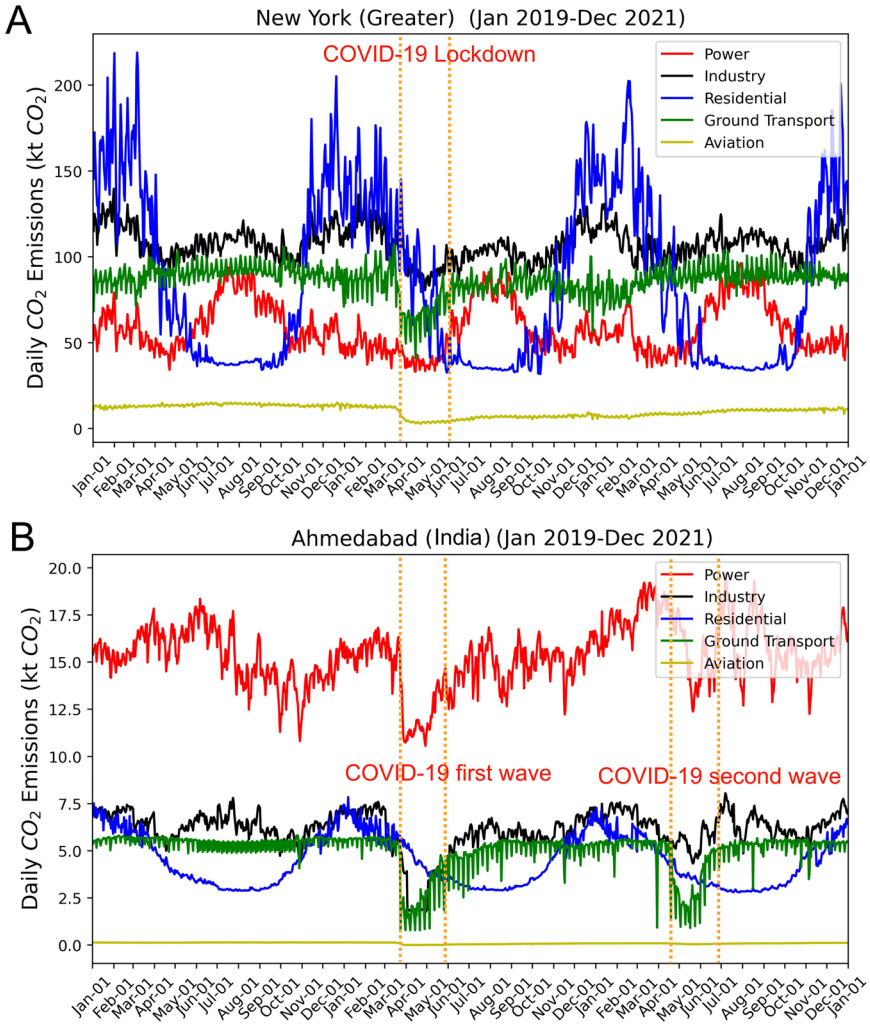

Fig. 5 Daily city-level CO_2 emissions show the impact of COVID-19 for (**A**) Greater New York in the U.S. and (**B**) Ahmedabad in India. The emissions from ground transportation and aviation decreased significantly during the lockdown period (between the dashed lines) in spring 2020, and also during the second wave between March 2021 to June 2021 in India.

multiple follow-up works. From a bottom-up perspective, we are collecting more city-level fossil fuel consumption data to better constrain the annual or monthly total emissions. From a top-down perspective, we plan to compare our results with field observations and satellite retrieval data. As proposed by previous study[1], remotely sensed urban atmospheric measurements can help us estimate and predict CO_2 emission fluxes, and we plan to leverage these research outcomes to improve this city-level emissions dataset. This will require a lot more effort to collect and harmonize inventories with observations[49], but some progress is being made, for example, several observation systems are being designed to monitor megacity CO_2 domes and surface-based observations of atmospheric CO_2 are commercially available for some major cities[1]. Future work will compare our results with observations from a set of surface, airborne, and satellite sensors. which would provide a foundation for more accurate validation of bottom-up city emission inventories.

Data Records

CM-Cities provides scope-1 NRT city-level emission inventories from 01/01/2019 to 31/12/2021 for 1500 cities in 46 counties. All data have gone through a validation process, in which we estimated the uncertainties and corrected errors. The attributes of the final dataset are listed in Table 4, and the emission data are organized into spreadsheets. The definitions for sectors are consistent with the Carbon Monitor national inventories. Brief descriptions of the methods, sectors, coverage and uncertainty are also provided in Table 5. Latest updates for selected cities and related information are available for view and download on our website https://cities.carbonmonitor.org. At the time of writing this article, this dataset has been updated to December 31, 2021 and the full dataset can be downloaded at Figshare[50]. Future updates will also be available on our website.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

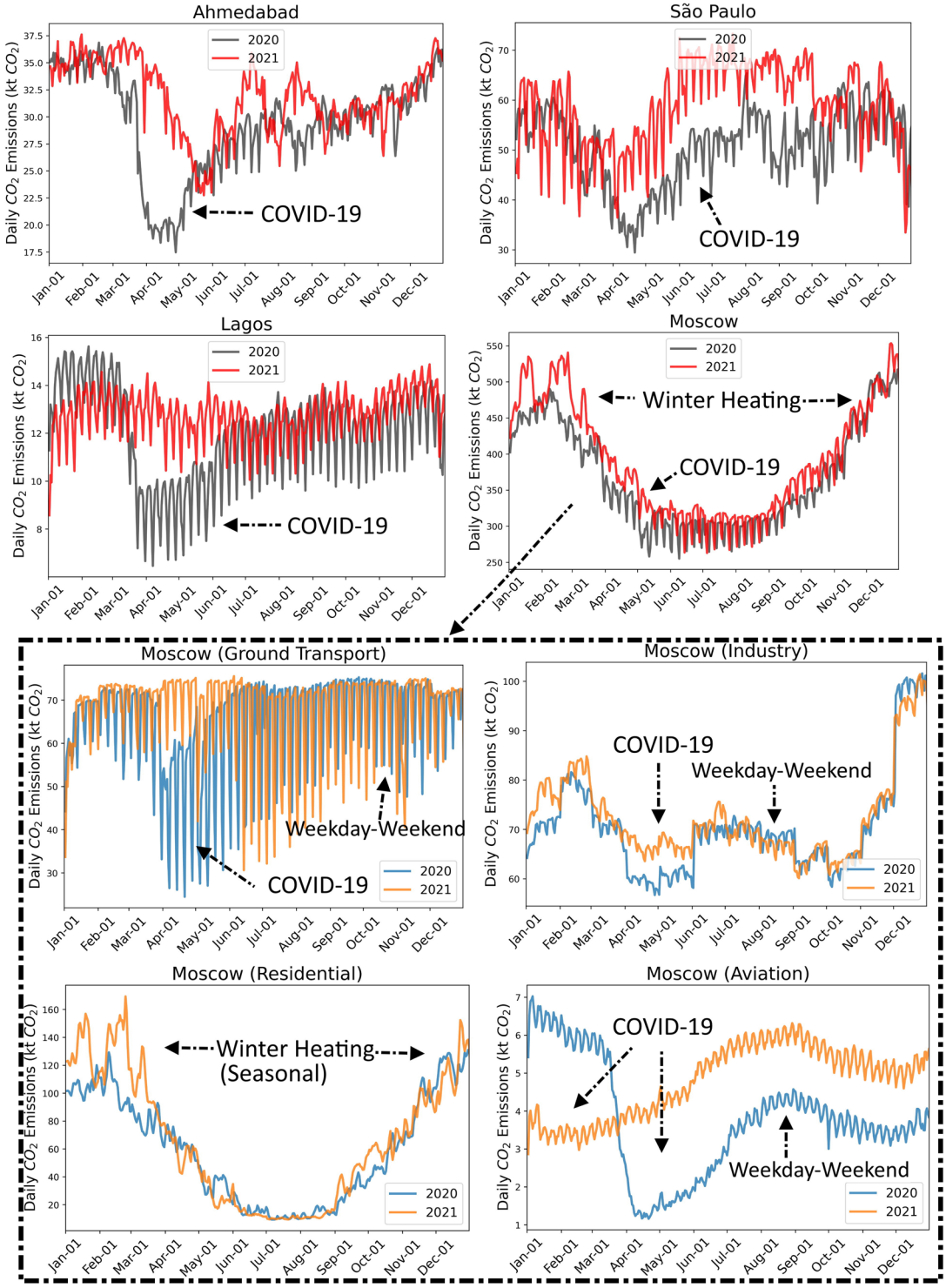

Fig. 6 Daily total CO_2 emissions for selected cities. Gray lines depict daily emissions for the year 2020 and red lines depict daily emissions for the year 2021. The impact of the COVID-19 pandemic on city-level emissions is highlighted. Subplots for Moscow show the seasonal and weekly emission patterns for each sector, which demonstrates the advantage of the high temporal resolution.

- The file that contains functional urban area results for all cities (carbon-monitor-cities-all-cities-FUA.csv) has 8,114,886 data records (some cities have missing values). Separate data files are also provided for each of the 46 countries (carbon-monitor-cities-"CountryName".csv).
- The file that contains all administrative area results for Chinese cities (carbon-monitor-cities-China-GADM-prefecture.csv) has 1,885,120 data records, including 344 prefecture-level cities, and each city has 5480 data records.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

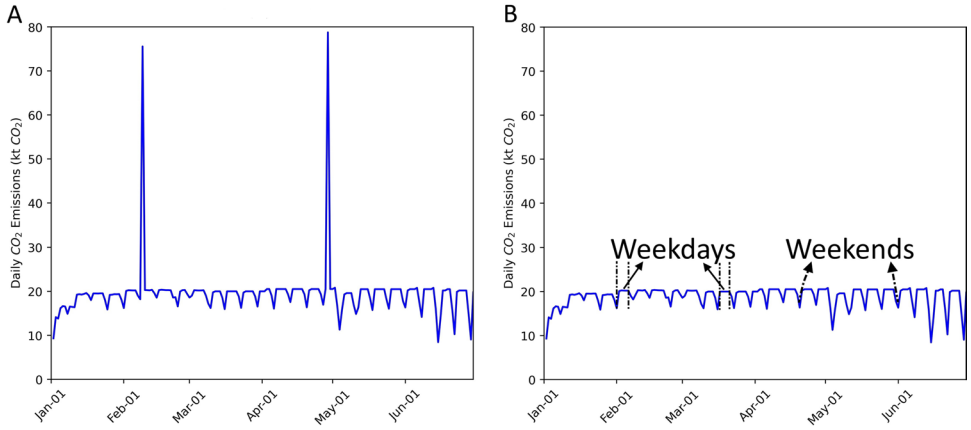

Fig. 7 Examples of outlier identification and correction for the ground transport data. (**A**) Two outliers clearly fall out of the typical range of weekday-weekend variation before the correction. (**B**) Outliers removed after the correction.

- The file that contains U.S. county-level results (carbon-monitor-cities-US-Counties.csv) has 1,720,720 data records, including 314 counties, and each county has 5480 data records.

Data examples. Daily CO_2 emission variations from a city reveal its geographic and socio-economic characteristics. Figure 4 shows the sectoral breakdown of daily CO_2 emissions for some major cities in different regions of the world, including East Asia, Middle East, Southeast Asia, West Europe, East Europe, Oceania, South America, North America, and Africa. As an example of the geographic influence on the emissions, cities in the Southern Hemisphere, such as Sydney in Australia, and Cape Town in Africa, exhibit higher emissions in the residential sector during the southern winter (northern summer) due to the increase in heating demand. Daily emissions also reveal certain events such as holidays and the COVID-19 outbreak. The emissions from the power sector show a surge in summer for many cities, which is likely due to the increased power consumption for cooling. As an example, we show the impact of COVID-19 on city emissions for Greater New York in the U.S. and Ahmedabad in India (Fig. 5), note the significant drop in emissions for the ground transport sector in spring 2020 (as indicated by the dashed lines) during the lockdown period and during the 2021 second wave of COVID-19 pandemic in India.

Figure 6 depicts the total daily emissions for the year 2020 versus the year 2021 for selected cities. By comparing the emissions in spring 2020 and spring 2021, we noted that for these cities, emissions rebound from the lower levels caused by the COVID-19 pandemic. Subplots for the city of Moscow are presented here as an example to show the seasonal and weekly patterns of the city-level emissions for different sectors, which highlights the advantage of the low latency and demonstrates that this high temporal resolution dataset can be very useful for investigating weekly and seasonal variations in city emissions.

Technical Validation

The quality of this dataset is evaluated by comparing it against existing datasets (Table 5). We also performed uncertainty analysis for our data and for each sector based on a synthesis analysis of input data uncertainties and the methodology used. Significant outliers were identified and corrected as shown in the examples of Fig. 7. The outlier occurrence rate for this dataset is 0.012%.

Validation against other datasets. Multiple datasets are used to validate our results, including 1. City inventories from the CDP-ICLEI Track, 2. Vulcan dataset for US counties, 3. CEADs and MEIC dataset for China, and 4. individual reports released by city governments. Note that only scope-1 emissions are compared. For cities in China, we validated our dataset by comparing it with the CEADs and MEIC datasets. CEADs provides annual provincial emission inventories for China for 2019, and we validated the data for each province by summing up emissions from all prefecture-level cities in each province (for China, the GADM level-2 is exactly the area of prefecture-level cities, and the sum of all prefecture-level cities within a province equals the total area of that province). Figure 8 depicts the comparison results for all the Chinese provinces including municipalities and most autonomous regions. The result indicates a good agreement between the CEADs and CM-Cities, with less than 10% difference in annual emissions for most of the provinces. Statistics (Table 5) indicate that the coefficient of determination (R^2) values between CEADs and CM-Cities are 0.96, 0.76, and 0.92 for total, power, and industry sectors, respectively, and the corresponding mean relative difference (Rd) are 11%, 30%, and 28%. R^2 values between MEIC and CM-Cities are 0.93 and 0.62 for the power sector and ground transport sector, respectively, and the corresponding Rd are 21% and 31%. Other sectors were not compared due to the large differences in sector definition and coverage. The mean relative differences are all within the uncertainty ranges, which indicates a relatively high accuracy for Chinese cities.

Content courtesy of Springer Nature, terms of use apply. Rights reserved

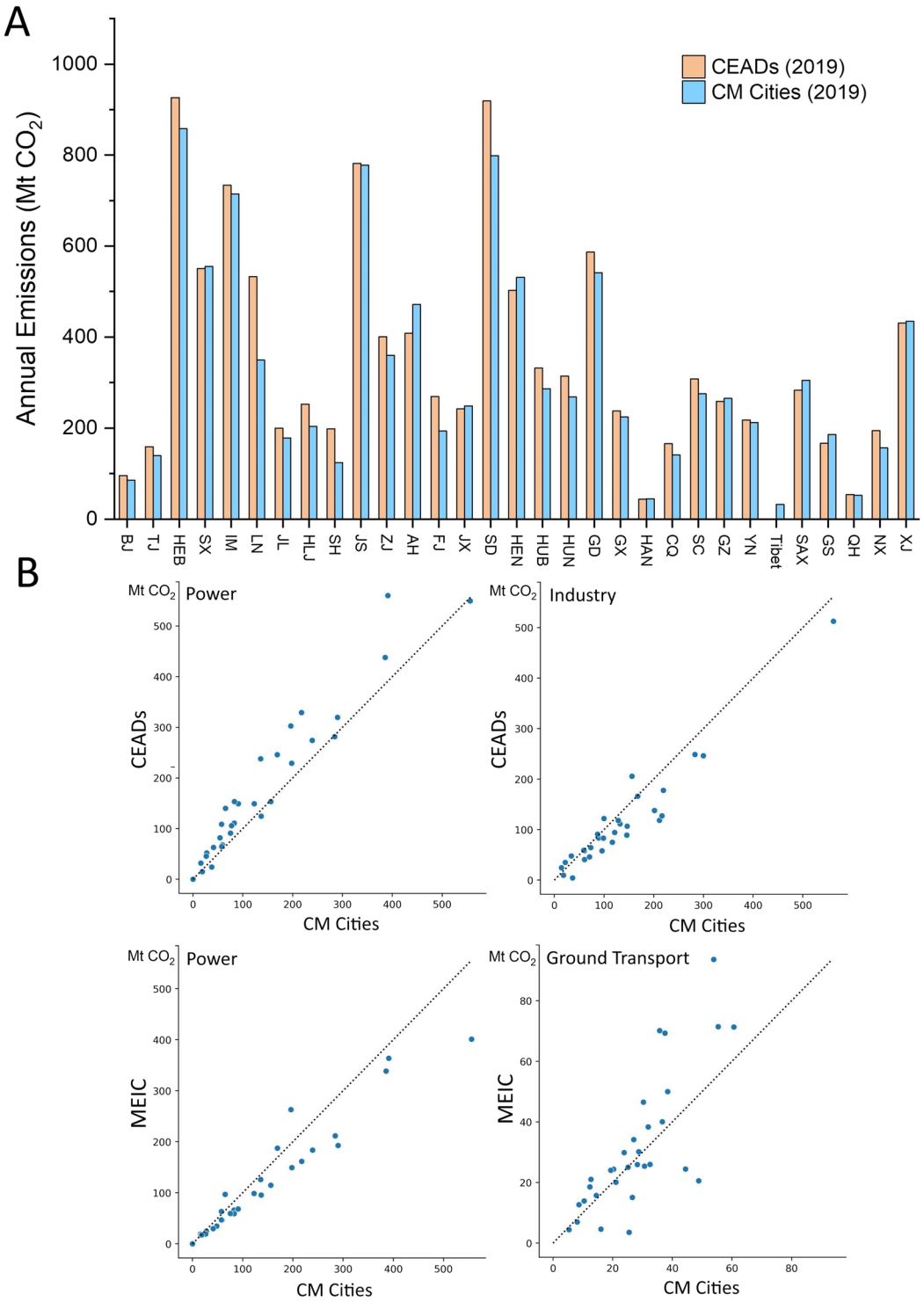

Fig. 8 Dataset comparison for cities in China. (**A**) Comparison of the sum of all prefecture-level cities within each Chinese province (including municipalities and autonomous regions) against the CEADs provincial datasets for year 2019. Note that the sum of all prefecture-level cities within a province equals the total area of that province in China (**B**) Sectoral comparison between CM-Cities, CEADs and MEIC for sectors that have similar coverages.

For cities in other countries, many datasets do not have recent (2019 or later) inventories, for example, the latest Vulcan dataset provides emissions in 2015 for United States counties, and the latest CDP-ICLEI Track inventories may range from 2010 to 2021 depending on reporting status of each city. For the completeness of the validation, we adjusted the area of accounting for CM-Cities to be as consistent as possible with these

Content courtesy of Springer Nature, terms of use apply. Rights reserved

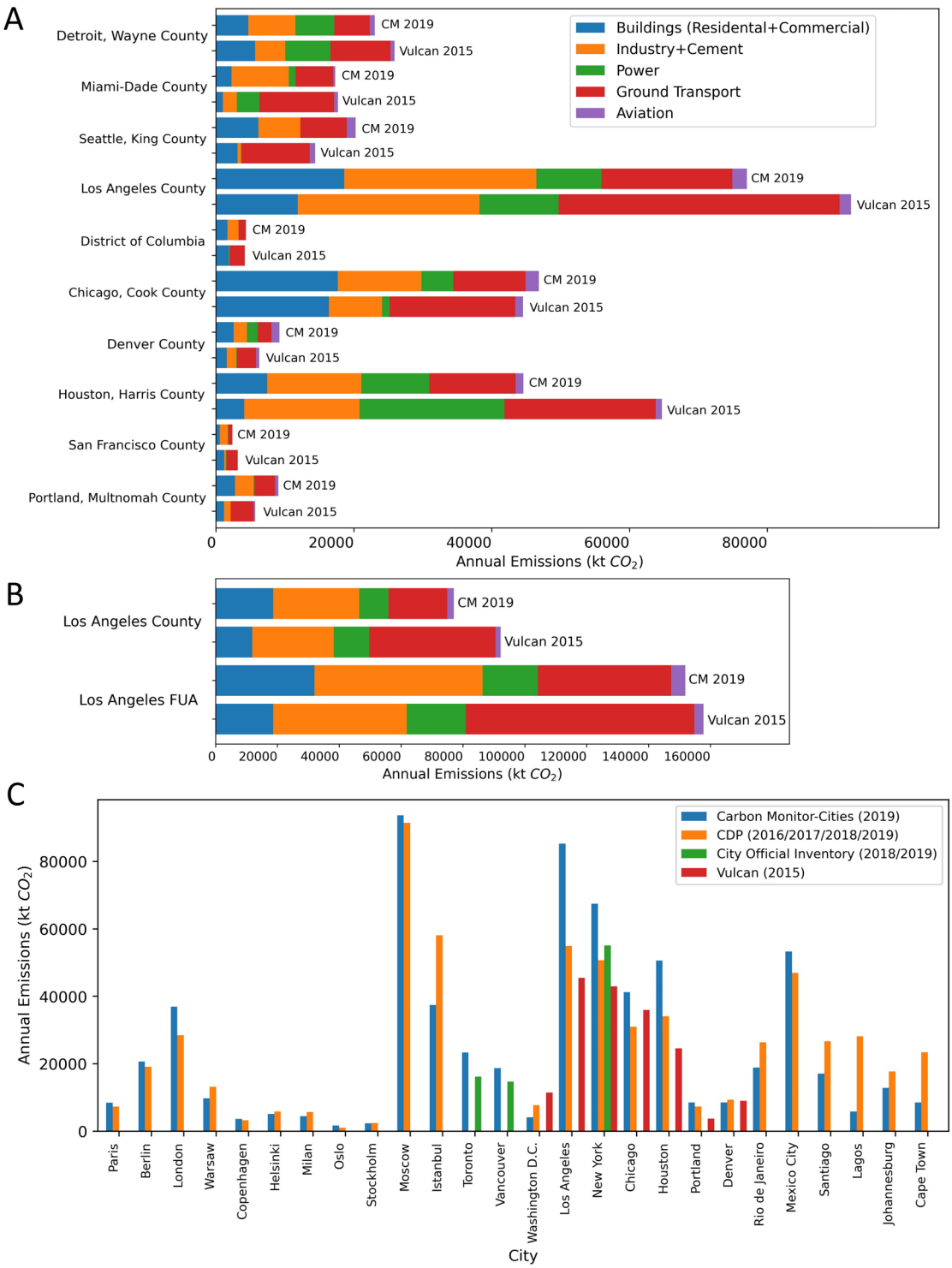

Fig. 9 Dataset comparison results. (**A**) Comparison of sectoral emissions between Vulcan and CM-Cities for selected counties in the United States, and (**B**) county-level and FUA-level comparison for Los Angeles. The year of accounting is 2015 for Vulcan inventories and 2019 for CM-Cities, which could partially explain the differences. (**C**) City annual total emission comparisons between CM-Cities, CDP-ICLEI Track, Vulcan and some other city self-reported inventories. Magnitudes represent total emissions from each dataset. The area of accounting is adjusted to be as consistent as possible across datasets.

datasets and compared city inventories from all available data sources regardless of the time of accounting. Note that GADM level-2 in the United States represents exactly the area of counties, so our GADM results were used

Content courtesy of Springer Nature, terms of use apply. Rights reserved

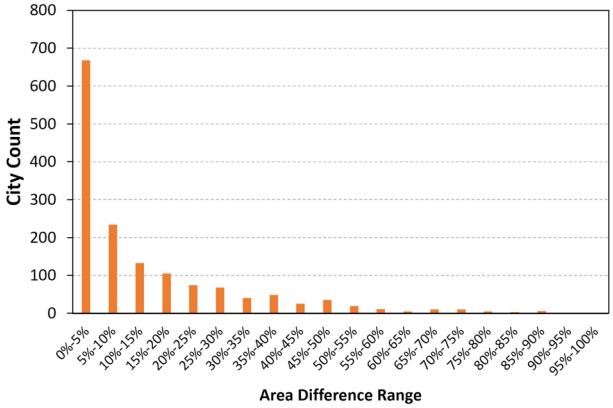

Fig. 10 The frequency distribution of city area (boundary) uncertainty ranges in the dataset. The mean area uncertainty for all cities is 13.55%.

for comparison with Vulcan county-level inventories. Figure 9 show examples of the annual total emission comparisons between CM-Cities and these datasets.

Comparison between Vulcan (2015) and CM-Cities (2019) covers top 50 counties with the highest emissions in the United States. We used the coefficient of determination (R^2) and mean relative difference (Rd) to evaluate the comparison results. R^2 values are 0.82, 0.60, 0.58, 0.82, 0.90, and 0.69 for total, power, industry, residential (and commercial buildings), ground transport, and aviation sectors, respectively, and the corresponding Rd values are 26%, 114%, 67%, 35%, 41%, and 58%, respectively (Table 5). The differences are mainly due to 1. the difference in the year of accounting, as the earliest estimates of CM-Cities for the year 2019 is compared to the latest Vulcan for the year 2015, multiple factors that govern emissions could have changed during the period, 2. the different accounting methods, as CM-Cities uses a territorial downscaling approach, while Vulcan uses a consumption-based bottom-up accounting approach, and 3. the differences in sector coverage definitions and source data (Tables 1, 2, 5), which partially explains why the total emission comparison show a better good agreement than the sectoral comparisons.

Direct comparisons with CDP-ICLEI Track were difficult due to several reasons: 1. CDP-ICLEI Track inventories are city self-reported data, which were typically estimated using different methods. 2. Most cities follow the GPC protocol and report in scopes rather than in sectors, therefore, we only compared total emissions. 3. CDP-ICLEI Track has not independently calculated the uncertainty range for these self-reported inventories, and self-reported uncertainties are expected to be variable. For example, 45% of cities reported "high confidence" in their emissions data for 2021, 35% reported "medium confidence", and 3% reported "low confidence". 4. Spatial coverage is unclear, as the definition of a "city" can vary across different countries, some cities report based on administrative areas, but others include adjacent areas, but no shapefiles or raster maps were provided to clarify the exact city boundary or area of accounting. Nonetheless, we performed comparisons for 24 large cities in different regions with available CDP-ICLEI Track inventories, the total emission comparisons for these cities show an agreement with $R^2 = 0.74$ and $Rd = 31\%$ (Fig. 9, Table 5).

Uncertainty analysis. The uncertainties in this dataset have two sources: 1. The uncertainties inherited from Carbon Monitor and GRACED. 2. The uncertainty introduced by the spatial downscaling process. The uncertainty analysis was conducted based on the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. For the power sector, uncertainty mainly comes from the emission factor and the variability of energy mix for power generation, the 1-sigma uncertainty of power emission from fossil fuel is estimated as ±10.0%. For the industry sector, monthly production data is the main source of uncertainty, especially the production in China, which accounts for more than 60% of world total industrial CO_2 emissions. Monte Carlo simulations were used to determine the confidence interval based on regression models between estimated monthly emissions and officially reported emissions. The 1-sigma uncertainty for the industry sector is estimated as ±36.0%. For the ground transport, uncertainty is estimated by applying the regression between the TomTom congestion index and traffic flux to other cities (other than Paris). The 1-sigma uncertainty for the ground transport sector is estimated as ±9.3%. For the residential sector, the uncertainty is calculated based on comparisons between estimated emissions and consumption-based accounting results for several countries in Europe. The 1-sigma uncertainty for the residential sector is estimated as ±40.0%. For the aviation sector, 1-sigma uncertainty is estimated as ±10.2%. These uncertainty estimates follow the methods used by Carbon Monitor[10].

Spatial downscaling introduces additional uncertainty because of the rasterization of city areas. Spatial computations are based on raster (gridded) files, but most cities and urban areas have irregular-shaped boundaries that are not fully overlapped with gridded cells. Area discrepancies are found along all city boundaries, and smaller cities typically suffer from higher levels of dissimilarities because few grid cells account for a large portion of the total urban area. We computed the area discrepancies for all cities in the dataset (Fig. 10), and found that 44.53% of cities show an area difference of 0%–5% and the count decreases as the discrepancy percentage gets higher. The mean area discrepancy for all cities is 13.55%. We then estimated the overall uncertainty by first

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Sectors Daily emissions uncertainty	
Power	12% to 43%
Industry	21% to 32%
Ground transport	19% to 27%
Residential	26% to 35%
Aviation	10% to 21%

Table 6. Uncertainty estimations at daily scale for each sector.

applying the error propagation equation provided by IPCC[51], and then combining the uncertainties of all sectors and the city area uncertainty:

$$U_{overall} = \sqrt{\frac{\sum(U_s\alpha_s)}{\sum\alpha_s^2}} + U_a, \tag{13}$$

where U_s and a_s are the percentage and quantity (daily mean emissions) of the uncertainty for sector s, respectively, and U_a is the city area uncertainty. Finally, the overall annual uncertainty range of CM-Cities is estimated as $\pm 21.66\%$.

Uncertainties at a daily scale are also estimated for each sector (Table 6). For example, the daily uncertainty from the power sector is estimated by comparing our results with emissions from real load curve data in several cities. Daily uncertainty for the ground transportation sector has two components, the regression model and the daily allocation of CO_2 emissions by traffic flow. The uncertainty quantification of the daily-scale allocation of emissions requires real daily emissions from ground transportation, which is difficult to obtain and ignored in this study. Therefore, we focus on the uncertainty generated by the regression model. We used the 95% confidence interval of the regression model to estimate the uncertainty generated by the model. Given the high temporal resolution of this dataset, the uncertainties from daily activity will increase the overall uncertainties on top of the annual uncertainties, which is not inconsistent, as data are temporally auto-correlated.

Usage Notes

The generated datasets[50] are available from https://doi.org/10.6084/m9.figshare.19425665.v1. The main data file has more than one million lines of data, which will take a long time to load in Excel. We recommend loading the data with a script that can handle large datasets. We have provided an example of Python code to help users read in and plot emissions for any city in the dataset (https://github.com/dh107/Carbon-Monitor-Cities/). Note that the raw TomTom and flightradar24 data are not included in this dataset as we only provide the estimated emissions. Users should also note that the unit of emissions in this dataset is $ktCO_2$. Filename indicates whether the data is based on administrative areas (GADM) or functional urban areas (default). The next update to this dataset is scheduled for May 2022, which will update the dataset to Feb 28, 2022.

Code availability

Python code for producing, reading and plotting data for any city in the dataset is provided at https://github.com/dh107/Carbon-Monitor-Cities/.

Received: 31 March 2022; Accepted: 22 August 2022;
Published online: 01 September 2022

References

1. Duren, R. M. & Miller, C. E. Measuring the carbon emissions of megacities. *Nature Climate Change* **2**, 560–562, https://doi.org/10.1038/nclimate1629 (2012).
2. Seto, K. C. *et al. Climate Change 2014: Mitigation of Climate Change Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change - Human settlements, infrastructure and spatial planning.* (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014).
3. Revi, A. *et al. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change - Urban Areas*, book section 8, 535–612 (Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014).
4. Gurney, K. R. *et al.* The vulcan version 3.0 high-resolution fossil fuel CO2 emissions for the United States. *Journal of Geophysical Research: Atmospheres* **125**, e2020JD032974, https://doi.org/10.1029/2020JD032974 (2020).
5. Gurney, K. R. *et al.* Under-reporting of greenhouse gas emissions in U.S. cities. *Nature Communications* **12**, 553, https://doi.org/10.1038/s41467-020-20871-0 (2021).
6. Seto, K. C. *et al.* From low- to net-zero carbon cities: The next global agenda. *Annual Review of Environment and Resources* **46**, 377–415, https://doi.org/10.1146/annurev-environ-050120-113117 (2021).
7. Kona, A., Bertoldi, P., Monforti-Ferrario, F., Rivas, S. & Dallemand, J. F. Covenant of mayors signatories leading the way towards 1.5 degree global warming pathway. *Sustainable Cities and Society* **41**, 568–575 (2018).
8. Nangini, C. *et al.* A global dataset of CO2 emissions and ancillary data related to emissions for 343 cities. *Scientific Data* **6**, 180280, https://doi.org/10.1038/sdata.2018.280 (2019).
9. Chen, J. *et al.* China's city-level carbon emissions during 1992–2017 based on the inter-calibration of nighttime light data. *Scientific Reports* **11**, 3323, https://doi.org/10.1038/s41598-021-81754-y (2021).
10. Liu, Z. *et al.* Carbon monitor, a near-real-time daily dataset of global CO2 emission from fossil fuel and cement production. *Scientific Data* **7**, 392, https://doi.org/10.1038/s41597-020-00708-7 (2020).

Content courtesy of Springer Nature, terms of use apply. Rights reserved

11. Liu, Z. *et al*. Global patterns of daily CO2 emissions reductions in the first year of COVID-19. *Nature Geoscience* https://doi.org/10.1038/s41561-022-00965-8 (2022).
12. Ramaswami, A. *et al*. Carbon analytics for net-zero emissions sustainable cities. *Nature Sustainability* **4**, 460–463, https://doi.org/10.1038/s41893-021-00715-5 (2021).
13. Bulkeley, H. Cities and the governing of climate change. *Annual Review of Environment and Resources* **35**, 229–253, https://doi.org/10.1146/annurev-environ-072809-101747 (2010).
14. D'Avignon, A., Carloni, F. A., Rovere, E. L. L. & Dubeux, C. B. S. Emission inventory: An urban public policy instrument and benchmark. *Energy Policy* **38**, 4838–4847, https://doi.org/10.1016/j.enpol.2009.10.002 (2010).
15. Tan, S. *et al*. A holistic low carbon city indicator framework for sustainable development. *Applied Energy* **185**, 1919–1930, https://doi.org/10.1016/j.apenergy.2016.03.041 (2017).
16. Ramaswami, A., Chavez, A., Ewing-Thiel, J. & Reeve, K. E. Two approaches to greenhouse gas emissions foot-printing at the city scale. *Environmental Science & Technology* **45**, 4205–4206, https://doi.org/10.1021/es201166n (2011).
17. Chen, G. *et al*. Review on city-level carbon accounting. *Environmental Science & Technology* **53**, 5545–5558, https://doi.org/10.1021/acs.est.8b07071 (2019).
18. Wiedmann, T. *et al*. Three-scope carbon emission inventories of global cities. *Journal of Industrial Ecology* **25**, 735–750, https://doi.org/10.1111/jiec.13063 (2021).
19. Moran, D. *et al*. Estimating CO2 emissions for 108,000 European cities. *Earth Syst. Sci. Data Discuss.* **2021**, 1–23, https://doi.org/10.5194/essd-2021-299 (2021).
20. Jing, Q., Bai, H., Luo, W., Cai, B. & Xu, H. A top-bottom method for city-scale energy-related CO2 emissions estimation: A case study of 41 Chinese cities. *Journal of Cleaner Production* **202**, 444–455, https://doi.org/10.1016/j.jclepro.2018.08.179 (2018).
21. Marcotullio, P. J., Sarzynski, A., Albrecht, J. & Schulz, N. A top-down regional assessment of urban greenhouse gas emissions in Europe. *AMBIO* **43**, 957–968, https://doi.org/10.1007/s13280-013-0467-6 (2014).
22. Crippa, M. *et al*. *Fossil CO2 and GHG emissions of all world countries* (Publications Office of the European Union, Luxembourg, 2019).
23. Crippa, M. *et al*. High resolution temporal profiles in the emissions database for global atmospheric research. *Scientific Data* **7**, 121, https://doi.org/10.1038/s41597-020-0462-2 (2020).
24. Fong, W. K. *et al*. *Global protocol for community-scale greenhouse gas emission inventories: An accounting and reporting standard for cities* (World Resources Institute, Washington, D.C., USA, 2015).
25. Kona, A. *et al*. Global covenant of mayors, a dataset of greenhouse gas emissions for 6200 cities in Europe and the Southern Mediterranean countries. *Earth Syst. Sci. Data* **13**, 3551–3564, https://doi.org/10.5194/essd-13-3551-2021 (2021).
26. Long, Y. *et al*. Monthly direct and indirect greenhouse gases emissions from household consumption in the major Japanese cities. *Scientific Data* **8**, 301, https://doi.org/10.1038/s41597-021-01086-4 (2021).
27. Kennedy, C. *et al*. Greenhouse gas emissions from global cities. *Environmental Science & Technology* **43**, 7297–7302, https://doi.org/10.1021/es900213p (2009).
28. Kennedy, C. *et al*. Methodology for inventorying greenhouse gas emissions from global cities. *Energy Policy* **38**, 4828–4837, https://doi.org/10.1016/j.enpol.2009.08.050 (2010).
29. Shan, Y., Huang, Q., Guan, D. & Hubacek, K. China CO2 emission accounts 2016–2017. *Scientific Data* **7**, 54, https://doi.org/10.1038/s41597-020-0393-y (2020).
30. Zheng, B. *et al*. High-resolution mapping of vehicle emissions in China in 2008. *Atmospheric Chemistry and Physics* **14**, 9787–9805, https://doi.org/10.5194/acp-14-9787-2014 (2014).
31. Liu, F. *et al*. High-resolution inventory of technologies, activities, and emissions of coal-fired power plants in China from 1990 to 2010. *Atmospheric Chemistry and Physics* **15**, 13299–13317, https://doi.org/10.5194/acp-15-13299-2015 (2015).
32. Croci, E., Lucchitta, B., Janssens-Maenhout, G., Martelli, S. & Molteni, T. Urban CO2 mitigation strategies under the covenant of mayors: An assessment of 124 European cities. *Journal of Cleaner Production* **169**, 161–177, https://doi.org/10.1016/j.jclepro.2017.05.165 (2017).
33. Gurney, K. R. *et al*. The hestia fossil fuel CO2 emissions data product for the Los Angeles megacity (Hestia-LA). *Earth Syst. Sci. Data* **11**, 1309–1335, https://doi.org/10.5194/essd-11-1309-2019 (2019).
34. Andrew, R. M. Timely estimates of india's annual and monthly fossil CO2 emissions. *Earth Syst. Sci. Data* **12**, 2411–2421, https://doi.org/10.5194/essd-12-2411-2020 (2020).
35. Wang, R., Assenova, V. A. & Hertwich, E. G. Energy system decarbonization and productivity gains reduced the coupling of CO2 emissions and economic growth in 73 countries between 1970 and 2016. *One Earth* **4**, 1614–1624, https://doi.org/10.1016/j.oneear.2021.10.010 (2021).
36. Liu, Z. *et al*. Near-real-time monitoring of global CO2 emissions reveals the effects of the COVID-19 pandemic. *Nature Communications* **11**, 5172, https://doi.org/10.1038/s41467-020-18922-7 (2020).
37. Le Quéré, C. *et al*. Temporary reduction in daily global CO2 emissions during the COVID-19 forced confinement. *Nature Climate Change* **10**, 647–653, https://doi.org/10.1038/s41558-020-0797-x (2020).
38. Liu, Z. *et al*. Near-real-time carbon emission accounting technology toward carbon neutrality. *Engineering* https://doi.org/10.1016/j.eng.2021.12.019 (2022).
39. Liu, Z., Deng, Z., Davis, S. J., Giron, C. & Ciais, P. Monitoring global carbon emissions in 2021. *Nature Reviews Earth & Environment* **3**, 217–219, https://doi.org/10.1038/s43017-022-00285-w (2022).
40. Dou, X. *et al*. Near-real-time global gridded daily CO2 emissions. *The Innovation* **3**, 100182, https://doi.org/10.1016/j.xinn.2021.100182 (2022).
41. Moreno-Monroy, A. I., Schiavina, M. & Veneri, P. Metropolitan areas in the world. delineation and population trends. *Journal of Urban Economics* **125**, 103242, https://doi.org/10.1016/j.jue.2020.103242 (2021).
42. Dijkstra, L., Poelman, H. & Veneri, P. The EU-OECD definition of a functional urban area, https://doi.org/10.1787/d58cb34d-en (2019).
43. Schiavina, M., Moreno-Monroy, A., Maffenini, L., Veneri, P. & Paolo. GHS-FUA R2019A–GHS functional urban areas, derived from GHS-UCDB R2019A, (2015), https://doi.org/10.2905/347F0337-F2DA-4592-87B3-E25975EC2C95 (2019).
44. Deng, Z. *et al*. Daily updated dataset of national and global CO2 emissions from fossil fuel and cement production. *figshare* https://doi.org/10.6084/m9.figshare.12685937.v4 (2020).
45. Spoladore, A., Borelli, D., Devia, F., Mora, F. & Schenone, C. Model for forecasting residential heat demand based on natural gas consumption and energy performance indicators. *Applied Energy* **182**, 488–499, https://doi.org/10.1016/j.apenergy.2016.0 (2016).
46. Copernicus Climate Change Service, C. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. *Copernicus Climate Change Service Climate Data Store (CDS)* (2019).
47. Doxsey-Whitfield, E. *et al*. Taking advantage of the improved availability of census data: A first look at the gridded population of the world, version 4. *Papers in Applied Geography* **1**, 226–234, https://doi.org/10.1080/23754931.2015.1014272 (2015).
48. Berezin, E. V. *et al*. Multiannual changes of CO2 emissions in China: indirect estimates derived from satellite measurements of tropospheric NO2 columns. *Atmos. Chem. Phys.* **13**, 9415–9438, https://doi.org/10.5194/acp-13-9415-2013 (2013).
49. Deng, Z. *et al*. Comparing national greenhouse gas budgets reported in UNFCCC inventories against atmospheric inversions. *Earth System Science Data* **14**, 1639–1675, https://doi.org/10.5194/essd-14-1639-2022 (2022).

Content courtesy of Springer Nature, terms of use apply. Rights reserved

50. Huo, D. *et al*. Near-real-time daily estimates of CO2 emissions from 1500 cities worldwide. *figshare* https://doi.org/10.6084/m9.figshare.19425665.v1 (2022).
51. Eggleston, S., Buendia, L., Miwa, K., Ngara, T. & Tanabe, K. 2006 IPCC guidelines for national greenhouse gas inventories. Report, IPCC (2006).

Acknowledgements

ZL acknowledge the National Natural Science Foundation of China (grant 71874097, 41921005, 72140002 and 72140002), Beijing Natural Science Foundation (JQ19032), and the Qiu Shi Science & Technologies Foundation. Authors acknowledges support from the CDP Worldwide.

Author contributions

D.H. and Z.L. designed the research. D.H., X.H. and X.D. conducted the data processing and wrote the manuscript. D.H., P.C. Z.D. and Z.L. designed the methods, and all authors contributed to data collection, discussion and analysis.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.H. or Z.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022

Content courtesy of Springer Nature, terms of use apply. Rights reserved

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com