Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

GEEA

Thematic set article: Analysis of Exploration Geochemical Data

Geochemistry: Exploration, Environment, Analysis

Published online June 1, 2017

https://doi.org/10.1144/geochem2016-012 | Vol. 17 | 2017 | pp. 204–216

Catchment-based gold prospectivity analysis combining geochemical, geophysical and geological data across northern Australia

M. J. Cracknell[1]* & P. de Caritat[2,3]

[1] CODES Centre of Excellence in Ore Deposits, University of Tasmania, Hobart TAS 7001, Australia
[2] Geoscience Australia, GPO Box 378, Canberra ACT 2601, Australia
[3] Research School of Earth Sciences, The Australian National University, Canberra ACT 2601, Australia
M.J.C., 0000-0001-9843-8251; P.d.C., 0000-0002-4185-9124
* Correspondence: m.j.cracknell@utas.edu.au

Abstract: The results of a pilot study into the application of an unsupervised clustering approach to the analysis of catchment-based National Geochemical Survey of Australia (NGSA) geochemical data combined with geophysical and geological data across northern Australia are documented. NGSA Mobile Metal Ion® (MMI) element concentrations and first and second order statistical summaries across catchments of geophysical data and geological data are integrated and analysed using Self-Organizing Maps (SOM). Input features that contribute significantly to the separation of catchment clusters are objectively identified and assessed.

A case study of the application of SOM for assessing the spatial relationships between Au mines and mineral occurrences in catchment clusters is presented. Catchments with high mean Au code-vector concentrations are found downstream of areas known to host Au mineralization. This knowledge is used to identify upstream catchments exhibiting geophysical and geological features that indicate likely Au mineralization. The approach documented here suggests that catchment-based geochemical data and summaries of geophysical and geological data can be combined to highlight areas that potentially host previously unrecognised Au mineralization.

Keywords: Regolith; Geochemistry; Geophysics; Geology; Self-Organizing Maps; National Geochemical Survey of Australia

Received 10 October 2016; **revised** 7 March 2017; **accepted** 8 March 2017

Regolith cover

Unweathered rock outcrop is rare across the Australian continent with more than 85% of its surface being covered by regolith (Wilford 2012). Although the majority of the continent is classified as having an arid to semi-arid climate, the wide range of regolith types found across Australia is a result of contrasting parent materials, long-term landscape evolution, diverse vegetation communities, and (palaeo-)climate extremes (Taylor & Butt 1998; Mann *et al*. 2012; Pain *et al*. 2012). The ubiquity of the regolith and its highly diverse characteristics present significant challenges to developing a clear understanding of the nature of these surface materials. In light of this challenge, a key theme of the UNCOVER initiative is to characterise regolith geochemical and geophysical properties (UNCOVER 2012). Information on regolith properties and formative (consistent with L1334) processes is crucial for developing new mineral exploration models in areas where prospective bedrock is concealed by surface material.

The National Geochemical Survey of Australia

Recently, Geoscience Australia with its State/NT partners embarked on a systematic continental-scale geochemical sampling program, the National Geochemical Survey of Australia (NGSA; Caritat & Cooper 2011*a*). The aim of the NGSA project was to provide ultra-low (spatial) density compositional data and information regarding the near-surface regolith to advance exploration for energy and mineral resources. The results of NGSA analyses are being used to improve our understanding of the concentration levels, spatial distribution, associations and their genesis and significance, transport processes, and sources and sinks of geochemical elements in the near-surface environment (e.g. see Caritat & Cooper 2016).

One of the key challenges to analysing the NGSA dataset is that it contains a large number of variables or features collected across nearly all the geological and biological regions of Australia. Recent research has explored the use of robust multivariate statistical analysis to identify and understand the distribution of geochemical elements analysed in the NGSA data (e.g. Caritat & Grunsky 2013; Scealy *et al*. 2015). Caritat & Cooper (2016) provide an up-to-date summary of studies using NGSA data for investigating geochemical processes for mineral exploration, agriculture and understanding contamination sources at continental and regional scales. Some of the studies that specifically relate to mineral exploration and employ methods for multivariate analysis are briefly summarized below as context to the present work.

Caritat *et al*. (2011) identified geochemical patterns in the Mobile Metal Ions® (MMI) analyses related to the spatial distribution of generalised lithological types by comparing geochemical patterns in the NGSA data to surface geology polygons. For example, samples taken from within the Great Artesian Basin and Murray-Darling Basin sedimentary provinces were found to exhibit elevated Ba, Ga and Sr concentrations. Conversely, elevated La, Ce and other rare earth elements (REEs) were found to be spatially coincident with areas where felsic intrusive rocks dominate, such as eastern Australia and SW Western Australia (WA). High-grade metamorphic terrains were spatially correlated with moderate concentrations of MMI Cs, K, Mo, Rb and W.

© 2017 Commonwealth of Australia (Geoscience Australia). This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/). Published by The Geological Society of London for GSL and AAG. Publishing disclaimer: www.geolsoc.org.uk/pub_ethics

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

Mann *et al.* (2013) analysed the relationship between mineral deposits and MMI element concentrations by identifying catchments with elevated commodity element concentrations and also containing mineral deposits. They found a reasonable spatial correlation with elevated Au concentrations and provinces known to host major Au mineralization. This observation was used to suggest that elevated Au concentrations in areas that do not contain known Au mineralization potentially form a useful exploration lead into areas such as the western Albany-Fraser belt in WA.

Caritat & Grunsky (2013) used principal component analysis to summarize and interpret NGSA geochemical data (total concentrations). They found that the first four principal components (PCs) accounted for 59% of the variance in the dataset. The element associations represented by these PCs relate to geological processes such as lithological controls, weathering, transport and secondary mineral precipitation. Based on these findings, Caritat & Grunsky (2013) identified lithological prediction (e.g. Grunsky *et al.* 2017) and mineral prospectivity analysis (e.g. this study) as potential uses of the NGSA dataset.

Spatial support

The interpolation of NGSA point data, which represent the overall sediment geochemical characteristics of catchments, into 2D maps via linear kriging of for instance the top ranked PCs generates continuously varying surfaces (rasters). The resulting surfaces do not conform to the discretised areal summary (watersheds or catchments) of geochemical characteristics that the NGSA data represent. While this may be a valid approach for defining broad continental-scale spatial trends in geochemical concentrations, it assumes a different spatial support, e.g. point locations, to that represented by catchment outlet sediments, which are transported sediments derived from upstream sources. Thus, NGSA data provide a representative indication of the overall geochemical characteristics of rocks and soils within a catchment (Caritat & Cooper 2011*a*) and potentially a catchment's entire upstream watershed.

In this study, we use an approach that does not assume the processes under consideration to be continuously varying in geographic space, that is, we generate spatial models using a multivariate statistical approach that not only preserves the catchment-based spatial support but also maintains the direct contribution of input features to quantifying the similarities and dissimilarities between catchments. Furthermore, we integrate geophysical and geological data, widely available across the Australian continent, with geochemical analyses to provide a deeper understanding of bedrock and regolith characteristics within NGSA catchments across much of northern Australia. This is to our knowledge only the second time such an approach has been applied to interrogating cross-disciplinary datasets with the aim of elucidating geological processes (see below).

Self-Organizing Maps

Self-Organizing Maps (SOM; Kohonen 1982, 2001) is an unsupervised clustering method useful for finding natural groups within complex multivariate data. SOM aids visualization and interpretation by reducing *n*-dimensional (*n*D) multivariate data to a two-dimensional 'map' where the spatial arrangement of neighbouring groups is representative of their similarities in *n*D space (Penn 2005; Bierlein *et al.* 2008). SOM uses vector quantization and measures of vector similarity, typically Euclidean distances, as a means of grouping input samples. The resultant groups or nodes are represented by a vector (code-vector) that summarises the properties of the associated input samples. Visualization of SOM component planes assists the interpretation of patterns and structures within the input data (Penn 2005; Bierlein *et al.* 2008; Löhr *et al.* 2010). For more detailed descriptions of SOM implementation and theory see Sun *et al.* (2009) and Cracknell *et al.* (2015).

Unlike other statistical clustering methods, such as factor analysis or k-means, SOM does not assume Gaussian distributions (Löhr *et al.* 2010; Žibret & Šajn, 2010). This is an important consideration for the analysis of geochemical data as these data are rarely normally or even log-normally distributed (Reimann & Filzmoser 2000). Previous research demonstrating the application of SOM for the analysis of geological and environmental patterns in geochemical data include Lacassie *et al.* (2004), Lacassie & Ruiz Del Solar (2006), Tsakovski *et al.* (2009), Sun *et al.* (2009), Löhr *et al.* (2010) and Žibret & Šajn (2010). In contrast to the research cited above that only analysed geochemical data, Cracknell *et al.* (2014) used SOM to combine interpolated soil geochemical and geophysical data. The resulting SOM clusters identified spatially consistent domains representing subtle geochemical contrasts related to changes in primary magmatic composition and hydrothermal alteration.

Study area

The study region covers *c.* 1.3 million km^2 across the Northern Territory and western Queensland in northern Australia (Fig. 1). This area was chosen primarily for the development of analysis methods documented in this pilot study as it is large enough to cover a substantial number of NGSA catchments but small enough to rapidly generate results. Furthermore, the study area contains a range of mineralization styles and includes geological materials from a wide variety of ages ranging from Proterozoic to Cenozoic (including Quaternary sediments; Fig. 2) and lithological types (Fig. 3). Finally, northern Australia is currently the subject of attention and investment (e.g. the 2016 – 2020 'Exploring for the Future' Programme of the Australian Government) and is thus both a topical and timely focus for the demonstration of the cross-disciplinary SOM approach developed here.

The oldest rocks in the study area are Palaeoproterozoic to Mesoproterozoic in age and are found in the Arunta, Tanami, Davenport, Tennant Creek, Isa, McArthur and Georgetown geological regions (see Fig. 2). Their dominant rock types are granulite facies metamorphosed felsic and mafic volcanics, fine grained clastic and carbonate metasediments, amphibolite facies turbidites and carbonaceous metasediments, folded greywackes and siltstones, clastic sedimentary rocks and metamorphosed volcanic rocks intruded by mafic and felsic (Blake *et al.* 1987; Ferenczi & Ahmad 1998; Wygralak & Bajwah 1998). The economically significant Isa and McArthur geological regions, as well as the Georgetown geological region, are hosts to a range of mineralization and deposit types, e.g. Au, base metals, Sn, W and Ta (Blake *et al.* 1987; Ahmad 1998; Ferenczi & Ahmad 1998; Wygralak & Bajwah 1998; Budd 2001; Withnall & Hutton 2013). The main Palaeozoic geological regions in the study area are the Georgina and Wiso geological regions (see Fig. 2). Their dominant rock types are clastic and carbonate sedimentary rocks, and regolith (Smith 1972; Kruse & Munson 2013). Known resources include phosphate and U, as well as groundwater, oil and gas (Smart *et al.* 1972; Radke 2009; Kruse & Munson 2013).

Aims

This study integrates catchment-based MMI geochemical data with geophysical imagery and geological information using SOM with the aim of objectively identifying groups of catchments with similar geochemical, geophysical and geological properties. Once identified, these catchment clusters are visually analysed in both data

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

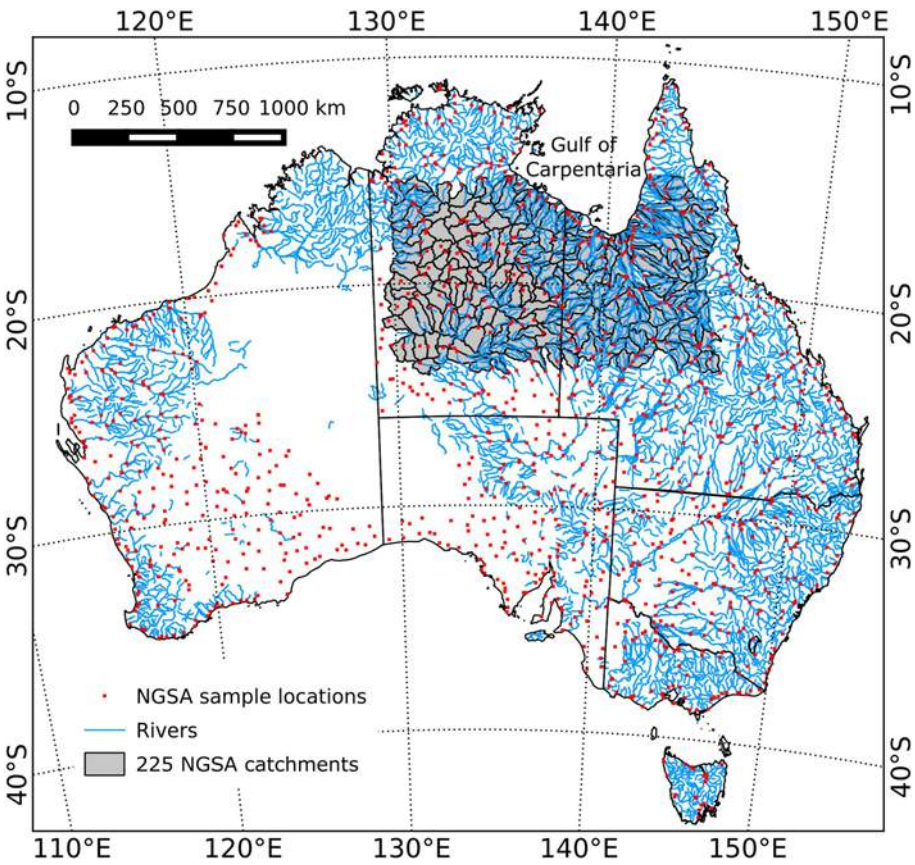

Fig. 1. Map of Australia showing selected National Geochemical Survey of Australia (NGSA; Caritat & Cooper 2011*a*) sample locations (red points) and the 225 catchments selected for this study (grey polygons) overlain with the Australian river network (blue polylines) (Geoscience Australia 2003). For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

space and geographic space. The integrated interrogation of catchment clusters – with respect to other geoscience data, including lithology, mineral deposits and mineral occurrences – are then used to formulate a Au mineral exploration model.

Materials and methods

All data used are publicly available from Geoscience Australia and were transformed to the Lambert Conformal Conical (Geoscience Australia) projection prior to analysis.

Geochemical data

Catchment-based geochemical data were sourced from the NGSA (Caritat & Cooper 2011*a*). A total of 225 NGSA catchments, *c.* 1/6 of the total number of NGSA catchments, were selected for the present SOM analysis. Analysis was performed on bulk properties (e.g. pH, electrical conductivity) and the MMI geochemical element assay data, the latter being determined on the coarse fraction (<2 mm) of top outlet sediment (TOS) samples (0 – 10 cm depth). A comprehensive quality assessment of the NGSA data describing

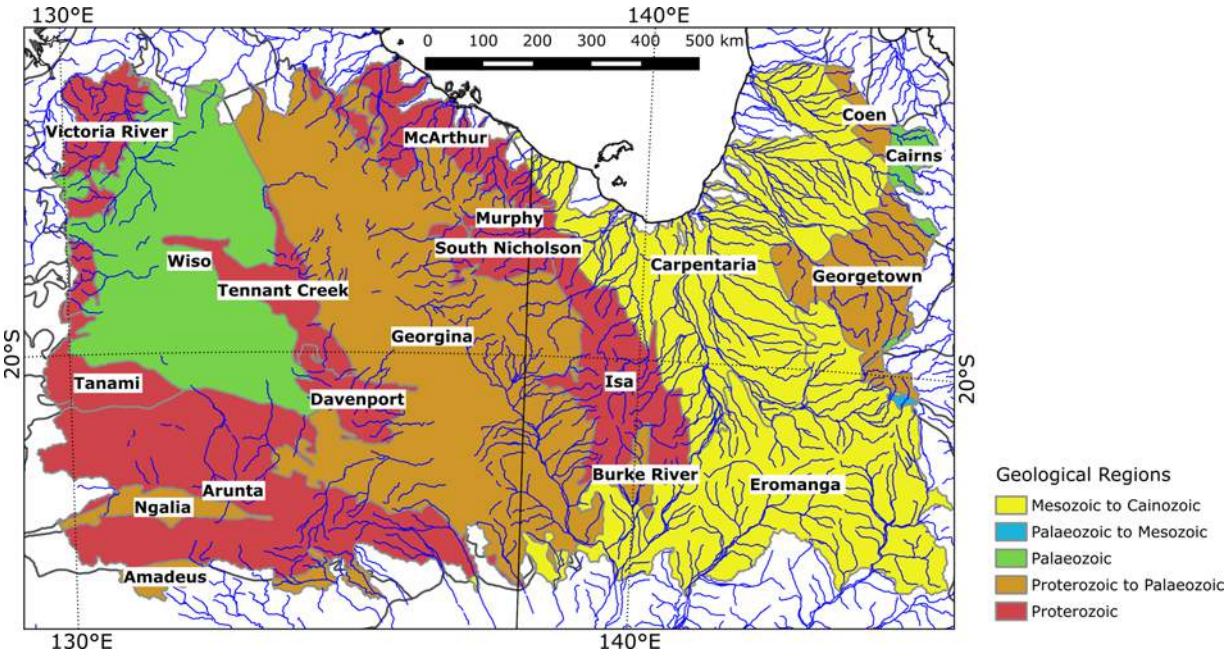

Fig. 2. Generalised geological regions (Blake & Kilgour 1998) across the study area, colour coded by age. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

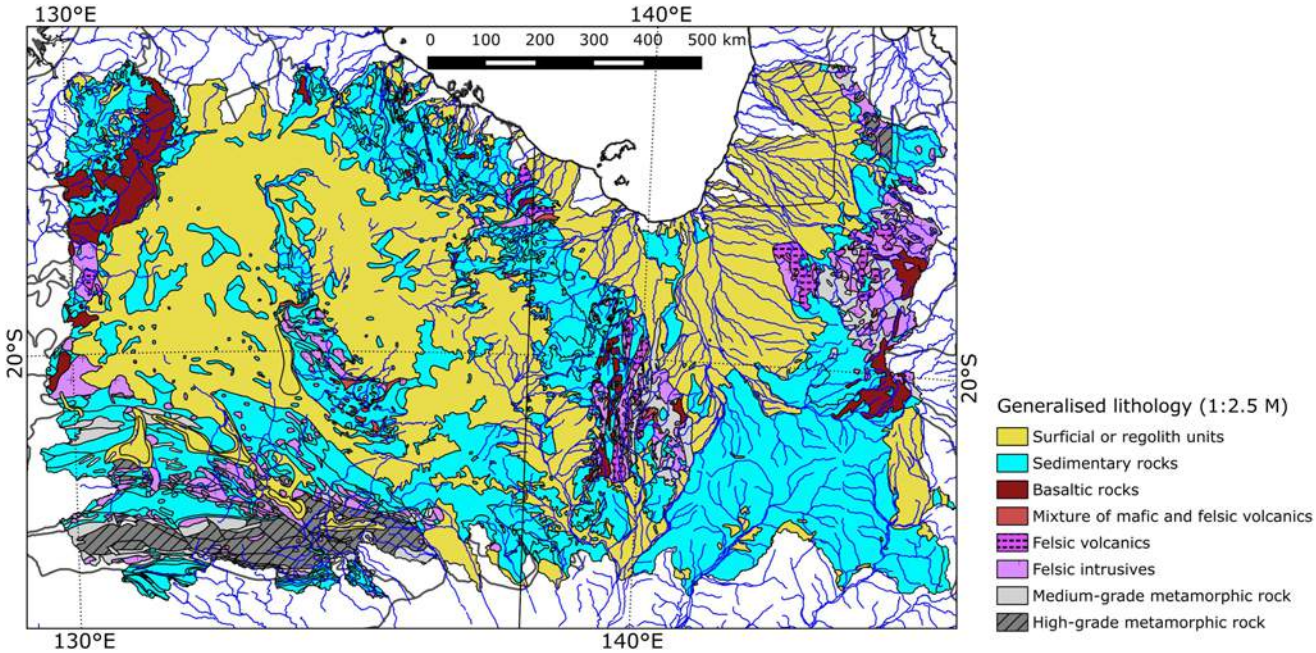

Fig. 3. Generalised lithologies (Raymond & Gallagher 2012) across the study area, colour coded by type. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

precision, bias, and censoring proportion is in the public domain (Caritat & Cooper 2011*b*). The scope of the present study was limited to the MMI data because this method extracts loosely adsorbed ions from the surfaces of minerals, organic matter and Fe-oxyhydroxides; thus, MMI results can be indicative of elements that have moved relatively recently through the regolith, which can reflect unusual element concentrations at depth potentially indicative of lithology or mineralization (Mann 2010). Consequently, these data are well suited to potentially identifying buried mineral deposits.

Bulk sediment properties data used included pH, electrical conductivity (EC), and percent fractions of clay, silt and sand. EC values were log transformed to approximate a normal distribution as these data are typically log-normally distributed (McKenzie *et al.* 2008). One catchment within the study area was excluded from analysis due to missing bulk properties data.

MMI element data that contained half or more samples with censored results (below the detection limit) were excluded from analysis. For the remaining 42 elements (Ag, Al, Au, Ba, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Ga, Gd, K, La, Li, Mg, Mn, Mo, Nd, Ni, P, Pb, Pr, Rb, Sc, Se, Sm, Sr, Tb, Th, Ti, U, V, Y, Yb, Zn and Zr), censored values were replaced by half the appropriate detection limit, a common practice in geochemistry (e.g. Botnick & White 1998; Helsel 2005; Antweiler & Taylor 2008; Carranza 2011). The data were then centred log-ratio (clr) transformed as described by Aitchison (1986) and in-line with other studies investigating the spatial variability of NGSA geochemical data (e.g. Caritat & Grunsky 2013; Mueller *et al.* 2014; Furman *et al.* 2016).

Geophysical data

The latest versions of total magnetic intensity (MAG; Percival 2014) with variable reduction-to-pole corrections applied (Version 6), filtered total count (dose) radiometrics Version 3 (TC; Minty *et al.* 2009) and spherical cap Bouguer gravity anomaly (GRAV; Tracey *et al.* 2007) raster data were clipped to the study area extent and resampled from their original resolutions to a 1000 m cell resolution using bilinear interpolation. Geophysical data resampling was carried out to avoid memory usage errors when processing the grey level co-occurrence matrix (GLCM; see below) textures and to enhance regional-scale geological features.

From the cells intersecting a given catchment, first order spatial statistics (mean and standard deviation) and second order spatial statistics (e.g. GLCM) for each geophysical input were calculated. The resulting first and second order spatial statistics were appended to the geochemical data acquired from each catchment. First order

Table 1. *Quantization and topological errors for 10 different combinations of X and Y dimensions with c. 200 node*

	X dim	Y dim	Nodes	Quantization error	Topological error
*	6	33	198	1.13	7.49
	8	25	200	1.14	7.81
	10	20	200	1.13	8.00
	12	17	204	1.13	8.20
	14	14	196	1.15	8.48
	16	12	192	1.14	8.28
	18	11	198	1.14	8.20
	20	10	200	1.14	7.96
	22	9	198	1.14	7.93
	24	8	192	1.14	8.00

Optimal SOM map dimensions (6 by 33) indicated by *.

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

spatial statistics were obtained by summarizing all cell values across a given catchment. Second order spatial statistics (texture) were obtained by assessing the spatial variability of all cell values at a particular scale (offset) within a given neighbourhood (Gonzalez & Woods 2008) using GLCM (Haralick *et al.* 1973). The mean GLCM contrast index averaged across the four principal directions (north–south, NE–SW, east–west and SE–NW) was used to represent spatial texture across a given catchment for 10 offsets with increments of *c.* 2 km (i.e. 2, 4, …, 20 km).

GLCM contrast is defined as (Baraldi & Parmiggiani 1995):

$$\mathrm{GLCM_{con}} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-j)^2 g(i,j)$$

where N_g is the number of grey levels in an image, i and j represent the ith and j^{th} grey levels respectively and $g(i, j)$ is defined as the $(i, j)^{\mathrm{th}}$ entry in the GLCM such that:

$$g(i,j) = \frac{p(i,j)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i,j)}$$

where $p(i,j)$ is the occurrence of unique pairwise combinations of grey levels i and j measured at two pixels separated by a given offset. GLCM contrast is correlated with spatial frequencies such that a high value of $\mathrm{GLCM_{con}}$ for a given offset and measured value indicates a large relative difference at that offset distance (Baraldi & Parmiggiani 1995).

Ancillary data

The OZMIN database (Ewers *et al.* 2002) of mines and mineral deposits and occurrences was used to assess the type and frequency of mineral occurrences within individual catchment clusters. Terrain slope was derived from GEODATA 9 second (*c.* 250 m resolution) digital elevation model (DEM) version 3 (Geoscience Australia 2008) using the slope function in QGIS version 2.12.1. The 1:2 500 000 scale Surface Geology of Australia (Raymond & Gallagher 2012) was used to summarize the dominant generalized lithological units intersecting a given catchment cluster.

SOM implementation

A total of 83 input variables representing geochemical, geophysical and geological properties were range normalised to 0–1 using a linear transformation. SOM was implemented using the R statistical programming language package *som* (Yan 2010), which is based on SOM-PAK (Kohonen *et al.* 1996). Multiple trials of different X and Y SOM map dimensions for *c.* 200 randomly seeded nodes with hexagonal topologies were initiated and run for over 10 000 iterations with a Gaussian neighbourhood function and inverse learning function (Kohonen *et al.* 1996). Optimal SOM map dimensions were identified by minimizing quantization and topological errors. For more information on the theory and derivation of SOM quantization and topological errors used here see Cracknell *et al.* (2015).

Cluster selection and properties

Once an optimal SOM model was selected for *c.* 200 nodes, a hierarchical dendrogram agglomerative clustering method was employed to merge SOM nodes based on their code-vectors (Vesanto & Alhoniemi 2000; Cracknell *et al.* 2015). The Davies-Bouldin Index (DBI; Davies & Bouldin 1979), which estimates cluster similarity using the maximum mean ratio of cluster dispersion and pairwise centroid distances, was used to identify an optimal number of clusters (i.e. merged SOM nodes).

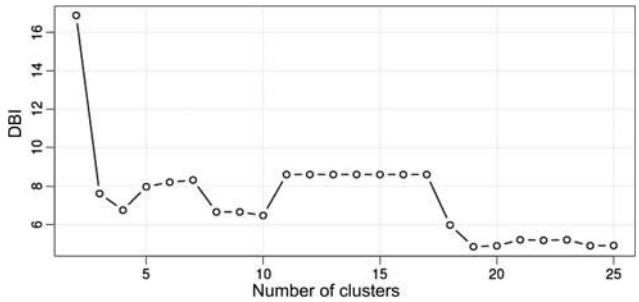

Fig. 4. Davies-Bouldin Index (DBI: Davies & Bouldin 1979) as a function of the number of merged SOM nodes for the 6 by 33 SOM model.

Input variables that contributed significantly high or low SOM code-vector values, with respect to other cluster code-vector values, were identified using the following formula, modified from Siponen *et al.* (2001):

$$r_i = \frac{s(i,k)}{\frac{1}{c-1}\sum_{j \neq i} s(j,k)} - 1$$

where $s(i,k)$ is the mean cluster i code-vector for input variable k and $s(j,k)$ is the mean cluster j code-vector for input variable k. This ratio provides an indication of the relative difference of k in cluster i as compared to the mean code-vector values of all other clusters j (Siponen *et al.* 2001). For example, values >>0 indicate clusters with substantially higher mean code-vector values compared to the mean values of all other clusters. Conversely, values <<0 indicate clusters with substantially lower mean code-vector values compared to the mean values of all other clusters. Variables contributing significantly higher (or lower) values to a given cluster are identified as those with code-vector ratios greater (or lower) than one standard deviation from the mean code-vector ratio.

Results

A 6 by 33 (X by Y) SOM map with 198 nodes was found to result in the minimum mean quantization and topological errors (Table 1). The DBI as a function of merged SOM nodes (up to 25 merged nodes) identified 19 as the optimal minimum number of clusters (Fig. 4), although other local minimum DBI values were observed for 4 and 8 clusters. A plot of the spatial distribution of resulting catchment clusters is shown in Figure 5.

The relative positions of catchment clusters on the Au component plane plot are presented in Figure 6a. The catchment clusters with significantly high mean Au code-vector ratios (warm colours) plot together near the top of the SOM map (refer to the online version of this article). Figure 6b plots cluster mean code-vector ratios of Au concentration as compared to all other clusters. Clusters 15, 14, 17, 16 and 12 are identified as exhibiting mean Au code-vector ratios greater than one standard deviation above the mean Au code-vector ratio.

Figure 7a maps the locations of clusters identified in Figure 6b that display mean Au code-vector ratios greater than one standard deviation above the mean (bold outlines), as well as the clr transformed values of MMI Au concentrations (colour scale), overlain with Au mines and mineral occurrences (red stars). Figure 7b plots catchment clusters with high Au mean code-vector overlain with terrain slope as greyscale pixels. At first glance there does not appear to be any spatial relationship between Au mines and high catchment Au concentrations or catchment clusters with high mean Au code-vector ratios, however, catchments with high mean Au code-vector ratios are positioned at the transition from relatively high slope to low slope (i.e. break in slope) downstream from Au mine and mineral occurrence locations. Visual comparison of catchment clusters 15, 14, 17, 16 or 12 indicates that 41 out of these

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

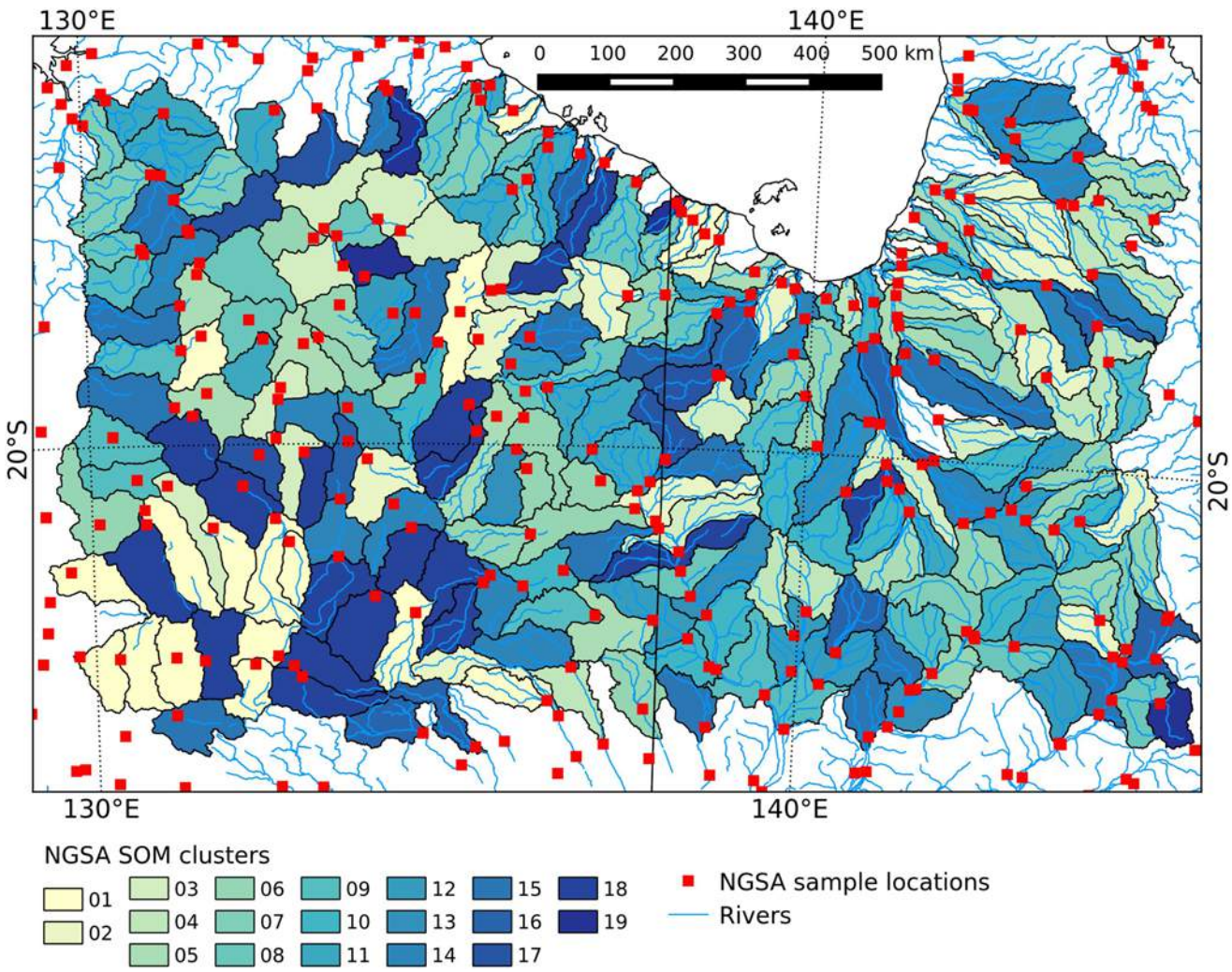

Fig. 5. Spatial distribution of the 19 catchment clusters for SOM map dimensions 6 by 33 and 198 nodes. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

54 catchments (76%) have a portion of their extent intersecting a break in slope.

Figure 8 shows catchments with high mean Au code-vectors (clusters 15, 14, 17, 16 or 12) and upstream catchment clusters 4, 6 and 9 identified by visually interrogating mapped relationships. Catchment clusters 4, 6 and 9 were selected by manually querying regions upstream of catchment clusters 15, 14, 17, 16 and 12 and taking note of regularly occurring cluster indices. These upstream

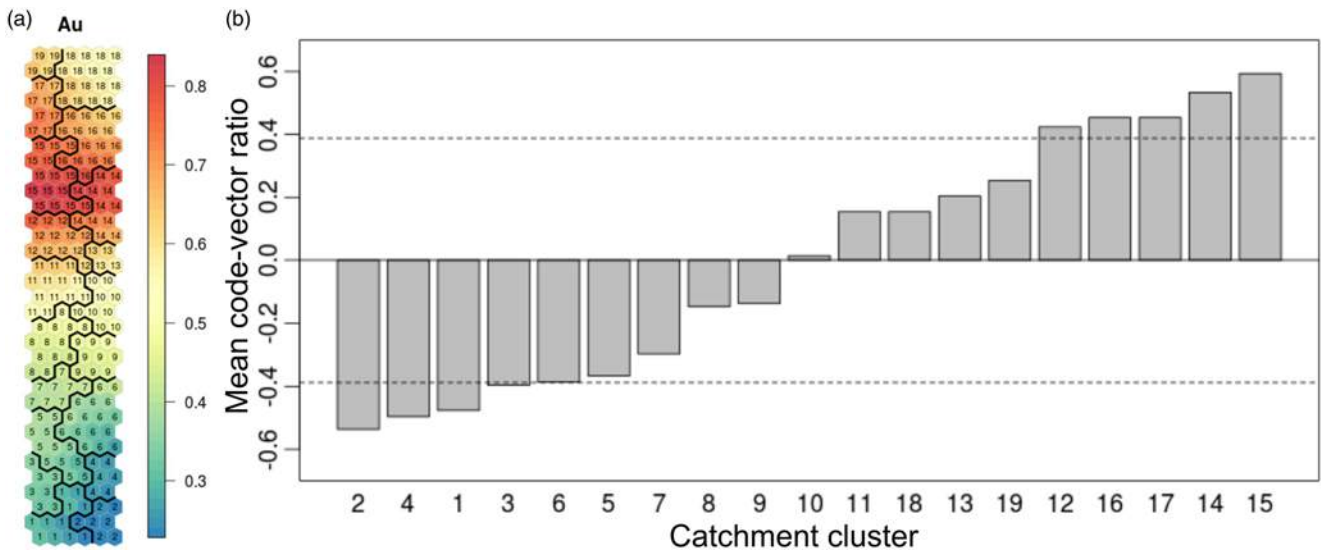

Fig. 6. (**a**) 2D SOM map and Au component plane with division of final catchment clusters and (**b**) mean code-vector ratios for cluster Au concentration. Five clusters display higher that one standard deviation from the mean (dashed line) code-vector ratio. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

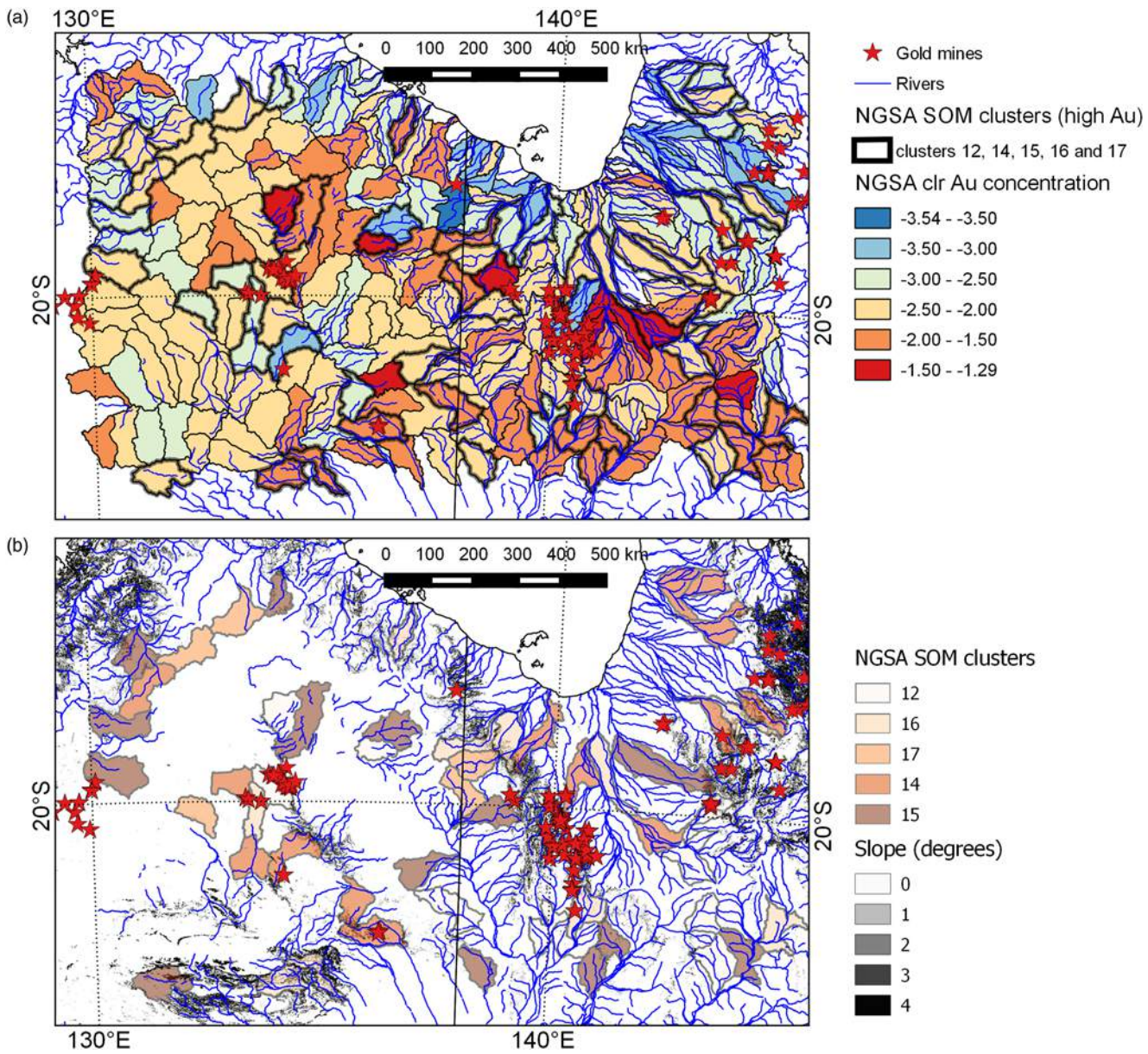

Fig. 7. (**a**) Spatial distribution of clusters with mean Au concentration code-vector ratios greater than one standard deviation from the mean, and Au mines and mineral occurrences locations overlain on colour coded NGSA catchment centred log-ratio transformed Au concentration. (**b**) NGSA catchment clusters with high Au mean code-vectors overlain on terrain slope. Note that clusters with high Au code-vector ratios are found immediately downstream of most Au mines at the break in slope. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

catchment clusters are found to have a close spatial association with catchments hosting Au mines and mineral occurrences and plot as immediate neighbours to each other on the SOM 2D map (Fig. 6a). By combining stream network information the proportion of catchments containing Au mines that are directly upstream of one or more NGSA catchment clusters 15, 14, 17, 16 or 12 was calculated. Of the 31 NGSA catchments in the study area that contain Au mineralization, 28 have flow paths that are not internally draining, i.e. we have omitted three catchments (located in the southern region of the Northern Territory) with confused flow paths. Of these 28 catchments (13% of the total number of catchments), 21 are linked downstream to NGSA catchment clusters with high mean Au code-vector (24% of the total number of catchments). This indicates that 75% of the catchments with Au mineralization are upstream of NGSA catchment clusters 15, 14, 17, 16 or 12. If catchment clusters 4, 6 and 9 (a further 18% of the total number of catchments) are included, 27 (96%) of the 28 Au mineralised catchments are upstream of catchment clusters identified in this study. In contrast, of the 63 catchments (28% of the total number of catchments) within the study area that display high Au concentrations (i.e. Au clr values between −2.00 to −1.50 and −1.50 to −1.29, see Fig. 7a) only 10 (16%) are located downstream of catchments with known Au mines and mineral occurrences.

Table 2 summarizes the significantly high and low code-vector values for all catchment clusters with the frequency of mines and mineral occurrences for a given (dominant) commodity that intersects these clusters. The upstream catchment clusters 4 and 6 are characterized by low concentrations in fine clastic components (i.e. clay and silt). Cluster 4 displays high contrast in magnetics for distances less than 10 km and low contrast in magnetics for greater than 10 km. Cluster 9 exhibits a high contrast in gravity for wavelengths less than 10 km and low contrast for wavelengths of 12 – 14 km. All upstream catchment clusters (4, 6 and 9) contain a high frequency of Au mines and mineral occurrences with cluster 4 also containing many Ag mines, cluster 6 Cu mines and cluster 9 Cu and U mines.

Table 3 ranks generalized lithological units within clusters 4, 6 and 9 based on differences in the proportion of area for a given unit

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

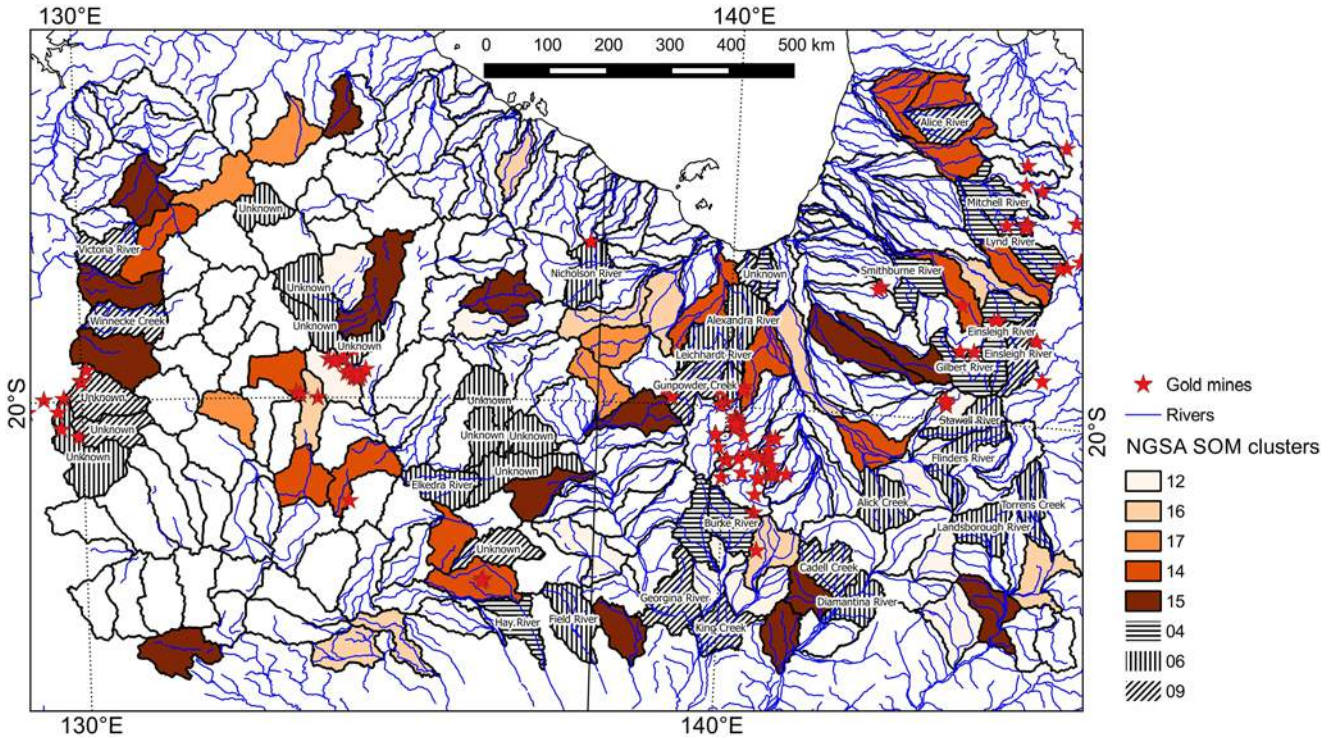

Fig. 8. Spatial distribution of catchments with high Au concentration (colour intensity indicates Au concentration rank, see Fig. 6) and neighbouring upstream catchment clusters 4, 6 and 9. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

compared to their overall proportion across the entire study area. Hence, positive values highlight lithological units that cover a larger proportion of the cluster area with respect to the mean of all catchments. Clusters 4 and 9 contain large proportions of felsic intrusive rocks and low proportions of surficial or regolith units. Clusters 4 and 6 contain high proportions of medium-graded metamorphic rocks, while clusters 6 and 9 show high proportions of sedimentary rocks and low proportions of high-grade metamorphic rocks.

Discussion

Present work

NGSA samples were collected as catchment outlet (overbank or floodplain) sediments. Overbank sediments have been shown to be more representative of the geochemical composition of the catchment than stream sediments (Ottesen *et al.* 1989). This is because the suspended sediment load in a flood event, from which the overbank or floodplain sediments are primarily derived, is sourced from a greater area than the sediments within the stream channel, which are typically derived from local point sources. Thus, catchment outlet sediments are assumed to represent an integrated sample of the entire catchment area (Ottesen *et al.* 1989; Bølviken *et al.* 2004). Furthermore, outlet sediments are ubiquitous across a diverse range of geomorphological and climatological regions. The results presented in this study indicate that the geochemical characteristics of outlet sediments sampled from large river systems are likely to be representative of both the immediate catchment watershed and the upstream drainage basin from which these sediments are potentially derived.

The MMI extraction, however, was developed to mainly mobilise the labile fraction of chemical elements, presumably from the outer surfaces of soil particles (Mann 2010). Accordingly the MMI response can be subdued after significant rain and flooding, but can also reform relatively quickly (Mann *et al.* 2005). Thus the system investigated geochemically here is a fairly dynamic one, especially in the region of interest where rainfall is seasonal (typical dry and wet seasons in winter and summer respectively). The reason why the MMI geochemical characteristics of outlet sediments sampled from large river systems are likely to be representative of both the immediate catchment watershed and the upstream drainage basin is through a combination of mechanical transport of sediment grains and hydromorphic dispersion of geochemical signatures through groundwater flow systems. Whilst the sediment matrix is physically inherited from both the catchment where the outlet sediment is sampled and potentially that upstream, the surface adsorbed, labile chemical (MMI) signature may form as groundwater rises to the surface at topographic breaks in slope. If groundwater is in direct or indirect contact with mineralised basement in the upstream (part of a) catchment it can acquire and transport downstream a geochemical signature diagnostic of this (e.g. Leybourne & Cameron 2010). In the case of Au, the MMI response in sediments likely arises from a combination of the etching of clastic gold grains (placer pathway) and the extraction of labile, adsorbed fine secondary Au on particle surfaces (hydromorphic pathway) (A. Mann, pers. comm. 2017).

The information in Table 2 summarizes catchment cluster characteristics that contribute significantly to their dissimilarities (or similarities) to other clusters. This information provides a tentative indication of the lithological origins of the outlet sediments analysed. For example, clusters 1, 2, and 4 have high Ce, La and REEs suggesting felsic igneous dominant sources (Caritat *et al.* 2011), while also exhibiting a high total count first order mean. Clusters 1, 3 and 4 display high sand content, high Th and Zr. Clusters 1 to 4 display significantly low pH, EC, Au, Cu, Ca, Ba, Co, Mg, Ni and Sr. Many of these elements are typically associated with mafic igneous lithological sources. These observations suggest felsic igneous sources for clusters 1–4. Clusters 1 and 3 display low contrast in gravity at wavelengths of less than or equal to 10 km and clusters 2 and 4 display low contrast in magnetism at wavelengths greater than 10 km. The geophysical characteristics of these clusters potentially provide an indication of the maximum 'size' of felsic

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

Table 2. *Summary of upstream catchment cluster high and low code-vector values and the frequency of intersecting mines and mineral occurrences based on the OZMIN database (Ewers et al. 2002) for a given mineral commodity*

	High code to vector ratio	Low code to vector ratio	Commodity (frequency)
Cluster 1	Sand, Ce, Cs, Eu, Fe, Gd, La, Nd, Pr, Rb, Sc, Sm, Tb, Th, Ti, Zr TC-mean, TC-14con to TC-20con, GRAV-std, GRAV-12con to GRAV-20con	pH, EC, Silt, Clay, Ag, Au, Ba, Ca, Cd, Co, Cu, Ga, K, Li, Mg, Mo, Ni, Sr MAG-mean, TC-8con to TC-10con, GRAV-2con to GRAV-10con	U_3O_8 (4), Fe (1)
Cluster 2	Ce, Cr, Cs, Dy, Er, Eu, Fe, Gd, La, Nd, Pb, Pr, Sc, Sm, Tb, Th, Ti, Y, Yb, Zr MAG-2con to MAG-8con, TC-mean, TC-std	pH, EC, Clay, Al, Au, Ba, Ca, Cd, Co, Cu, Ga, K, Li, Mg, Mn, Mo, Ni, Sr, U, Zn MAG-mean, MAG-std, MAG-12con to MAG-18con	Au (1), U_3O_8 (1)
Cluster 3	Sand, Eu, La, Nd, P, Pr, Rb, Sc, Th, Zr GRAV-std, GRAV-12con to GRAV-18con	pH, EC, Silt, Clay, Ag, Au, Ba, Ca, Co, Cu, Ga, Mg, Ni, Pb, Sr MAG-mean, MAG-8con to MAG-10con, TC-2con to TC-10con, GRAV-6con to GRAV-10con	P_2O_5 (3), Mo (1)
Cluster 4	Sand, Ce, Cr, Cs, Dy, Eu, Fe, Gd, La, Nd, Pr, Sc, Sm, Tb, Th, Ti, Zr MAG-2con to MAG-10con, TC-mean, TC-std	pH, EC, Silt, Clay, Al, Au, Ba, Ca, Cd, Co, Cu, Ga, K, Mg, Mn, Ni, Sr, V MAG-12con to MAG-20con	Ag (5), Au (4) Co (1), Cu (1), P_2O_5 (1)
Cluster 5	Sand, P, Rb	Silt, Clay, Co, Pb	Au (2), Ag (1)
Cluster 6	Nd, Zn MAG-10con	Clay, Co, Cs, Pb, Se, V TC-mean	Au (5), Cu (3), Mn (1), P_2O_5 (1), U_3O_8 (1)
Cluster 7	TC-2con to TC-6con	Cr, Cs, Pb MAG-std, TC-18con to TC-20con	Au (2), Cu (1)
Cluster 8	MnMAG-12con to MAG-20con, TC-2con to TC-10con	Cr, Cs, Rb, U MAG-std, MAG-2con to MAG-10con, TC-12con to TC-16con	P_2O_5 (4), Ag (1), Fe (1), Pb (1)
Cluster 9	Dy, Er, Gd, Tb, Y, Yb TC-4con to TC-10con, GRAV-2con to GRAV-10con	Cr, Cs, Mo, Rb, Se to Ti, V GRAV-12con to GRAV-14con	Au (4), Cu (7), Fe (1), P_2O_5 (1), U_3O_8 (4)
Cluster 10	MAG-mean, MAG-std, MAG-8con to MAG-10con, GRAV-mean, GRAV-2con to GRAV-10con	Fe, Rb, Zn GRAV-12con to GRAV-20con	Cu (20), U_3O_8 (7), Au (12), Co (3), Fe (3), Mn (3), Ag (2), P_2O_5 (2), Pb (1)
Cluster 11	MAG-12con to MAG-20con, TC-2con to TC-10con	Rb, UMAG-2con to MAG-4con, TC-std, TC-12con, GRAV-std	Ag (1), Dmd (1)
Cluster 12	Silt, Clay, Ag, Au, Co, Ga, Pb TC-2con to TC-10con, GRAV-10con	Sand, La, P, Pr, Zr TC-mean, TC-std, TC-12con to TC-16con	Au (10), Cu (1), Mn (1), Y_2O_3 (1)
Cluster 13	MAG-mean, MAG-std, MAG-4con to MAG-10con, GRAV-2con to GRAV-10con	Fe, P, Ti MAG-12con to MAG-16con, GRAV-12con to GRAV-20con	Cu (6), Au (4), REO (1)
Cluster 14	pH, Silt, Clay, Ag, Al, Au, Cd, Co, Cu, Li, Mo, Pb, Se, Sr, U, V, Yb MAG-2con to MAG-10con	Sand, Ce, Eu, Fe, La, Nd, P, Pr, Sm, Th, Ti, ZrMAG-12con to MAG-20con	Au (3), Ag (2),Cu (2), P_2O_5 (2), Mo (1), U_3O_8 (1), WO_3 (1), Zn (1)
Cluster 15	pH, EC, Silt, Clay, Ag, Au, Ba, Cd, Co, Cu, Ga, Li, Mg, Mo, Pb, Sr, V	Sand, Ce, Dy, Eu, Gd, La, Nd, P, Pr, Sm, Tb, Th, Y, ZrTC-mean, TC-std	P_2O_5 (2), Cu (1)
Cluster 16	pH, EC, Clay, Au, Ca, Cd, Cu, K, Li, Mg, Mo, Rb, Se, Sr, U, V, Zn	Ce, Dy, Eu, Fe, Gd, La, Nd, Pr, Sm, Tb, Th, YMAG-20con	P_2O_5 (5), Ag (3), Au (2), Pb (1), U_3O_8 (1), Zn (1)
Cluster 17	pH, Clay, Ag, Al, Au, Ba, Ca, Cd, Cu, Ga, K, Li, Mg, Mn, Mo, Ni, Sr, U, V, Zn MAG-12con to MAG-20con, TC-16con to TC-18con to TC-20con, GRAV-20con	Ce, Dy, Er, Eu, Gd, La, Nd, Pr, Sm, Tb, Th, Y, YbMAG-2con to MAG-10con, TC-mean, TC-2con, GRAV-mean	–
Cluster 18	EC, K, U TC-mean, TC-std, TC-12con to TC-20con, GRAV-std, GRAV-12con to GRAV-20con	TC-2con to TC-10con, GRAV-mean, GRAV-2con to GRAV-10con	Au (8), U_3O_8 (3), Cu (2), Fe (1), Ni (1), P_2O_5 (1)
Cluster 19	Al, K, Mn, Mo, P, Zn MAG-12con to MAG-20con, TC-12con to TC-20con, GRAV-12con to GRAV-20con	Dy, Er, Eu, Gd, Tb, Y, Yb MAG-2con to MAG-10con, TC-2con to TC-8con, GRAV-mean, GRAV-2con to GRAV-8con	–

Abbreviations for geophysical code-vectors: MAG, total magnetic intensity; GRAV, spherical cap Bouguer anomaly; TC, total count radiometrics; mean, first order mean across catchment; std, first order standard deviation across catchment; -Xcon, second order GLCM contrast at a given offset (X) in km. Abbreviations for commodities: Ag, silver; Au, gold; Co, cobalt; Cu, copper; Dmd, diamond; Fe, iron; Mn, manganese; Mo, molybdenum; Ni, nickel; Pb, lead; P_2O_5, phosphate; REO, rare earth oxides; U_3O_8, uranium; WO_3, tungsten; Y_2O_3, yttrium; Zn, zinc.

igneous features, e.g. low gravity contrasts within plutons and high magnetic 'alteration' in contact zones. Furthermore, the majority of clusters 1–4 either intersect Proterozoic geological regions such as the Arunta, Isa or Georgetown regions, or are immediately downstream of one, e.g. north and NE of the Tennant Creek, and east of the South Nicholson geological regions. These clusters also

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

Table 3. *Ranked order of difference in average area proportion of generalised lithological groups with respect to the proportion covering catchment clusters 4, 6, and 9*

	Rank	Generalised lithologies	Difference to average
Cluster 4	1	Felsic intrusive rocks	0.112
	2	Medium-grade metamorphic rocks	0.088
	3	Felsic volcanic rocks	0.042
	4	Dolerite	0.011
	5	Mixture of mafic and felsic volcanic rocks	0.000
	6	Basaltic rocks	−0.005
	7	High-grade metamorphic rocks	−0.026
	8	Sedimentary rocks	−0.044
	9	Surficial or regolith units	−0.176
Cluster 6	1	Sedimentary rocks	0.082
	2	Mixture of mafic and felsic volcanic rocks	0.005
	3	Medium-grade metamorphic rocks	0.002
	4	Dolerite	−0.002
	5	Surficial or regolith units	−0.004
	6	Felsic volcanic rocks	−0.010
	7	Felsic intrusive rocks	−0.019
	8	Basaltic rocks	−0.023
	9	High-grade metamorphic rocks	−0.030
Cluster 9	1	Felsic intrusive rocks	0.087
	2	Sedimentary rocks	0.080
	3	Basaltic rocks	0.066
	4	Volcanoclastic sedimentary rocks	0.004
	5	Felsic volcanic rocks	0.003
	6	Mixture of mafic and felsic volcanic rocks	−0.003
	7	Medium-grade metamorphic rocks	−0.015
	8	High-grade metamorphic rocks	−0.028
	9	Surficial or regolith units	−0.191

occur together in the lower half of the SOM map (Fig. 6a) with clusters 1 and 3 on the left hand side and clusters 2 and 4 on the right.

The high-grade metamorphic terrain of the Arunta geological region in the SE of the study area is predominantly intersected by clusters 1 and 18 (Fig. 9). These two clusters are at opposite ends of the SOM map in Figure 6a and appear to be linked based on their low contrast in gravity for wavelengths less than 10 km, high contrast in total count radiometrics for wavelengths greater than 10 km and high total count mean (Table 2). Some of these geophysical characteristics are coincident with those identified for catchment clusters 4, 6 and 9, i.e. high contrast in gravity at wavelengths less than 10 km and high variability in total count radiometrics. This suggests that a high proportion of metamorphic rocks within catchment clusters corresponds to unique geophysical signals.

Clusters 15 and 17 display high Ba, Ga and Sr suggesting regions with source rocks dominated by sedimentary basins (Caritat *et al.* 2011), while also displaying high pH and silt and clay materials further supporting this observation. However, these two clusters exhibit high Ag, Cd, Cu, Li, Mg, and V and low concentrations of REEs, Ce and Th. Given that these clusters have been identified as potential Au mineralized catchments they share similarities with other clusters displaying high Au concentration: high values for pH, clay-silt material and EC; low sand concentration; and highly variable magnetics and gravity, i.e. a mixture of low contrast at long and short wavelengths. These cluster similarities suggest that the bulk of the separation is based on the geochemical data for these clusters.

Future work

In future studies the implementation of additional processing of the geochemical, geophysical and geological data prior to input into SOM and more sophisticated analysis of stream networks will greatly improve the interpretability of catchment cluster characteristics and aid positive mineral exploration outcomes. In the data pre-processing phase we suggest imputing censored values, i.e. those below detection limits, based on the methods described in Caritat & Grunsky (2013) or similar. This will provide additional geochemical features to analyse. We then believe that the removal of regional trends in geophysical signals, effectively calculating residual fields,

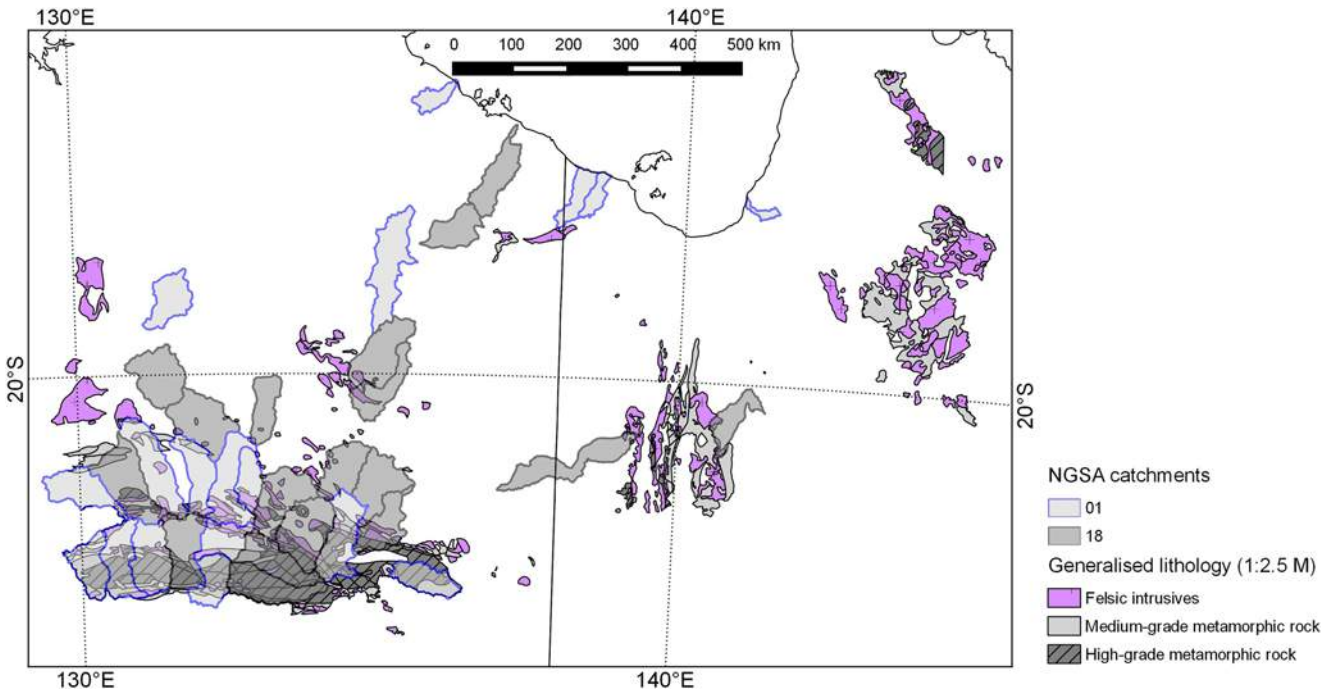

Fig. 9. Map of the spatial distribution of felsic intrusive rocks and medium- and high-grade metamorphic rocks. Map overlain with catchment clusters 1 and 18. For interpretation of the references to colour in this figure, the reader is referred to the online version of this article.

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

will reduce the potential for unreasonable textural outputs. However, a change in the input features will likely lead to variations in the optimal number of clusters (Dy & Brodley 2004). Moreover, an increase in the number of input features in an already high-dimensional data space will exacerbate the effect of the curse of dimensionality (Bellman 1961).

The curse of dimensionality describes the increase in the distance between samples as the number of features increases. Hence, a focus of further research should investigate dimensionality reduction approaches such as the removal of correlated features and unsupervised feature selection, which identifies features that contribute most to the separation of clusters (e.g. Dash *et al.* 2002*a*,*b*; Dy & Brodley 2004; Alelyani *et al.* 2013). For example, a simple filter search method proposed by Dash *et al.* (2002*a*) uses an information entropy metric based on sample distances in data space to identify the overall disorder of a system given the iterative exclusion of individual input features. This metric is then used to rank the input features that contribute most to the separation of samples. Subsequently a wrapper method (Alelyani *et al.* 2013) can be used to iteratively obtain cluster separation metrics, e.g. based on the DBI, for a number of different SOM map dimensions which are used to identify the optimal number of ranked features for a given dataset. Alternatively, weights of evidence (WofE) analysis could be used to determine the relative importance of input features based on mineral occurrence information (Bonham-Carter *et al.* 1989; Carranza 2009) and may offer a means of identifying relevant input features.

Due to the spatially restricted nature of stream networks, the analysis of a subset of catchments located within a regional drainage basin will aid the development and interpretation of models of geochemical, geophysical and geological catchment characteristics. The analysis of catchments located within regional drainage basins will simplify the construction stream network geometry models. Carranza (2010*a*,*b*) demonstrated the crucial influence that stream network geometries have on the analysis of stream sediment geochemical samples. The resulting regional drainage basin models of catchment clusters across Australia may then be compared in order to identify similarities and dissimilarities between catchments.

A further refinement in future work would be to incorporate interpretations of the potential influence of climate, vegetation and topography on geochemical data across large areas. A number of national datasets exist that capture such information and could be integrated into the next generation of SOM based prospectivity analysis research.

Conclusions

The search for undiscovered mineral deposits across Australia is shifting to regolith dominated terrains. As a result, today's mineral explorers require knowledge of regolith sources and formative processes in order to develop appropriate prospectivity models. The ever increasing volume and variety of digital geoscience data available in the public domain, such as catchment-based geochemical analyses collected for the National Geochemical Survey of Australia (NGSA), provide an opportunity to formulate new prospectivity models where bedrock is covered. However, the challenge is to integrate these multivariate data in a meaningful and interpretable way.

Unsupervised clustering methods, such as Self-Organising Maps (SOM), provide an opportunity to identify and visualise patterns in diverse multivariate data that are not apparent in a low-dimensional data space. In this study, SOM was used to integrate NGSA geochemical data with first and second order summaries of geophysical data and geological information across regional-scale catchments. Groups of catchment clusters identified from the analyses of SOM code-vectors can be linked to regional lithological trends and Au mineralization potential, however, these catchment clusters must be interpreted with consideration of the contributing upstream area both through mechanical transport of sediment grains and via hydromorphic dispersion of chemical elements that can bind to particle surfaces where groundwater intersects the land surface. This finding is demonstrated by analysing and visualizing catchment clusters that exhibit substantially high mean Au MMI code-vector values.

The identification of a high percentage of catchments with high mean Au code-vector ratios being located downstream from Au mines and mineral occurrences is a significant result as it suggests Au is being liberated from areas of Au mineralization and transported downstream potentially both mechanically and hydromorphically. Au is subsequently detected by MMI extraction in sediments at or below the break in slope where hydrological energy decreases (mechanical transport pathway) and aquifers potentially intersect the land surface (hydromorphic transport pathway). This information has been used to define catchments upstream of those with high mean Au MMI code-vector ratios in outlet sediments as potential hosts of Au mineralization. Three upstream catchment clusters that have a close spatial relationship with high mean Au code-vector ratios clusters are identified. These three clusters intersect a high frequency of Au mines in areas that contain a mixture of felsic intrusive, sedimentary and medium-grade metamorphic rocks. The geophysical characteristics of these prospective catchment clusters, and others with high proportions of metamorphic rocks, indicate high contrast in either magnetic and gravity signals at wavelengths of less than 10 km, or elevated and highly variable total count radiometrics signals.

Further investigation into the role that geochemical element mobility plays in governing the relative contribution of catchment characteristics is required. In addition, understanding the significance of the spatial frequency characteristics of geophysical data is needed to clarify the relationship between code-vector ratios and mineralization.

Acknowledgements We acknowledge the R project for statistical computing (http://www.r-project.org). The NGSA was lead and managed by Geoscience Australia and carried out in collaboration with the geological surveys of every State and the Northern Territory under National Geoscience Agreements. The authors acknowledge and thank all landowners for granting access to the sampling sites and all those who took part in sample collection. The sample preparation and analysis team at Geoscience Australia is thanked for its contributions, as is analytical staff at SGS Perth laboratories. We thank SGS Perth laboratories for providing the MMI analyses and Alan Mann for his advice on MMI geochemistry interpretation. We are grateful to Geoscience Australia internal reviewers Chris Lewis, Evgeniy Bastrakov, and Karol Czarnota for clarifying and improving the draft manuscript. John Carranza and one anonymous reviewer's comments contributed to the clarity and scientific rigour of this manuscript. P. de Caritat publishes with permission from the Chief Executive Officer, Geoscience Australia.

Funding The NGSA project was part of the Australian Government's Onshore Energy Security Program 2006 – 2011, from which funding support is gratefully acknowledged. This research has used data that are © Commonwealth of Australia and have been used with the permission of Geoscience Australia.

References

Ahmad, M. 1998. Geology and mineral deposits of the Pine Creek Inlier and McArthur Basin, Northern Territory. *AGSO Journal of Australian Geology and Geophysics*, **17**, 1–17.

Aitchison, J. 1986. *The statistical analysis of compositional data*. Chapman & Hall, Ltd, London, UK.

Alelyani, S., Tang, J. & Liu, H. 2013. Feature Selection for Clustering: A Review. *In:* Charu, C.A. & Chandan, K.R. (eds) *Data Clustering: Algorithms and Applications, Data Mining and Knowledge Discovery Series*. CRC Press, Boca Raton, Florida, 30–55.

Antweiler, R.C. & Taylor, H.E. 2008. Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets: I. Summary statistics. *Environmental Science and Technology*, **42**, 3732–3738.

Baraldi, A. & Parmiggiani, F. 1995. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

parameters. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 293–304, https://doi.org/10.1109/36.377929

Bellman, R.E. 1961. *Adaptive Control Processes: A Guided Tour*. Princeton University Press, New Jersey.

Bierlein, F.P., Fraser, S.J., Brown, W.M. & Lees, T. 2008. Advanced methodologies for the analysis of databases of mineral deposits and major faults. *Australian Journal of Earth Science*, **55**, 79–99, https://doi.org/10.1080/08120090701581406

Blake, D. & Kilgour, B. 1998. Geological Regions of Australia 1:5,000,000 scale [Dataset]. Geoscience Australia, Canberra. Available at: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_a05f7892-b237-7506-e044-00144fdd4fa6/Geological+Regions+of+Australia+2C+1%3A5+000+000+scale

Blake, D.H., Stewart, A.J., Sweet, I.P. & Hone, I.G. 1987. Geology of the Proterozoic Davenport province, central Australia (Bulletin 226). *In:* Department of Resources and Energy, Bureau of Mineral Resources, Geology and Geophysics, Canberra, ACT.

Bølviken, B., Bogen, J., Jartun, M., Langedal, M., Ottesen, R.T. & Volden, T. 2004. Overbank sediments: a natural bed blending sampling medium for large-scale geochemical mapping. *Chemometrics and Intelligent Laboratory Systems*, **74**, 183–199.

Bonham-Carter, G.F., Agterberg, F.P. & Wright, D.F. 1989. Weights of evidence modeling: a new approach to mapping mineral potential. *In:* Agterberg, F.P. & Bonham-Carter, G.F. (eds) *Statistical Applications in the Earth Sciences: Geological Survey Canada Paper*, **89–9**, 171–183.

Botnick, E. & White, K.T. 1998. Discussion of terms used by laboratories when speaking about detection, quantification and reporting limits. *The Synergist*, **1998**, 15–18.

Budd, A. 2001. *Georgetown Inlier Synthesis (No. 3777)*. Geoscience Australia, Canberra, ACT.

Carranza, E.J.M. 2009. *Geochemical anomaly and mineral prospectivity mapping in GIS. Handbook of exploration and environmental geochemistry 11*. Elsevier, UK, 351.

Carranza, E.J.M. 2010*a*. Catchment basin modelling of stream sediment anomalies revisited: incorporation of EDA and fractal analysis. *Geochemistry: Exploration, Environment, Analysis*, **10**, 365–381, https://doi.org/10.1144/1467-7873/09-224

Carranza, E.J.M. 2010*b*. Mapping of anomalies in continuous and discrete fields of stream sediment geochemical landscapes. *Geochemistry: Exploration, Environment, Analysis*, **10**, 171–187, https://doi.org/10.1144/1467-7873/09-223

Carranza, E.J.M. 2011. Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. *Journal of Geochemical Exploration*, **110**, 167–185, https://doi.org/10.1016/j.gexplo.2011.05.007

Cracknell, M.J., Reading, A.M. & McNeill, A.W. 2014. Mapping geology and volcanic-hosted massive sulfide alteration in the Hellyer–Mt Charter region, Tasmania, using Random Forests™ and Self-Organising Maps. *Australian Journal of Earth Science*, **61**, 287–304, https://doi.org/10.1080/08120099.2014.858081

Cracknell, M.J., Reading, A.M. & Caritat, P. de 2015. Multiple influences on regolith characteristics from continental-scale geophysical and mineralogical remote sensing data using Self-Organizing Maps. *Remote Sensing of Environment*, **165**, 86–99, https://doi.org/10.1016/j.rse.2015.04.029

Caritat, P. de & Cooper, M., 2011*a*. *National Geochemical Survey of Australia: The Geochemical Atlas of Australia*. (Record 2011/20). Geoscience Australia, Canberra, ACT.

Caritat, P. de & Cooper, M. 2011*b*. *National Geochemical Survey of Australia: Data Quality Assessment*. (Record 2011/21). Geoscience Australia, Canberra, ACT.

Caritat, P. de & Cooper, M. 2016. A continental-scale geochemical atlas for resource exploration and environmental management: the National Geochemical Survey of Australia. *Geochemistry: Exploration, Environment, Analysis*, **16**, 3–13, https://doi.org/10.1144/geochem2014-322

Caritat, P. de & Grunsky, E.C. 2013. Defining element associations and inferring geological processes from total element concentrations in Australian catchment outlet sediments: Multivariate analysis of continental-scale geochemical data. *Applied Geochemistry*, **33**, 104–126.

Caritat, P. de, Cooper, M., Mann, A. & Prince, P. 2011. Lithological signatures and exploration implications from overbank sediment sampling and MMI Analysis: Preliminary findings from the National Geochemical Survey of Australia. *In*: Programme & Abstracts. Presented at the 25th International Applied Geochemistry Symposium, Association of Applied Geochemists, Rovaniemi, Finland.

Dash, M., Choi, K., Scheuermann, P. & Liu, H. 2002*a*. Feature Selection for Clustering - A Filter Solution. *In:* Proceedings of the Second International Conference on Data Mining. 115–122.

Dash, M., Liu, H. & Yao, J. 2002*b*. Unsupervised Feature Ranking and Selection. *In:* Abramowicz, W. & Zurada, J. (eds) *Knowledge Discovery for Business Information Systems, The International Series in Engineering and Computer Science*. Springer, USA, 67–87.

Davies, D.L. & Bouldin, D.W. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-1**, 224–227.

Dy, J.G. & Brodley, C.E. 2004. Feature Selection for Unsupervised Learning. *Journal of Machine Learning Research*, **5**, 845–889.

Ewers, G.R., Evans, N., Hazell, M. & Kilgour, B. 2002. *OZMIN Mineral Deposits Database*. Geoscience Australia, Canberra, ACT.

Ferenczi, P.A. & Ahmad, M. 1998. Geology and mineral deposits of the Granites-Tanami and Tennant Creek Inliers, Northern Territory. *AGSO Journal of Australian Geology and Geophysics*, **17**, 19–33.

Furman, O., Caritat, P. de, Maher, W., Foster, S., Gruber, B. & Thompson, R.M. 2016. Mercury distribution in Australian catchment outlet sediments at the continental scale. *In:* Goldschmidt Conference Abstracts. Presented at the Goldschmidt2016, Yokohama, Japan, 881.

Geoscience Australia 2003. GEODATA TOPO 2.5M 2003 [Dataset]. Canberra, ACT. Available at: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_60804/GEODATA+TOPO+2.5M+2003

Geoscience Australia 2008. GEODATA 9 second Digital Elevation Model [Dataset], version 3. Canberra, ACT. Available at: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_66006

Gonzalez, R.C. & Woods, R.E. 2008. *Digital image processing.* 3rd ed. Prentice Hall Inc., Upper Saddle River, New Jersey.

Grunsky, E.C., Caritat, P. de & Mueller, U. 2017. Using surface regolith geochemistry to map the major crustal blocks of the Australian continent. *Gondwana Research*, **46**, 227–239.

Haralick, R.M., Shanmugam, K. & Dinstein, I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics. Systems, Man, and Cybernetics*, **3**, 610–621, https://doi.org/10.1109/tsmc.1973.4309314

Helsel, D.R. 2005. *Nondetects and Data Analysis: Statistics for Censored Environmental Data*. Wiley and Sons, New York.

Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, **43**, 59–69, https://doi.org/10.1007/bf00337288

Kohonen, T. 2001. *Self-Organizing Maps*, 3rd ed. Springer series in information sciences. Springer, Berlin.

Kohonen, T., Hynninen, J., Kangas, J. & Laaksonen, J. 1996. *SOM-PAK: The self-organizing map program package (Report A31)*. Helsinki University of Technology, Laboratory of Computer and Information Science, Helsinki, Finland.

Kruse, P.D. & Munson, T.J. 2013. Chapter 32: Wiso Basin. *In:* Ahmad, M. & Munson, T.J. (eds) *Geology and Mineral Resources of the Northern Territory, Special Publication 5*. Northern Territory Geological Survey, Darwin, NT, 12.

Lacassie, J.P. & Ruiz Del Solar, J. 2006. Knowledge extraction in geochemical data by using self-organizing maps. Presented at the 2006 International Joint Conference on Neural Networks, IJCNN, IEEE, Vancouver, British Columbia, 4878–4883.

Lacassie, J.P., Roser, B., Ruiz Del Solar, J. & Hervé, F. 2004. Discovering geochemical patterns using self-organizing neural networks: a new perspective for sedimentary provenance analysis. *Sedimentary Geology*, **165**, 175–191, https://doi.org/10.1016/j.sedgeo.2003.12.001

Leybourne, M.I. & Cameron, E.M. 2010. Groundwater in geochemical exploration. *Geochemistry: Exploration, Environment, Analysis*, **10**, 99–118, https://doi.org/10.1144/1467-7873/09-222

Löhr, S.C., Grigorescu, M., Hodgkinson, J.H., Cox, M.E. & Fraser, S.J. 2010. Iron occurrence in soils and sediments of a coastal catchment: A multivariate approach using self organising maps. *Geoderma*, **156**, 253–266, https://doi.org/10.1016/j.geoderma.2010.02.025

Mann, A.W. 2010. Strong versus weak digestions: ligand-based soil extraction geochemistry. *Geochemistry: Exploration, Environment, Analysis*, **10**, 17–26, https://doi/org/10.1144/1467-7873/09-216

Mann, A.W., Birrell, R.D., Fedikow, M.A.F. & Souza, H.A.F. de 2005. Vertical ionic migration: mechanisms, soil anomalies, and sampling depth for mineral exploration. *Geochemistry: Exploration, Environment, Analysis*, **5**, 201–210, https://doi.org/10.1144/1467-7873/03-045

Mann, A., Caritat, P. de & Prince, P. 2012. Bioavailability of nutrients in Australia from Mobile Metal Ion® analysis of catchment outlet sediment samples: continental-scale patterns and processes. *Geochemistry: Exploration, Environment, Analysis*, **12**, 277–292, https://doi.org/10.1144/geochem2011-090

Mann, A., Davidson, A. & Caritat, P. de 2013. High resolution soil geochemistry for gold exploration at the continental, regional and prospect scale. AusIMM World Gold Conference 2013 (Brisbane, Queensland, 26–29 September 2013), The Minerals Institute, Proceedings Volume.

McKenzie, N.J., Grundy, M.J., Webster, R. & Ringrose-Voase, A.J. 2008. *Guidelines for Surveying Soil and Land Resources, 2nd ed, Australian Soil and Land Survey Handbook*. CSIRO Publishing, Collingwood, Victoria.

Minty, B., Franklin, R., Milligan, P., Richardson, L.M. & Wilford, J. 2009. The Radiometric Map of Australia. *Exploration Geophysics*, **40**, 325–333.

Mueller, U., Lo, J., Caritat, P. de & Grunsky, E. 2014. Structural Analysis of the National Geochemical Survey of Australia Data. *In:* Pardo-Igúzquiza, E., Guardiola-Albert, C., Heredia, J., Moreno-Merino, L., Durán, J.J. & Vargas-Guzmán, A.J. (eds) *Mathematics of Planet Earth: Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences*. Springer Berlin Heidelberg, Berlin, Heidelberg, 99–102.

Ottesen, R.T., Bølviken, B., Volden, T. & Bogen, J. 1989. Overbank sediment: a representative sample medium for regional geochemical mapping. *Journal of Geochemical Exploration*, **32**, 257–277, https://doi.org/10.1016/0375-6742(89)90061-7

Downloaded from http://geea.lyellcollection.org/ by Michael David Campbell on April 3, 2019

Pain, C.F., Pillans, B.J., Roach, I.C., Worrall, L. & Wilford, J.R. 2012. Old, flat and red—Australia's distinctive landscape. *In:* Blewett, R.S. (ed.) *Shaping a Nation: A Geology of Australia*. Geoscience Australia and ANU E Press, Canberra, ACT, 227–275.

Penn, B.S. 2005. Using self-organizing maps to visualize high-dimensional data. *Computers & Geosciences*, **31**, 531–544, https://doi.org/10.1016/j.cageo.2004.10.009

Percival, P.J. 2014. *Index of Airborne Geophysical Surveys. 14th edn.* Record 2014/014, Geoscience Australia, Canberra, ACT.

Radke, B. 2009. *Hydrocarbon and Geothermal Prospectivity of Sedimentary Basins in Central Australia; Warburton, Cooper, Perdirka, Galilee, Simpson and Eromanga Basins.* Geoscience Australia Record 2009/25. Geoscience Australia, Canberra, ACT.

Raymond, O.L. & Gallagher, R. (eds) 2012. Surface Geology of Australia 1:2.5 million scale dataset 2012 edition [Dataset]. Geoscience Australia, Canberra. Available at: http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_b4088aa1-f875-2444-e044-00144fdd4fa6/Surface+Geology+of+Australia+1%3A2.5+million+scale+dataset+2012+edition

Reimann, C. & Filzmoser, P. 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statisitical treatment of geochemical and environmental data. *Environmental Geology*, **39**, 1001–1014, https://doi.org/10.1007/s002549900081

Scealy, J.L., Caritat, P. de, Grunsky, E.C., Tsagris, M.T. & Welsh, A.H. 2015. Robust principal component analysis for power transformed compositional data. *Journal of the American Statistical Association*, **110**, 136–148.

Siponen, M., Vesanto, J., Simula, O. & Vasara, P. 2001. An approach to automated interpretation of SOM. *In:* Allinson, N., Yin, H., Allinson, L. & Slack, J. (eds) *Advances in Self-Organising Maps*. Springer, London, 89–94.

Smart, J., Grimes, K.G., Doutch, H.F. & Pinchin, J. 1972. *The Mesozoic Carpentaria Basin and the Cainozoic Karumba Basin, North Queensland (Bulletin 202)*. Department of National Development & Energy, Bureau of Mineral Resources, Geology and Geophysics, Canberra, ACT.

Smith, K.G. 1972. *Stratigraphy of the Georgina Basin (Bulletin 111)*. Department of National Development, Bureau of Mineral Resources, Geology and Geophysics, Canberra, ACT.

Sun, X., Deng, J., Gong, Q., Wang, Q., Yang, L. & Zhao, Z. 2009. Kohonen neural network and factor analysis based approach to geochemical data pattern recognition. *Journal of Geochemical Exploration*, **103**, 6–16, https://doi.org/10.1016/j.gexplo.2009.04.002

Taylor, G. & Butt, C.R.M. 1998. The Australian regolith and mineral exploration. *AGSO Journal of Australian Geology and Geophysics*, **17**, 55–67.

Tracey, R., Bacchin, M. & Wynne, P. 2007. *AAGD07: A new absolute gravity datum for Australian gravity and new standards for the Australian National Gravity Database*. Presented at the ASEG2007 – 19th Geophysical Conference, CSIRO Publishing, 1–3.

Tsakovski, S., Kudlak, B., Simeonov, V., Wolska, L. & Namiesnik, J. 2009. Ecotoxicity and chemical sediment data classification by the use of self-organising maps. *Analytica Chimica Acta*, **631**, 142–152, https://doi.org/10.1016/j.aca.2008.10.053

UNCOVER 2012. Searching the Deep Earth: A Vision for Exploration Geoscience in Australia. Australian Academy of Science, Canberra, ACT. Available at: https://www.science.org.au/supporting-science/science-sector-analysis/reports-and-publications/searching-deep-earth-vision

Vesanto, J. & Alhoniemi, E. 2000. Clustering of the self-organizing map. *IEEE Transactions On Neural Networks*, **11**, 586–600, https://doi.org/10.1109/72.846731

Wilford, J. 2012. A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis. *Geoderma*, **183–184**, 124–142, https://doi.org/10.1016/j.geoderma.2010.12.022

Withnall, I.W. & Hutton, I.J. 2013. Chapter 2 North Australian Craton. *In:* Jell, P. A. (ed.) *Geology of Queensland*. Geological Survey of Queensland, Brisbane, QLD, 23–111.

Wygralak, A.S. & Bajwah, Z.U. 1998. Geology and mineralisation of the Arunta Inlier, Northern Territory. *AGSO Journal of Australian Geology and Geophysics*, **17**, 35–45.

Yan, J. 2010. som: Self-Organizing Map, R package version 0.3-5. Available at: http://CRAN.R-project.org/package=som

Žibret, G. & Šajn, R. 2010. Hunting for Geochemical Associations of Elements: Factor Analysis and Self-Organising Maps. *Mathematical Geosciences*, **42**, 681–703, https://doi.org/10.1007/s11004-010-9288-3