2017

Using surface regolith geochemistry to map the major crustal blocks of the Australian continent

E. C. Grunsky

P. de Caritat

Ute A. Mueller
Edith Cowan University, u.mueller@ecu.edu.au

10.1016/j.gr.2017.02.011
Grunsky, E. C., De Caritat, P., & Mueller, U. A. (2017). Using surface regolith geochemistry to map the major crustal blocks of the Australian continent. *Gondwana Research*, *46*, 227-239.

Gondwana Research 46 (2017) 227–239

Contents lists available at ScienceDirect

Gondwana Research

journal homepage: www.elsevier.com/locate/gr

Using surface regolith geochemistry to map the major crustal blocks of the Australian continent

E.C. Grunsky [a,*], P. de Caritat [b,c], U.A. Mueller [d]

[a] Department of Earth and Environmental Sciences, University of Waterloo, N2L 3G1, Canada
[b] Geoscience Australia, GPO Box 378, Canberra, ACT 2601, Australia
[c] Research School of Earth Sciences, The Australian National University, Canberra, ACT 2601, Australia
[d] School of Science, Edith Cowan University, Joondalup, WA 6027, Australia

ARTICLE INFO

Article history:
Received 19 November 2016
Accepted 21 February 2017
Available online 22 March 2017

Handling Editor: F. Pirajno

Keywords:
National Geochemical Survey of Australia
Compositional data
Multivariate statistics
Posterior probability
Minimum/maximum autocorrelation factor analysis
Major crustal boundaries

ABSTRACT

Multi-element near-surface geochemistry from the National Geochemical Survey of Australia has been evaluated in the context of mapping the exposed to deeply buried major crustal blocks of the Australian continent. The major crustal blocks, interpreted from geophysical and geological data, reflect distinct tectonic domains comprised of early Archean to recent Cenozoic igneous, metamorphic and sedimentary rock assemblages. The geochemical data have been treated as compositional data to uniquely describe and characterize the geochemistry of the regolith overlying the major crustal blocks across Australia according to the following workflow: imputation of missing/censored data, log-ratio transformation, multivariate statistical analysis, multivariate geospatial (minimum/maximum autocorrelation factor) analysis, and classification. Using cross validation techniques, the uniqueness of each major crustal block has been quantified. The ability to predict the membership of a surface regolith sample to one or more of the major crustal blocks is demonstrated. The predicted crustal block assignments define spatially coherent regions that coincide with the known crustal blocks. In some areas, inaccurate predictions are due to uncertainty in the initial crustal boundary definition or from surficial processes that mask the crustal block geochemical signature. In conclusion, the geochemical composition of the Australian surface regolith generally can be used to map the underlying crustal architecture, despite secondary modifications due to physical transport and chemical weathering effects. This methodology is however less effective where extensive and thick sedimentary basins such as the Eromanga and Eucla basins overlie crustal blocks.

1. Introduction

1.1. Background

The composition of surface sediment or soil (s) varies spatially as a result of changes in climate (cl), organisms (o), relief (r), parent material (p) and time (t). This was expressed over 75 years ago by Jenny (1941) in the well-known expression: $s = f(cl, o, r, p, t)$. Although this relationship has been understood conceptually for a long time, it is only recently that its quantitative prediction has become possible through (1) the development of large digital databases, (2) advances in analytical capabilities, and (3) application of multivariate and data mining statistical methods. The digital soil mapping community has been quick to show how this predictive capability can benefit knowledge discovery in pedology (e.g., McBratney et al., 2003). Applications of this new 'big data' approach in the geosciences are emerging. In a lake sediment study in Nunavut, Canada, Grunsky et al. (2014) concluded that multivariate statistical analysis of the geochemical data provided an objective basis for validating and potentially improving existing geological maps. Thus insights into parent material (p) of the sediments could be gained. Similarly, Drew et al. (2010) and Grunsky et al. (2013a, 2013b) showed that soil geochemistry over the United States of America could accurately map climate zones (cl), ecosystems ($\sim o$), landforms ($\sim r$), and surface geology ($\sim p$) at the continental scale. All these variables are proxies for the soil formation controls listed above. Previous work by Caritat and Grunsky (2013) and Grunsky et al. (2014) demonstrated the usefulness of applying multivariate statistical methods such as principal component analysis to describe linear combinations of elements that ultimately are controlled by the stoichiometry of minerals and associated geochemical/geological processes.

The Australian continent preserves a diverse geological record spanning over 3 Ga, and is composed of different crustal blocks delimited by major crustal boundaries (e.g., Korsch and Doublier, 2016). In the present contribution we apply statistical analysis to the National Geochemical Survey of Australia (NGSA) dataset. We compare the geochemical

* Corresponding author.
E-mail addresses: egrunsky@gmail.com, egrunsky@uwaterloo.ca (E.C. Grunsky).

http://dx.doi.org/10.1016/j.gr.2017.02.011

composition of surface sediments collected on different crustal blocks, then calculate and map the posterior probabilities (PP) of all NGSA samples relative to those crustal blocks. The scientific question motivating this investigation can be stated as follows: Can we predict the location and extent (spatial distribution) of Australian crustal blocks based on surface regolith geochemistry using multivariate statistical and geospatial techniques? A positive answer to this question opens up an avenue for mapping deeper geological features from surface materials even with poor/absent outcrop, with implications for geological mapping as well as mineral exploration through cover. This effort aligns directly with two main themes of the current Australian government-industry-academia UNCOVER strategic initiative for '*providing the knowledge base and technology that will substantially increase the success rate of mineral exploration beneath post-mineralisation cover in Australia*' (UNCOVER, 2016). The present contribution represents an extension of the earlier statistical analysis of NGSA data presented by Caritat and Grunsky (2013) and Mueller et al. (2014).

1.2. The National Geochemical Survey of Australia

The National Geochemical Survey of Australia (NGSA) project, a cooperation between the Federal and State/Northern Territory (NT) geoscience agencies in Australia, was part of the 5-year Onshore Energy Security Program managed at Geoscience Australia between 2006 and 2011 (Johnson, 2006). The NGSA was initiated to bridge a significant knowledge gap in Australia about the composition of surface regolith at the continental scale by providing internally consistent pre-competitive (i.e., government sponsored, regional and publicly available) data and knowledge to aid exploration for energy and mineral resources (Caritat and Cooper, 2016). The NGSA delivered a continental geochemical atlas (Caritat and Cooper, 2011), a series of reports, and a public-domain geochemical dataset (www.ga.gov.au/ngsa).

Previously existing geochemical data in Australia were perceived to have limited spatial coverage, poor internal comparability and/or inadequate extent of reported elemental compositions (Caritat et al., 2008a). During a series of pilot projects that preceded the NGSA a variety of sampling media, depths and grain-size fractions were tested at different sampling densities (Caritat et al., 2005, 2007, 2008b; Caritat and Lech, 2007; Lech and Caritat, 2007a, 2007b). It was found that regolith samples collected on floodplains or similar low-lying depositional landforms near the outlet of large catchments, or 'catchment outlet sediments', provided a ubiquitous sampling medium suitable for low-density geochemical mapping in Australia. These materials usually are a fine-grained mixture of detrital material originating from the main rock and soil types found within a catchment (Ottesen et al., 1989; Bølviken et al., 2004). In places, a

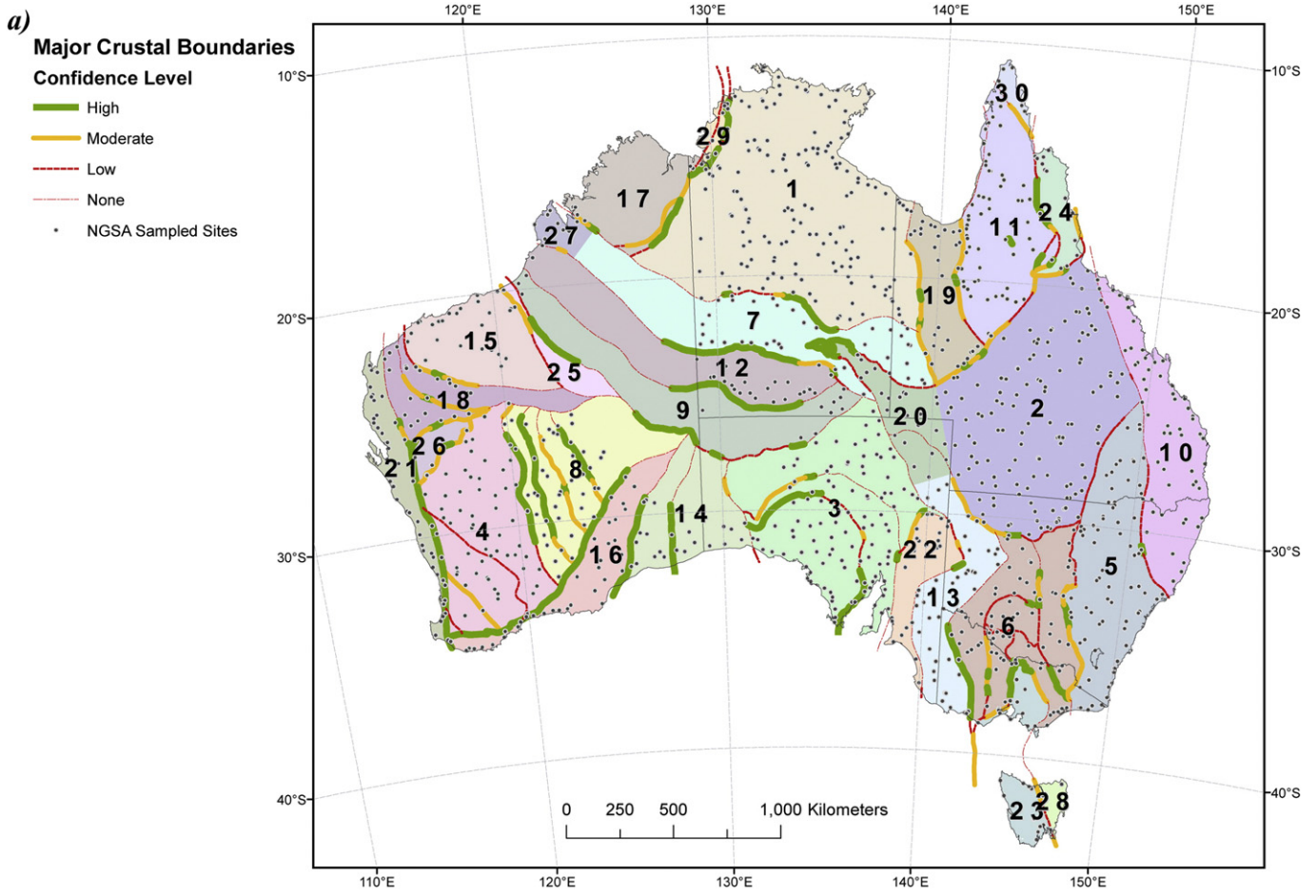


Fig. 1. (a) Map showing the major crustal blocks (MCBs) of Australia used in the present analysis (coloured and numbered)). The line styles of the MCB boundaries reflect the confidence level in their position/existence (solid thick: high; solid thin: moderate; dashed: low; dot-dashed: none). (b) Map showing the surface geology and the geological regions of Australia. The NGSA sample site locations are shown as black dots on both maps. All maps are presented here in the Lambert conformal conic projection of Australia (with standard parallels at 18 and 36°S latitude, central meridian at 134°E longitude and earth ellipsoid GRS80). Sources: Blake and Kilgour (1998), Caritat and Cooper (2011), Korsch and Doublier (2015, 2016), Nakamura and Milligan (2015), Raymond (2012).

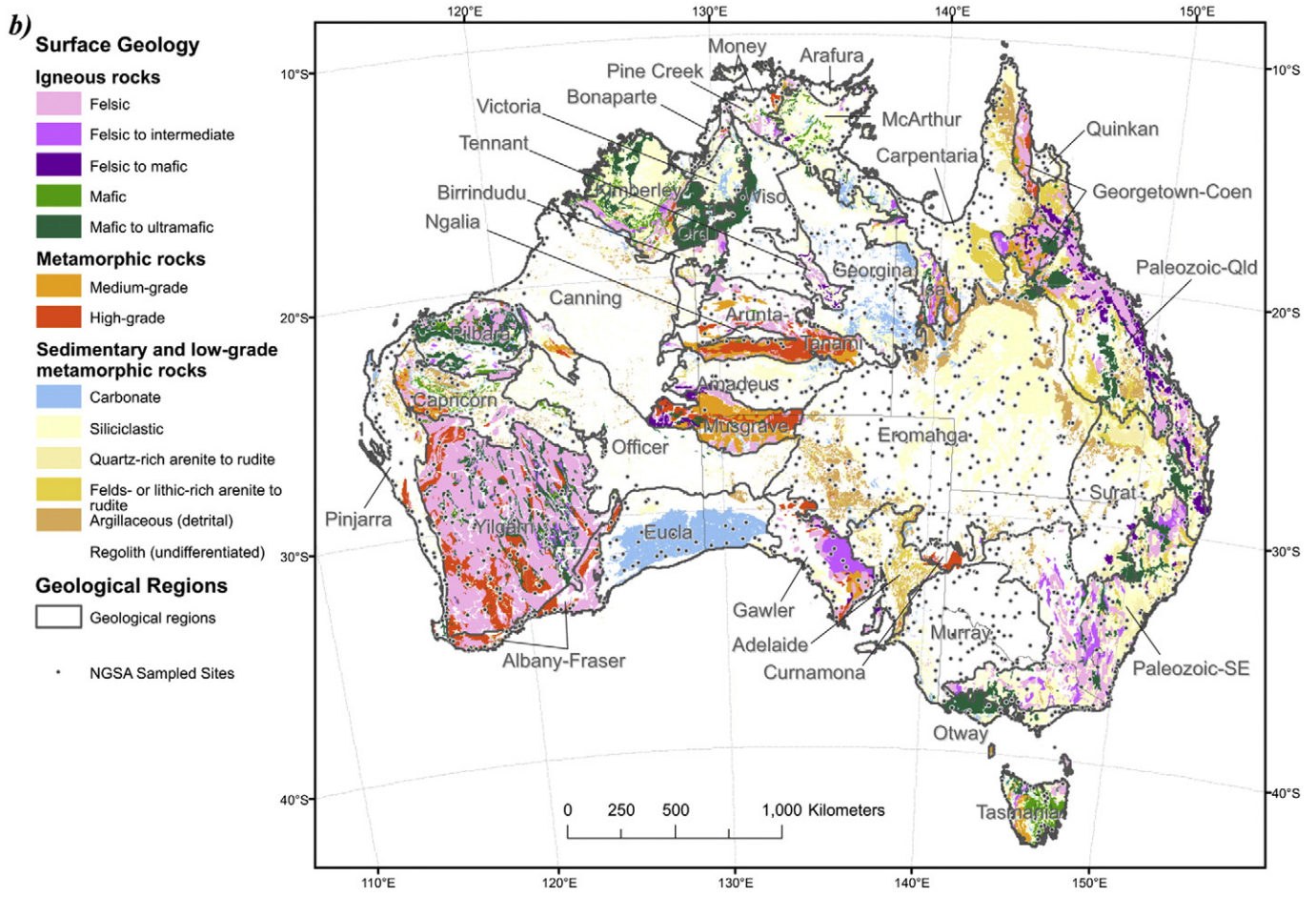


Fig. 1 (*continued*).

component of eolian material is inevitably present in these settings in Australia. The effect of quartz-rich sand undoubtedly will interfere with the statistical analysis of geochemical data from surficial materials, however, the origin of the quartz-rich sand is likely the result of weathering from nearby crustal blocks. Thus, the eolian effect is expected to 'smear' the boundaries of the blocks. Pedogenic processes have occurred to some extent in those landforms and to those materials over time, and the sample material collected can be referred to as soil (topsoil and subsoil) developed on (allochthonous) fluvial ± eolian sediment as the parent material (Caritat et al., 2011, 2012).

Many geochemical surveys have been carried out in Australia by Federal and State/NT geoscience agencies, industry, and/or academic institutions (e.g., Cox and Curtis, 1977; Ryall and Nicholas, 1979; Grunsky, 1990, 1991; Bain and Draper, 1997; Cohen et al., 1999; Davy et al., 1999; Morris et al., 2003; Cornelius et al., 2008). Before the NGSA, none was continental in scope mainly because a 'universal' sampling medium had not been recognised and also because of the perception that a relatively high sampling density was required, which would have been prohibitively expensive (Walker, 1978; L. Wyborn, pers. comm., 2003). Thus, the NGSA (Caritat and Cooper, 2011) is the first local attempt at delivering a uniform, internally consistent geochemical atlas and database using a common sampling medium, harmonised sampling protocols, and state-of-the-art multi-element analytical equipment for an area of 6,172,000 km^2 in extent covering ~81% of the continent (Fig. 1). By necessity, the sampling density of NGSA was very low, on average 1 site/5200 km^2, similar to, for instance, the European FOREGS project (stream sediments and soils collected at 1 site/4700 km^2; Salminen, 2005; De Vos and Tarvainen, 2006) and the 'Shacklette dataset' collected over the conterminous United States of America (soil and other surficial materials collected at 1 site/5925 km^2; Shacklette and Boerngen, 1984; Gustavsson et al., 2001). Recent applications of the NGSA data were reviewed in Caritat and Cooper (2016).

Table 1
Sample Site Count for Major Crustal Blocks. Shaded cells were not considered for further evaluation because of insufficient number of sample sites.

Block	1	2	3	4	5	6	7	8	9	10
NGSA Count	568	448	380	244	284	268	144	156	40	208
Block	11	12	13	14	15	16	17	18	19	20
NGSA Count	228	92	152	72	100	124	20	80	100	76
Block	21	22	23	24	25	26	27	28	29	30
NGSA Count	156	80	84	56	4	8	36	19	20	24

1.3. Major crustal blocks of Australia

We classify the surface geochemistry dataset according to the major crustal blocks (MCBs) of Australia. The MCBs were obtained from polygonising and simplifying the major crustal boundaries of Korsch and Doublier (2015, 2016), which were defined to better

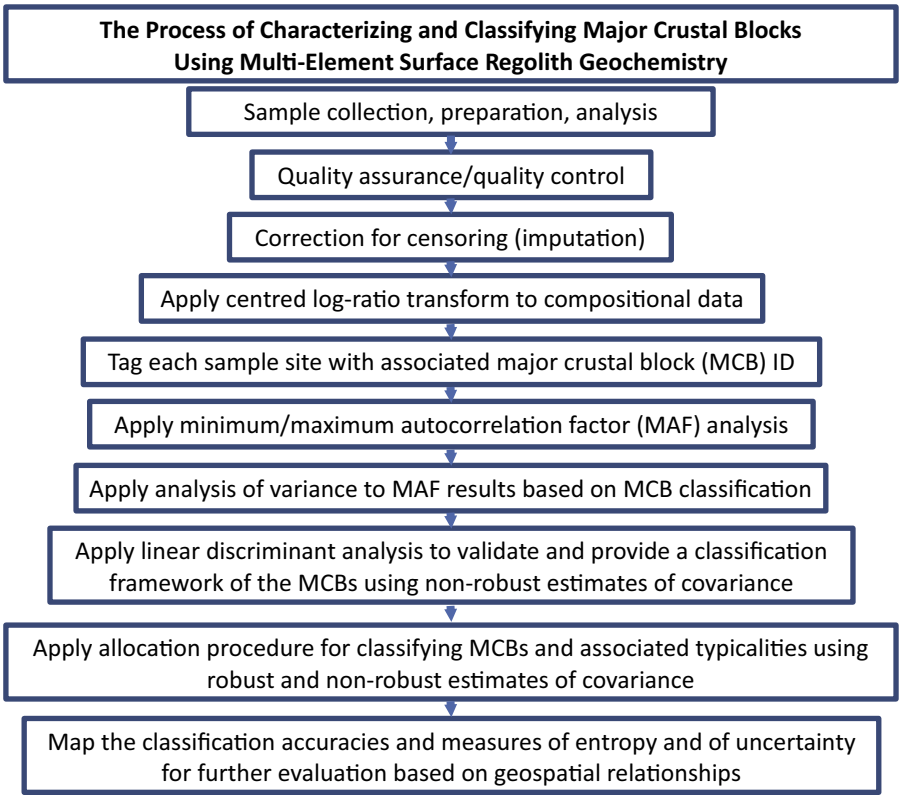


Fig. 2. Workflow developed for evaluating and predicting the major crustal blocks (MCBs) based on the NGSA data, minimum/maximum autocorrelation factor analysis and subsequent classification.

understand the crustal architecture and geodynamic evolution of Australia's key geological provinces and basins. For the generation of the major crustal boundaries dataset, the main crustal breaks interpreted from regional, deep seismic reflection lines collected across the continent since 1980 were used as starting points. Next, those major crustal boundaries were mapped in plan view away from the seismic profiles, using geological (e.g., outcrop maps, drill hole information, geochronology, isotope data) and geophysical (e.g., gravity, aeromagnetic, magnetotelluric) data. While some boundaries are exposed (e.g., Yilgarn Craton), others are covered by younger sediments/regolith and occur at greater depth. In this case the position of the boundary in map view represents the vertical surface projection of the position interpreted at depth in the deep seismic reflection data. Because a boundary between crustal blocks is not necessarily vertical and can be oblique, the surface projection of a contact between blocks can be located at some horizontal distance from the oblique upward continuation of the boundary, especially under deep sedimentary basins such as the Eromanga Basin in northeastern Australia. It is important to note that most deep crustal materials are crystalline basement rocks, except in some cases where metamorphosed sedimentary rocks may exist at quite some depth, e.g., under the Thomson Orogen in northeastern Australia.

Where crustal blocks were too small relative to the density of NGSA sampling points, they were merged either together or with an adjacent, larger block in the most geologically sensible fashion for the present analysis (M. Doublier, pers. comm., 2015). A map of the resulting 30 MCBs is shown in Fig. 1a. Table 1 shows the number of NGSA samples included in each MCB. Of the 30 MCBs, eight (MCBs 9, 17, 25–30) did not have a sufficient number of samples to warrant further statistical processing and were consequently dropped from the dataset. Supplementary material Table 1 shows the summary statistics of the chemical composition for each of the 22 MCBs that were retained for statistical evaluation and prediction. Fig. 1b shows the distribution of surface lithology (Raymond, 2012) and of the geological regions (Blake and Kilgour, 1998) over Australia, which will be referred to in the Results and discussion below.

1.4. Motivation for this study

Caritat and Grunsky (2013) applied principal component analysis (PCA) to the four combined sample types of the NGSA dataset (Caritat and Cooper, 2011) after a centred log-ratio (clr) transformation was applied (see Supplementary material). The resulting principal components enabled the description and interpretation of a number of geological and geochemical processes, and formed the basis for extending the NGSA data analysis into the prediction of spatially and geochemically defined regions, as developed here. For the present study, MCBs interpreted from geological and geophysical datasets (Korsch and Doublier, 2015, 2016) were used as a framework to validate any geochemically distinct blocks. Of particular significance in the present study is the ability to identify and predict these distinct crustal blocks from the near-surface soil geochemistry. Since their formation these crustal blocks have been subjected to various physical and chemical processes (e.g., uplift, weathering, erosion, alluvial, colluvial and eolian deposition, and locally glaciation), implying that their primary geochemical signature has been in part obscured/modified by secondary processes.

The use of the centred log-ratio transformation ensures that the data are represented within the real number space along with the relative relationships of the data. The compromise with this approach is the possibility of not achieving multivariate normality in

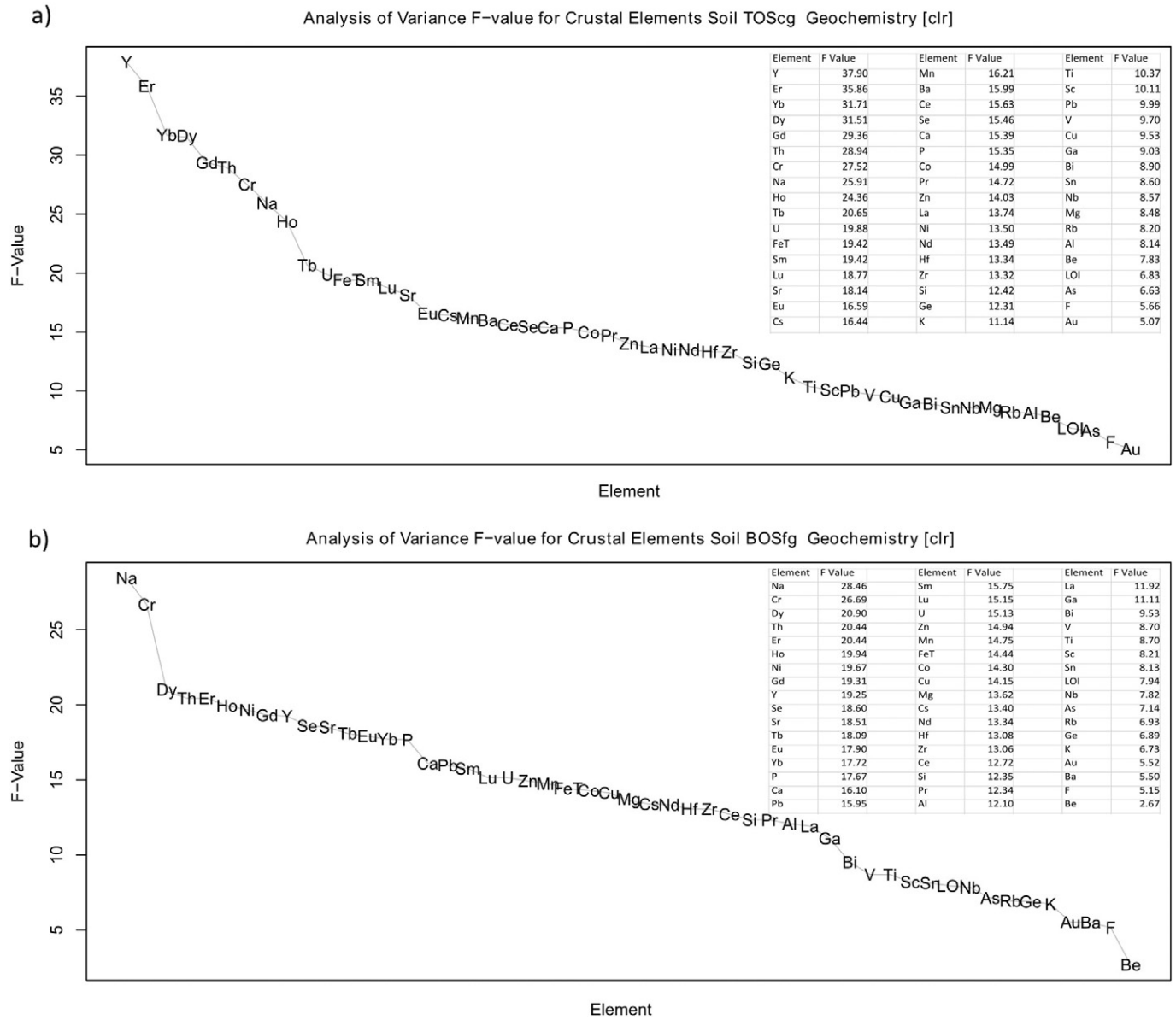


Fig. 3. Plots of ordered F-values for elements (a: TOS c/g; b: BOS f/g) as a measure of discrimination power. Elements with high F-values are better at discriminating between the major crustal blocks. The concentrations were centred log-ratio transformed.

the transformed data. However, we are not able to achieve perfection in the mathematical/statistical treatment of the data, as seen in the summary statistics of Supplementary Table 2, which contains summary statistics for the minimum/maximum autocorrelation factors (MAF) for each of the major crustal blocks. In this study, we intend to show that the definition of the MCBs is controlled by the relative relationships of the soil geochemistry. Non-normality may have an influence on this but our conclusion is that the influence is not significant. Our experiments with robust versus non-robust methods indicated that the results of the classification do not change by any significant amount. Our conclusions are the same regardless of the use of robust versus non-robust methods.

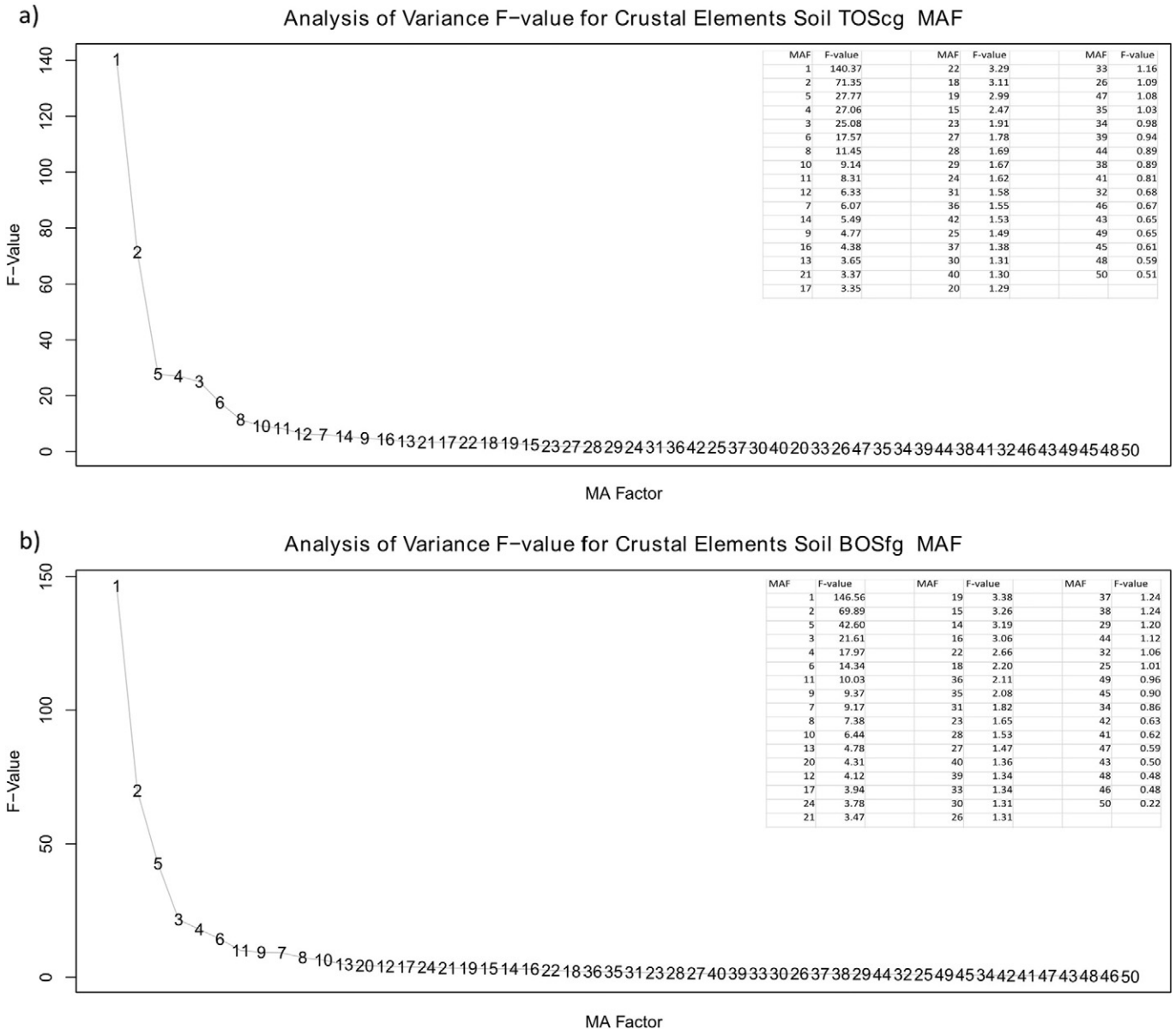
2. Methods

Details on the sample collection, preparation and analysis, as well as the methods of data evaluation, analysis and classification are provided as Supplementary material. The workflow developed from the methods employed in this study is graphically represented in Fig. 2. In the following, 'TOS' refers to 'top outlet sediment' (0–10 cm depth), 'BOS' to 'bottom outlet sediment' (~60–80 cm depth), 'c/g' to 'coarse-grained' (<2 mm grain-size fraction), and 'f/g' to 'fine-grained' (<75 μm grain-size fraction), as detailed in the Supplementary material.

3. Results and discussion

3.1. Process discovery using minimum/maximum autocorrelation factor (MAF) analysis

MAF analysis does not offer the equivalent 'process discovery' of patterns or trends of multivariate data that can be observed in PCA biplots. Based on the previous work by Caritat and Grunsky (2013) and preliminary research for this paper, PCA biplots of the NGSA data coded by



MAF	F-value	MAF	F-value	MAF	F-value
1	140.37	22	3.29	33	1.16
2	71.35	18	3.11	26	1.09
5	27.77	19	2.99	47	1.08
4	27.06	15	2.47	35	1.03
3	25.08	23	1.91	34	0.98
6	17.57	27	1.78	39	0.94
8	11.45	28	1.69	44	0.89
10	9.14	29	1.67	38	0.89
11	8.31	24	1.62	41	0.81
12	6.33	31	1.58	32	0.68
7	6.07	36	1.55	46	0.67
14	5.49	42	1.53	43	0.65
9	4.77	25	1.49	49	0.65
16	4.38	37	1.38	45	0.61
13	3.65	30	1.31	48	0.59
21	3.37	40	1.30	50	0.51
17	3.35	20	1.29		

MAF	F-value	MAF	F-value	MAF	F-value
1	146.56	19	3.38	37	1.24
2	69.89	15	3.26	38	1.24
5	42.60	14	3.19	29	1.20
3	21.61	16	3.06	44	1.12
4	17.97	22	2.66	32	1.06
6	14.34	18	2.20	25	1.01
11	10.03	36	2.11	49	0.96
9	9.37	35	2.08	45	0.90
7	9.17	31	1.82	34	0.86
8	7.38	23	1.65	42	0.63
10	6.44	28	1.53	41	0.62
13	4.78	27	1.47	47	0.59
20	4.31	40	1.36	43	0.50
12	4.12	39	1.34	48	0.48
17	3.94	33	1.34	46	0.48
24	3.78	30	1.31	50	0.22
21	3.47	26	1.31		

Fig. 4. Plots of ordered F-values for MAF (a: TOS c/g; b: BOS f/g) as a measure of discrimination power. Factors with high F-values are better at discriminating between the major crustal blocks.

MCB identifiers demonstrated similar linear trends and patterns that exhibit distinct differences and overlaps amongst the MCBs. However, MAF analysis also includes the spatial context of the multi-element geochemistry, which will provide a spatially coherent pattern of trends associated with the MCBs, forming the basis for the following discussion on process validation.

3.2. Process validation of MAF using analysis of variance (ANOVA)

A critical step in process validation is the identification of variables that assist in uniquely defining the classes assigned to the data through the application of an ANOVA. ANOVA can be carried out on the clr transformed concentrations or the MAFs derived from the clr transformed data. The results of the ANOVA based on the clr transformed elements for the TOS c/g and BOS f/g are shown in Fig. 3a,b. The figures show the ordered F-values, which provide a measure of group uniqueness (see Supplementary material section Methods 2.5), of the elements based on their ability to separate the 22 MCBs used here.

The plots of F-values for the elements describe the “relative” discriminating power for the variables (elements). Maximum discriminating power is obtained by using all of the elements. However, this can create degrees-of-freedom problems if there are too few observations. Thus, as a rule-of-thumb, a reduced set of elements are selected based on the identification of “inflection points” along the F-value trend. In Fig. 3a, Y-Er-Yb-Dy-Gd-Th-Cr-Na-Ho-Tb could be considered the “best” suite of elements.

For both media there is a gradual decay of F values for the elements. For TOS c/g the ten best discriminators are the elements Y, Er, Yb, Dy, Gd, Th, Cr, Na, Ho and Tb whereas the variables Au, F, As, LOI, Be, Al, Rb and Mg are poorest for group separation.

For BOS f/g eight of the ten best discriminators are the same as for TOS c/g albeit the discriminatory power differs. Similarly, five of the

Table 2
Diagonal of predictive accuracies (%) for the four sample types of the NGSA and the 22 major crustal blocks (MCBs), ordered by decreasing average accuracy for TOS c/g. Cells are coloured from highest (red) to lowest (blue).

MCB	TOS c/g	TOS f/g	BOS c/g	BOS f/g
MCB14	0.94	0.89	0.83	0.89
MCB01	0.87	0.82	0.81	0.85
MCB02	0.74	0.80	0.71	0.72
MCB04	0.72	0.69	0.57	0.69
MCB08	0.69	0.69	0.62	0.72
MCB03	0.67	0.74	0.67	0.69
MCB15	0.60	0.48	0.72	0.64
MCB18	0.60	0.70	0.60	0.65
MCB07	0.53	0.56	0.42	0.53
MCB21	0.49	0.62	0.56	0.44
MCB10	0.46	0.63	0.42	0.44
MCB05	0.42	0.44	0.41	0.24
MCB06	0.40	0.45	0.55	0.54
MCB12	0.39	0.30	0.48	0.48
MCB11	0.30	0.44	0.32	0.32
MCB23	0.29	0.24	0.33	0.43
MCB24	0.29	0.36	0.14	0.14
MCB20	0.26	0.47	0.32	0.21
MCB16	0.23	0.19	0.19	0.26
MCB13	0.21	0.26	0.18	0.29
MCB22	0.15	0.25	0.10	0.15
MCB19	0.04	0.16	0.16	0.12
Overall Accuracy	TOS c/g	TOS f/g	BOS c/g	BOS f/g
	0.55	0.59	0.54	0.55

elements (As, Au, Be, F and Rb) are poor at group separation for both media.

Figs. 4a,b show the ordered F-values based on the MAFs for the TOS c/g and BOS f/g sample types. These plots are significantly different from the ordered plots of F-values for the variables of Figs. 3a,b. The charts that accompany Figs. 4a,b show that MAF1, followed by MAFs 1, 2, 5, 4, 3 and 6 have high F-values relative to the other MAFs for the TOS c/g sample type. The remaining MAF F-values are quite low and the graph shows a gradual decay in their values. For the BOS f/g the chart shows that MAFs 1, 2, 5, 3, 4 and 6 account for most of the discrimination between the MCBs. We found that the grain-size based F-value patterns (c/g vs f/g) have a greater contrast and discrimination ability compared to the depth based ones (TOS vs BOS), a result echoing the PCA based findings of Caritat and Grunsky (2013). For the sake of brevity, the remainder of the discussion of the results therefore will focus on the TOS c/g and BOS f/g classifications.

3.3. Linear discriminant analysis (LDA)

LDA and classification of the MCBs were applied to the MAFs of the NGSA samples. Supplementary Table 3 shows a matrix of classification accuracies from the LDA for the four sample types and the 22 classes. The diagonal along each matrix gives the proportion of correctly classified samples for each MCB, based on the training sets derived from cross validation. These classification accuracies range from 0.04 (MCB 19 TOS c/g) to 0.94 (MCB 14 TOS c/g). The overlap (confusion) of the classification is shown in the off-diagonal elements. The overall accuracies of the classification, which are defined as the sum of the diagonal of the classification accuracy matrix, divided by the row sums of the classification accuracy matrix, are shown in Table 2 with values of: 0.55 (TOS c/g), 0.59 (TOS f/g), 0.54 (BOS c/g) and 0.55 (BOS f/g). These averages are weighted according to the number of sites within each MCB. These nearly identical accuracies suggest that there is little difference between the four media with respect to their ability to correctly classify the MCBs. For some of the MCBs the predictive accuracies are different between the media and this is discussed below.

In the following we are going to discuss maps of PPs for selected MCBs. Although the PPs are compositional in nature, the application of a log-ratio transformation is problematic because of the existence of 0 probabilities and collinearity issues so that compositional co-kriging cannot be used for the generation of maps. Thus, the PP maps were interpolated using soft indicator co-kriging. The following MCBs were selected for closer scrutiny (listed from west to

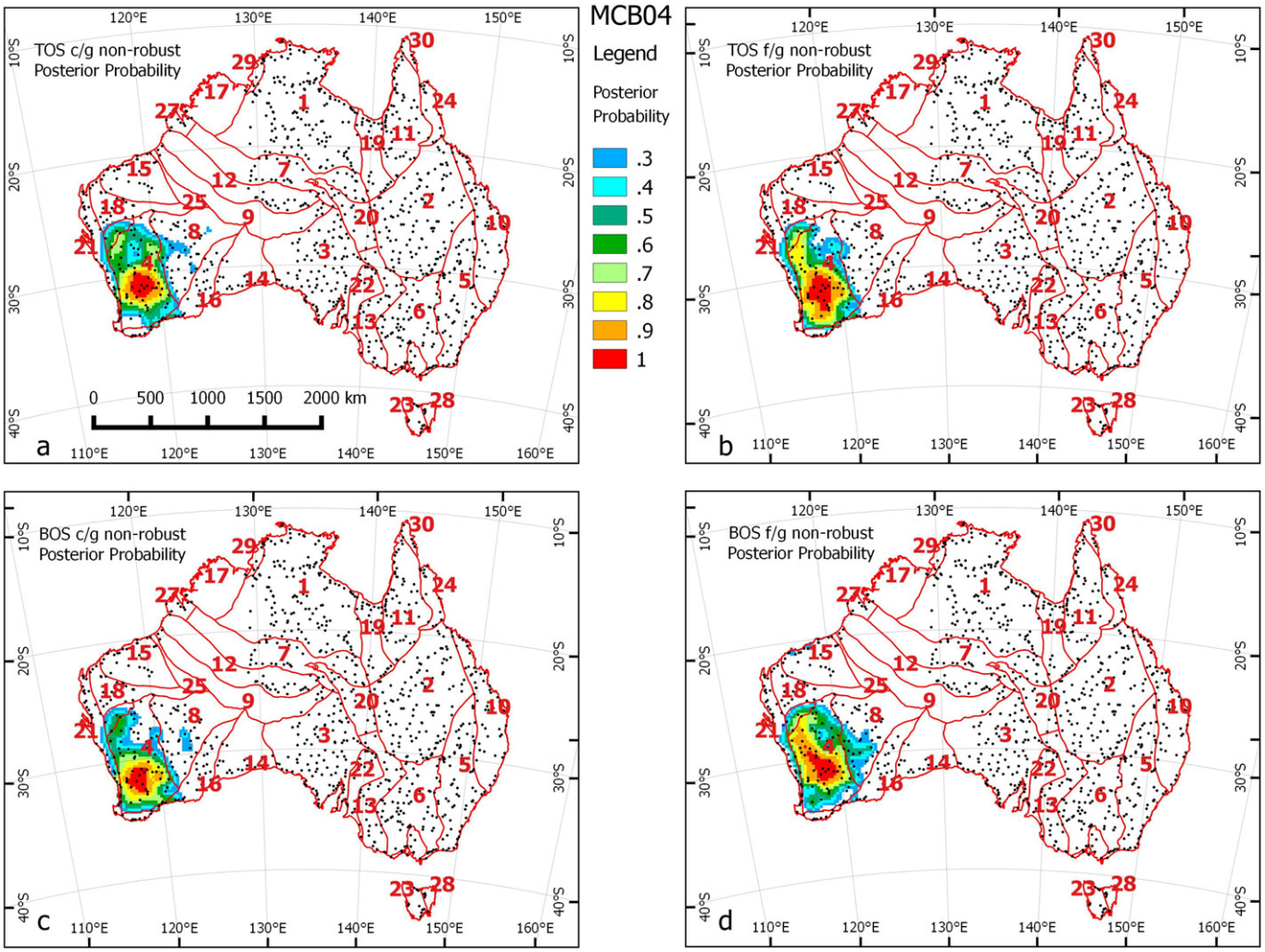


Fig. 5. Maps of the co-kriged PP predictions for major crustal block (MCB) 4 for the TOS c/g (a), TOS f/g (b), BOS c/g (c) and BOS f/g (d) sample types.

east): 4, 15, 18, 1, 3, 19, 11, 2, 6, and 10, because they cover a representative range of larger to smaller blocks, and situations with greater to lesser degrees of accuracy in the predictions. For brevity, only three maps of the co-kriged MCB predictions are shown, namely for MCBs 4, 2 and 6. The results for all MCBs are available in the Supplementary material.

Fig. 5 shows maps of the co-kriged PP for MCB 4 based on a non-robust LDA. Only PP values > 0.2 are shown. Regardless of sample type, the predictive accuracy of MCB 4 is high especially in the southwestern part of the block. High PPs (warm colours) are marginally more extensive for the f/g than for the c/g sample types. PPs > 0.2 only occur within and near the border of MCB 4, indicating that its composition is geochemically quite unique based on the statistical methods applied in this study.

Fig. 6 shows the PPs for MCB 2. This block shows PPs as high as 0.9 (c/g fractions) and overlaps with MCBs 1 (BOS c/g), 5, 6, 7, 10, 11, 13, 19, and 20. This is also reflected in Supplementary material Table 3. The highest probabilities (>0.6), however, are contained within the boundary of MCB 2. The smearing into the adjacent crustal blocks likely reflects the fact that much of the surface material for the sediments atop this crustal block consists of the Mesozoic Great Artesian (Eromanga) Basin, which contains clastic material from the surrounding highlands, particularly to the southwest (MCB 13), northwest (MCB 19) and east (MCBs 10, 5). Conversely, younger (mostly Cenozoic) sedimentary basins to the south (Murray-Darling Basin; MCB 6), west (Lake Eyre Basin; MCB 20), and north (Carpentaria/Karumba Basin; MCB 11), partly contain recycled clastic material from the Great Artesian Basin, explaining the compositional overlap in these directions. Thus, over the area of MCB 2, very little basement material proper crops out (e.g., the Anakie Inlier) and thus the surface geochemistry in this case more likely reflects the overlying basins' geochemical signature rather than the true basement signature. Note the similarity, in places, between the PP contours and the Eromanga geological region (see Fig. 1b). For the BOS c/g sample type there is minor confusion with the central region of MCB 1, which is likely attributable to similar mineralogy.

In Fig. 7, the PPs of MCB 6 reach a maximum of 0.6, which is low relative to many of the other MCBs. There is some overlap observed with MCBs 2, 5, 11, 13, and 24, which is also indicated by the confusion values in Supplementary material Table 3. Most of these overlaps, which are the result of recycled clastic material and recent erosional/depositional processes have been discussed above.

Table 2 also shows the diagonal elements of the four matrices coloured according to the accuracy of prediction. They are also sorted according to the predictive accuracy of the TOS c/g sample type. The table shows that the MCBs with the highest classification accuracies are: MCBs 14, 1, 2, 4, 8, 3, 15 and 18, whilst those MCBs with the lowest classification accuracies are MCBs 19, 22, 13, 16, 20, 24, 23 and 11.

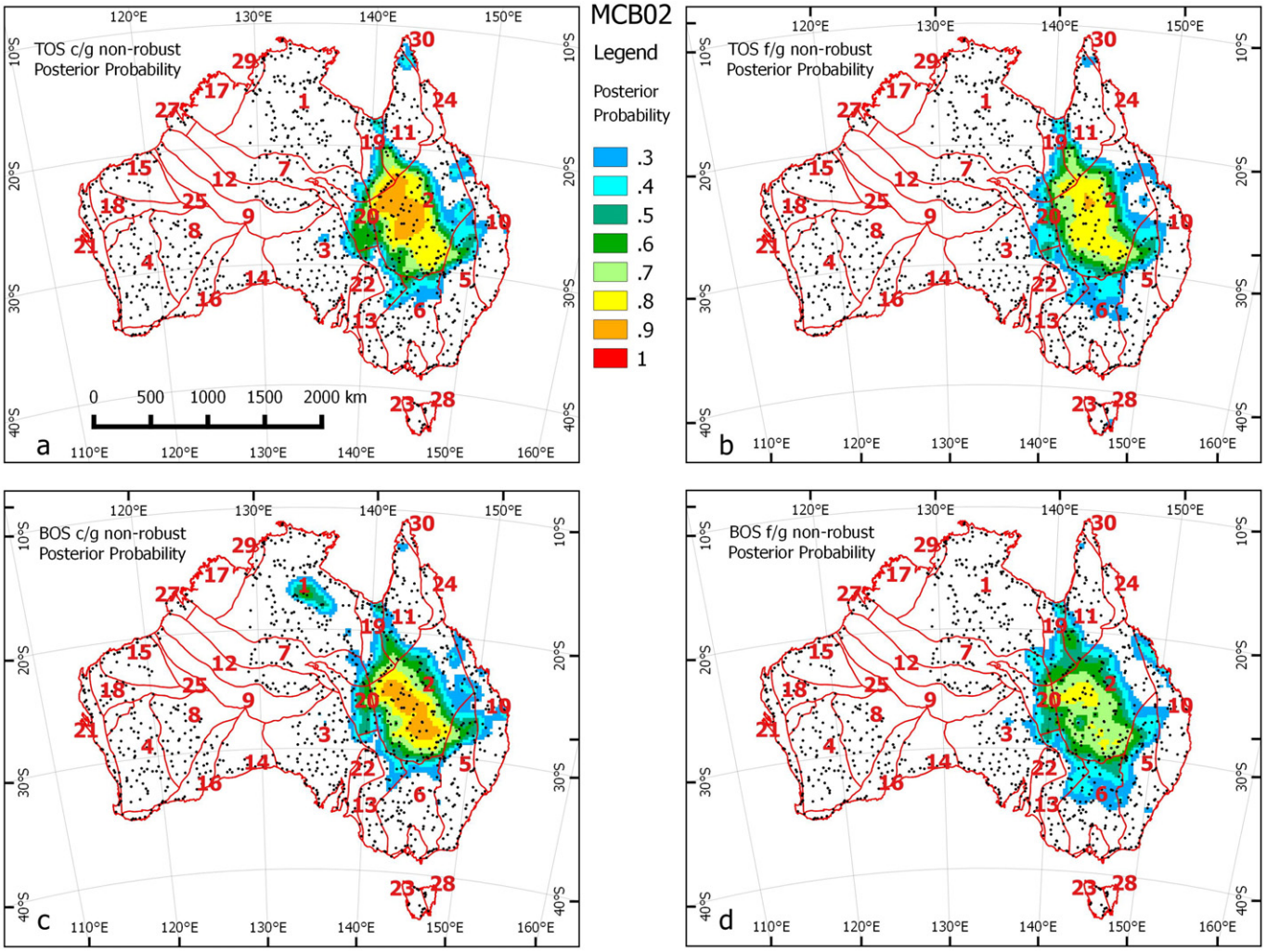


Fig. 6. Maps of the co-kriged PP predictions for major crustal block (MCB) 2 for the TOS c/g (a), TOS f/g (b), BOS c/g (c) and BOS f/g (d) sample types.

This is also geographically expressed in Figs. 8a and 9a mapping the accuracies for the TOS c/g and BOS f/g sample types for each of the MCBs. Both Figs. 8a and 9a show similar features, namely that the crustal blocks with the highest prediction accuracies are MCBs 14, 1, 2, 4, 8 and 3 regardless of sample type.

Measures of entropy (i.e., randomness; see Supplementary material), φ (local classification uncertainty), and 'most likely' MCB assignment for the TOS c/g and BOS f/g sample types are mapped spatially in Figs. 8b–d and 9b–d respectively.

In Fig. 8b, the lowest values of entropy (0.5 and below) occur mostly within MCBs 1, 2, 4, 7, 14, 15, 16 and 18. The highest entropies (1 and above) occur mainly in parts of MCBs 2, 5, 6, 8, 11, 13, 16, 22, 23 and 28. Fig. 9b shows the values of entropy for the BOS f/g sample media. The patterns of high and low measures of entropy are similar to the TOS c/g patterns of Fig. 8b. In both maps, the entropy values indicate greater variability in the northeast and southeast regions as well as in a central region and along some of the margins of the Yilgarn craton in the west. Conversely, the lowest entropy is displayed in the central north and south of the continent, as well as in the southwest and central east. The cause of this lower entropy (decreased disorder) may be the greater compositional stability or homogeneity of the (1) Eromanga, (2) Eucla, (3) Yilgarn and (4) Pilbara geological regions (see Fig. 1b), which are dominated, respectively (and in a simplified way) by (1) clastic sediments, (2) clastic and carbonate sediments, (3) granite and greenstone bedrock, and (4) felsic and mafic intrusive bedrock with sedimentary rocks and banded iron formations. MCBs 1 and 7 in northern central Australia also have low entropy in all four sample types, but the composition of surface materials and lithology of the basement is not as clearly homogeneous as for the other regions mentioned.

Local uncertainties (φ) for the TOS c/g and BOS f/g (Fig. 8c and 9c) are quite similar. Low uncertainties (0.4 and below) are associated with MCBs 1, 2, 3, 4, 7, 8, 9, 14, 15, 16 and 18. Elevated values of φ (0.8 and above) representing greater uncertainty are associated with MCBs 6/13/22, 2/11, 16 and 19/20. Fig. 9b (BOS f/g) shows patterns nearly identical to those in Fig. 8b, with larger areas of low entropy in MCBs 1, 4, 8 and 14. As for TOS c/g low values φ (0.4 and below) are associated with MCBs 1, 2, 3, 4, 7, 8, 9, 14, 16 and 18. The overall accuracy for sample type TOS c/g (0.47) is nearly equal to that of the BOS f/g sample type (0.48). Elevated values of φ (0.8 and above) are associated with MCBs 5/6/13/22, 2/11, 16, 18 and 19/20. The spatial patterns of φ (uncertainty) and of increased entropy (randomness) in Figs. 8 and 9 are very similar. The preliminary interpretation is suggested to be that increased uncertainty relates to a greater heterogeneity/mixture of the lithological source material in the above crustal blocks/regions. Regardless of interpretation, it is clear that greater uncertainty and disorder exist as spatially coherent regions for all four sample types.

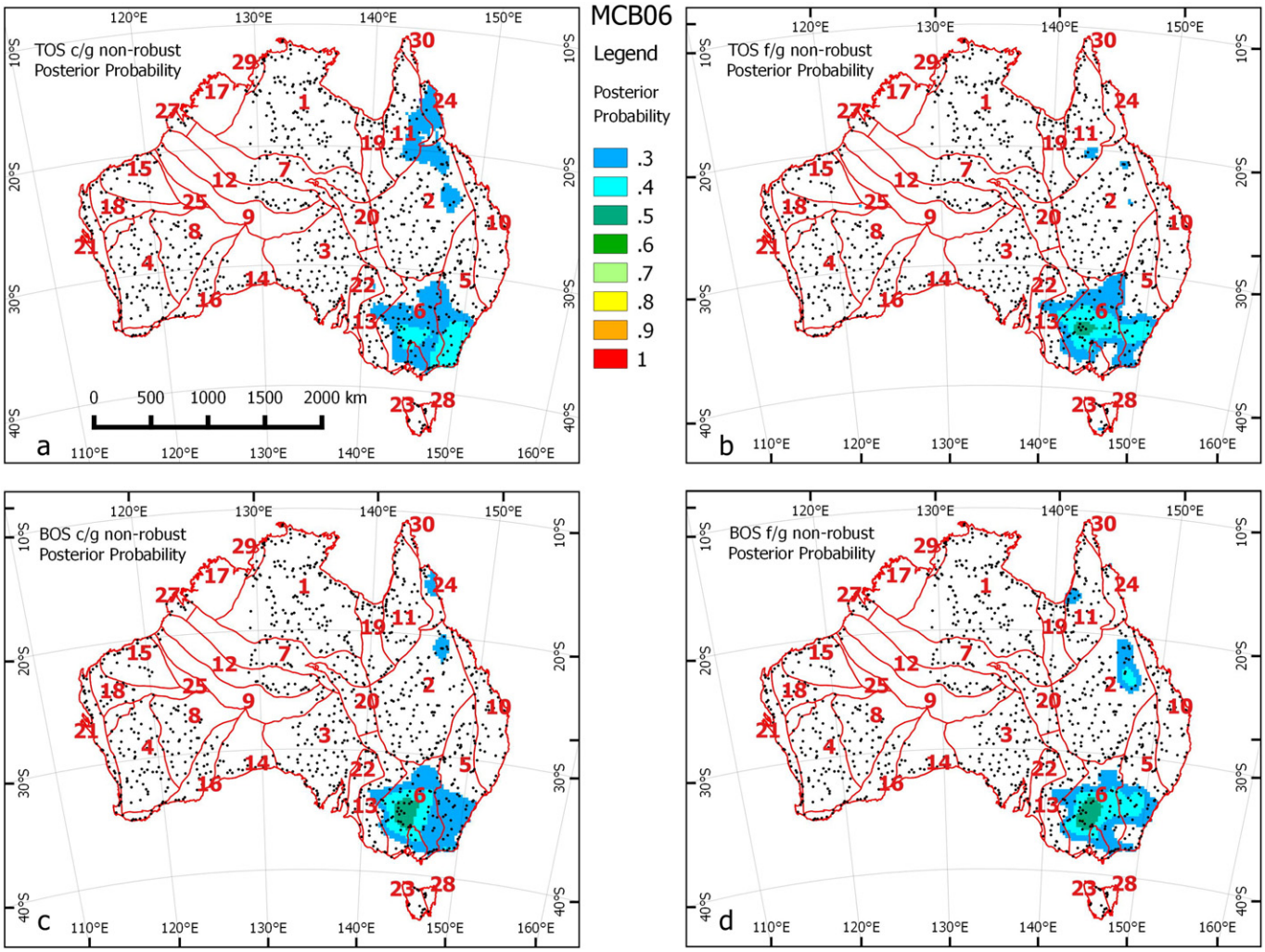


Fig. 7. Maps of the co-kriged PP predictions for major crustal block (MCB) 06 for the TOS c/g (a), TOS f/g (b), BOS c/g (c) and BOS f/g (d) sample types.

Figs. 8d and 9d show maps of the 'most likely' MCB based on the compositional similarity of a sample to a MCB from the PPs described above. Most of the cells match the MCB in which they are located, (e.g., MCBs 1, 2, 3, etc.). However, there are a few exceptions. MCBs 19 and 20 appear to be in large parts compositionally similar to MCB 2 for all four sample types. Here, transport of clastic sediments across these MCB boundaries in the Eromanga geological region (see Fig. 1b) can explain the compositional similarity. The southeastern part of MCB 6 appears to have a composition similar to that of MCB 24 to varying degrees in the four sample types. These regions being far removed from one another (eastern Victoria and northeastern Queensland), the compositional similarity cannot be explained by fluvial or eolian transport, but must instead be caused by coincidental similarity in the basement rocks (source materials for the sediments collected).

4. Conclusions

The results presented here identify two important features. First, the MCBs as defined through a combination of geophysical data and geological information are, for the most part, geochemically distinct. Despite long histories of weathering and deposition, the geochemical character of near-surface sediment atop these blocks persists and is by-and-large statistically unique, except where thick (several km) and widespread (100,000s km^2) younger sedimentary sequences cover the basement (e.g., Eromanga Basin over the Thomson basement, MCB 2). Second, the use of MAF analysis, which represents linear relationships of the data along with a measure of geospatial proximity, defines a metric from which the prediction of MCBs can be carried out with slightly greater prediction accuracy than PCA (Mueller and Grunsky, 2016). Use of MAFs is effective at describing linear combinations of elements that define distinct geochemical compositions.

The accuracy of prediction for the MCBs is highly variable. This variability is the result of compositional overlap between the MCBs. Those blocks with relatively low PPs show overlap with other crustal blocks. The off-diagonal elements of Supplementary material Table 3 show which blocks overlap with each other. The overlap of PPs can be due to a primary similarity of compositions between the MCBs or it can be due to compositional mixing through surficial processes. In most cases, the latter prevails in Australia (long periods for surface weathering and transport) especially with the sampling medium analysed here (fluvial sediment locally with eolian influence). Application of this method to outcrop rocks and drilled, fresher bedrock in areas of cover would likely reflect to a greater degree the inherent mineralogical/geochemical dissimilarities between crustal blocks.

Compositional overlap occurs in both contiguous blocks (overlap with adjacent MCBs) and, less commonly, in a spatially discontinuous

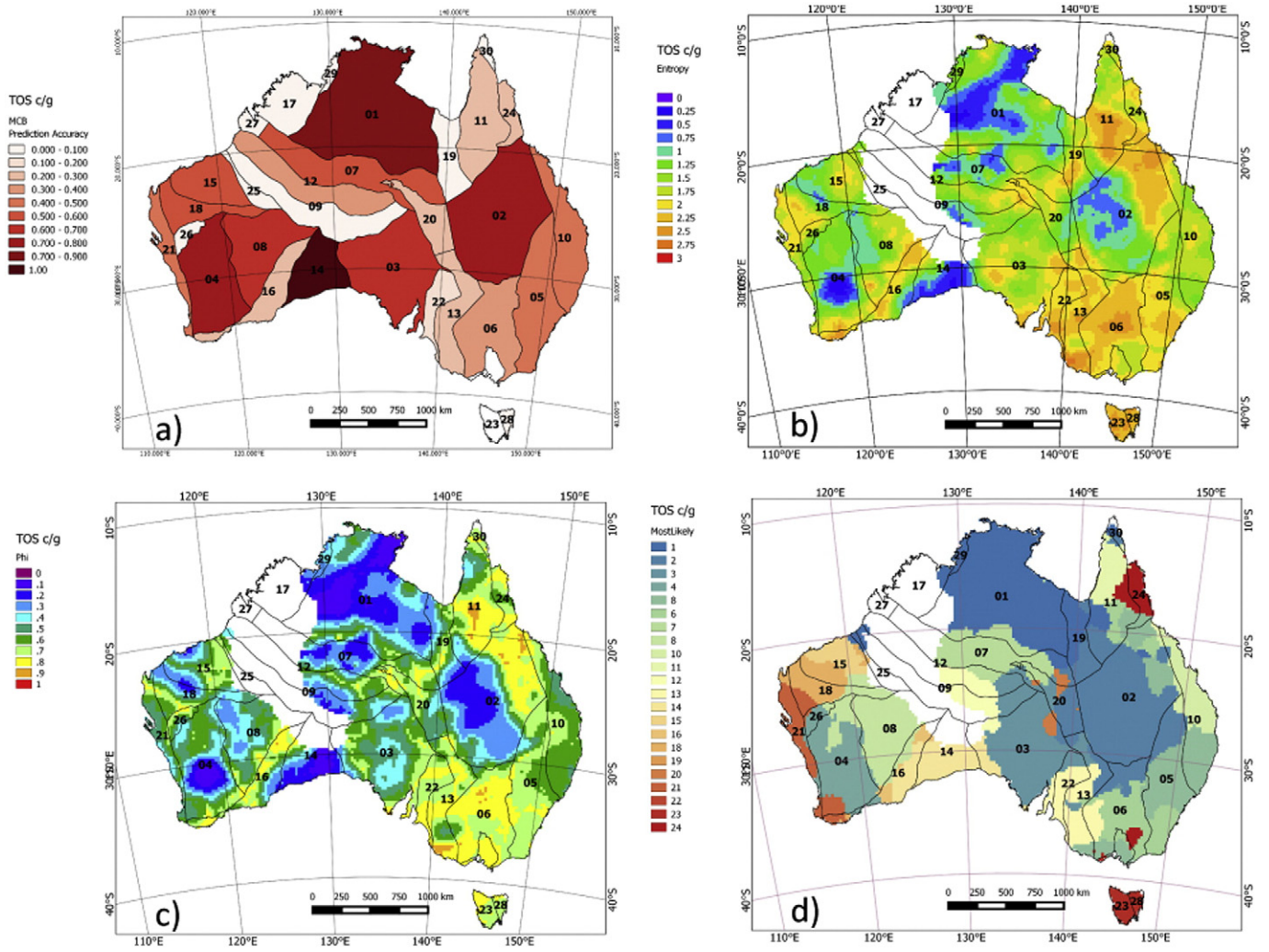


Fig. 8. Maps of average prediction accuracy (a), entropy (b), phi (c) and 'most likely' (d) for TOS c/g sample types.

form; the latter represents a similarity of compositions based on similar geochemistry (mineralogy). Thus, compositional overlap that is not geospatially associated can be interpreted as similar mineralogy and/or lithologies within these MCBs or the cover material (e.g., Cenozoic sedimentary or eolian cover). Compositional overlap between adjoining MCBs may be due to a combination of similar geochemistry of the source lithology and modification and mixing of the original compositions through mechanical and/or hydromorphic transport processes.

In areas where the PP is low (e.g., MCB 19), the overlap with other blocks (MCBs 1 and 2) suggests that surficial processes may mask the unique character of the block, particularly its potential to host base metal mineralisation.

The findings from this study demonstrate that the MCBs that form continental Australia are largely geochemically distinct and that despite long histories of uplift, burial, metamorphism, alteration through hydrothermal and groundwater interaction, with subsequent weathering, erosion, transport and deposition, these compositions persist throughout the geological history of the continent. The implications of this for mineral exploration, for instance, include the prospect of being able to geochemically recognise different parts of the architecture of Australia and use regionally meaningful geochemical background values against which anomalous samples, perhaps reflecting mineral deposits, can be identified.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gr.2017.02.011.

Acknowledgements

The National Geochemical Survey of Australia project would not have been possible without Commonwealth funding through the Onshore Energy Security Program, and Geoscience Australia appropriation. Collaboration with the geoscience agencies of all States and the Northern Territory was essential to the success of the NGSA project and is gratefully acknowledged. We also thank all the landowners for granting access to their properties for the purposes of sampling, and the laboratory staff for assistance with preparing and analysing the samples. We thank Michael Doublier and David Huston (Geoscience Australia) for their constructive internal reviews of the original manuscript, which was significantly improved as a result. Two anonymous journal referees and Editor Franco Pirajno are gratefully acknowledged for their efforts in critically reviewing and further enhancing this paper. PdC publishes with permission from the Chief Executive Officer, Geoscience Australia.

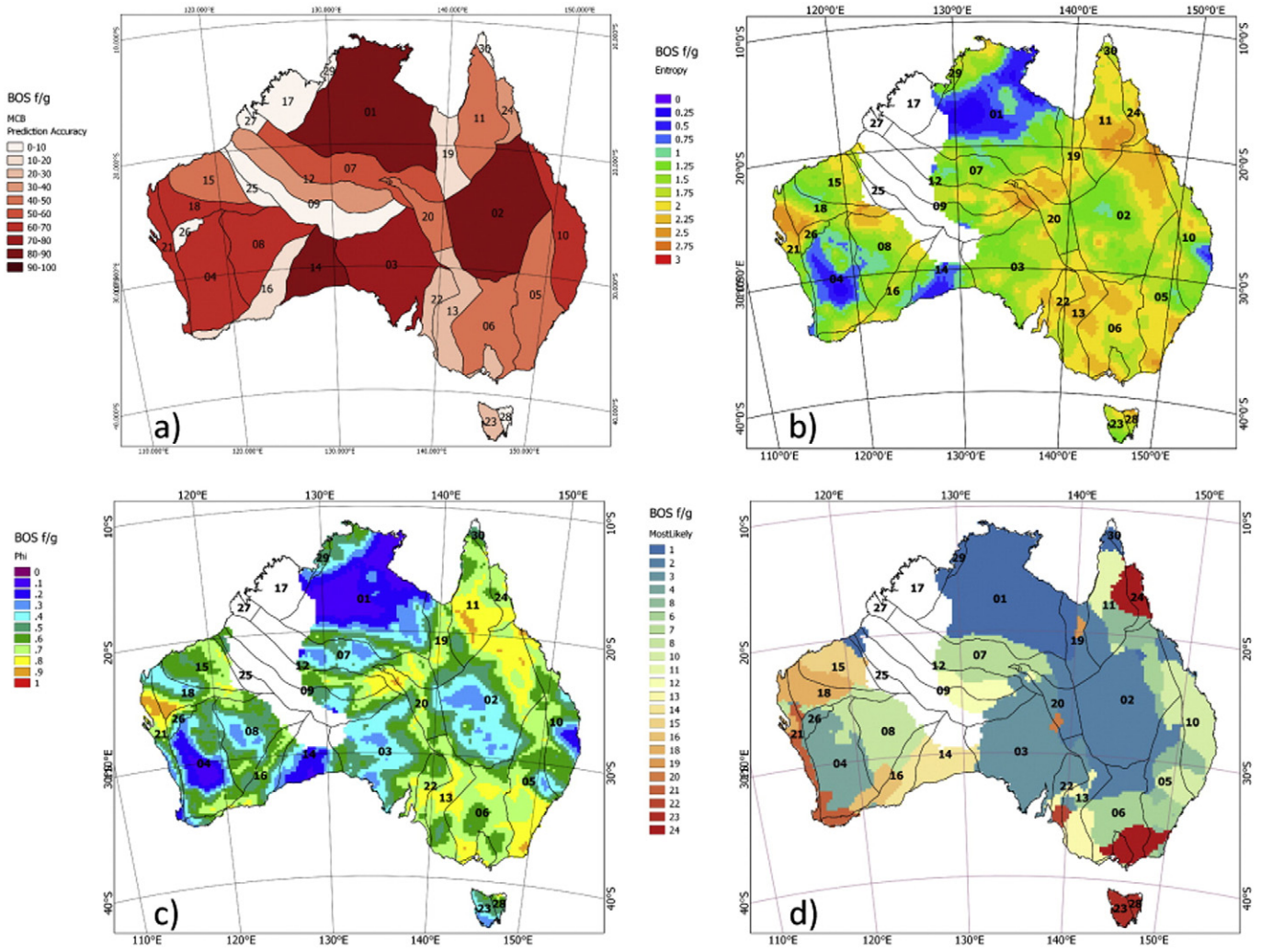


Fig. 9. Maps of average prediction accuracy (a), entropy (b), phi (c) and 'most likely' (d) for BOS f/g sample types.

References

Bain, J.H.C., Draper, J.J., 1997. Atlas of North Queensland geology 1:3 million scale. Australian Geological Survey Organisation Bulletin 240 95 p. Available at:. http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_a05f7892-9894-7506-e044-00144fdd4fa6/Atlas+of+North+Queensland+geology+1%3A3+million+scale.

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.

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. http://dx.doi.org/10.1016/j.chemolab.2004.06.006.

Caritat, P. de, Cooper, M., 2011. National Geochemical Survey of Australia: The Geochemical Atlas of Australia. Geoscience Australia Record 2011/20. 557 p. (2 Volumes). Available at:. http://www.ga.gov.au/metadata-gateway/metadata/record/gcat_71973.

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. http://dx.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. http://dx.doi.org/10.1016/j.apgeochem.2013.02.005.

Caritat, P. de, Lech, M.E., 2007. Thomson Region Geochemical Survey, Northwestern New South Wales. Cooperative Research Centre for Landscape Environments and Mineral Exploration Open File Report 145, 1021 p. + CD-ROM. Available at:. http://crcleme.org.au/Pubs/OFRSindex.html.

Caritat, P. de, Lech, M., Jaireth, S., Pyke, J., Lambert, I., 2005. Riverina geochemical survey - a national first. AusGeo News 78 Available at:. http://www.ga.gov.au/ausgeonews/ausgeonews200506/geochem.jsp.

Caritat, P. de, Lech, M.E., Jaireth, S., Pyke, J., Fisher, A., 2007. Riverina Region Geochemical Survey, Southern New South Wales and Northern Victoria. Cooperative Research Centre for Landscape Environments and Mineral Exploration Open File Report 234, 843 p. + CD-ROM. Available at:. http://crcleme.org.au/Pubs/OFRSindex.html.

Caritat, P. de, Lech, M.E., McPherson, A.A., 2008a. Geochemical mapping 'down under': selected results from pilot projects and strategy outline for the National Geochemical Survey of Australia. Geochemistry: Exploration, Environment, Analysis 8:301–312. http://dx.doi.org/10.1144/1467-7873/08-178.

Caritat, P. de, Lech, M.E., Kernich, A., 2008b. Gawler Region Geochemical Survey, South Australia. Cooperative Research Centre for Landscape Environments and Mineral Exploration Open File Report 211, 987 p. + CD-ROM. Available at:. http://crcleme.org.au/Pubs/OFRSindex.html.

Caritat, P. de, Cooper, M., Wilford, J., 2011. The pH of Australian soils: field results from a national survey. Soil Research 49:173–182. http://dx.doi.org/10.1071/SR10121.

Caritat, P. de, Reimann, C., NGSA Project Team, GEMAS Project Team, 2012. Comparing results from two continental geochemical surveys to world soil composition and deriving Predicted Empirical Global Soil (PEGS2) reference values. Earth and Planetary Science Letters 319:269–276. http://dx.doi.org/10.1016/j.epsl.2011.12.033.

Cohen, D.R., Silva-Santisteban, C.M., Rutherford, N.F., Garnett, D.L., Waldron, H.M., 1999. Comparison of vegetation and stream sediment geochemical patterns in northeastern New South Wales. Journal of Geochemical Exploration 66:469–489. http://dx.doi.org/10.1016/S0375-6742(99)00042-4.

Cornelius, M., Robertson, I.D.M., Cornelius, A.J., Morris, P.A., 2008. Geochemical mapping of the deeply weathered western Yilgarn Craton of Western Australia, using laterite geochemistry. Geochemistry: Exploration, Environment, Analysis 8:241–254. http://dx.doi.org/10.1144/1467-7873/08-172.

Cox, R., Curtis, R., 1977. The discovery of the Lady Loretta zinc-lead-silver deposit, northwest Queensland, Australia - a geochemical exploration case history. Journal of Geochemical Exploration 8:189–202. http://dx.doi.org/10.1016/0375-6742(77)90051-6.

Davy, R., Pirajno, F., Sanders, A.J., Morris, P.A., 1999. Regolith geochemical mapping as an adjunct to geological mapping and exploration; examples from three contiguous

Proterozoic basins in Western Australia. Journal of Geochemical Exploration 66: 37–53. http://dx.doi.org/10.1016/S0375-6742(99)00006-0.
De Vos, W., Tarvainen, T., 2006. Geochemical Atlas of Europe. Part 2 - Interpretation of Geochemical Maps, Additional Tables, Figures, Maps and Related Publications. Finland, Geological Survey of Finland, Espoo Available at:. http://weppi.gtk.fi/publ/foregsatlas/part2.php.
Drew, L.D., Grunsky, E.C., Sutphin, D.M., Woodruff, L.G., 2010. Multivariate analysis of the geochemistry and mineralogy of soils along two continental-scale transects in North America. Science of the Total Environment 409:218–227. http://dx.doi.org/10.1016/j.scitotenv.2010.08.004.
Grunsky, E.C., 1990. Laterite Geochemistry in the CSIRO-AGE Database for the Central Yilgarn Region (Barlee, Bencubbin, Corrigin, Hyden, Jackson, Kalgoorlie, Kellerberrin, Southern Cross Sheets). CSIRO Division of Exploration Geoscience Restricted Report 121R. Reprinted (1998) as Cooperative Research Centre for Landscape Environments and Mineral Exploration Open File Report 22 (2 Volumes). Available at:. http://crcleme.org.au/Pubs/OFRSindex.html.
Grunsky, E.C., 1991. Laterite Geochemistry in the CSIRO-AGE Database for the Albany-Fraser Region (Collie, Dumbleyung, Mt Barker, Pemberton Sheets). CSIRO Division of Exploration Geoscience Restricted Report 161R. Reprinted (1998) as Cooperative Research Centre for Landscape Environments and Mineral Exploration Open File Report 28 (2 Volumes + Dataset). Available at:. http://crcleme.org.au/Pubs/OFRSindex.html.
Grunsky, E.C., Drew, L.J., Smith, D.V., Sutphin, D.M., 2013a. Multivariate analysis of the United States portion of the North American Soil Geochemical Landscapes Project - a compositional approach. International Association for Mathematical Geosciences Meeting, August 2013, Madrid, Spain.
Grunsky, E.C., Drew, L.J., Woodruff, L.G., Friske, P.W.B., Sutphin, D.M., 2013b. Statistical variability of the geochemistry and mineralogy of soils in the maritime provinces of Canada and part of the northeast United States. Geochemistry: Exploration, Environment, Analysis 13:249–266. http://dx.doi.org/10.1144/geochem2012-138.
Grunsky, E.C., Mueller, U.A., Corrigan, D., 2014. A study of the lake sediment geochemistry of the Melville Peninsula using multivariate methods: applications for predictive geological mapping. Journal of Geochemical Exploration 141:15–41. http://dx.doi.org/10.1016/j.gexplo.2013.07.013.
Gustavsson, N., Bølviken, B., Smith, D.B., Severson, R.C., 2001. Geochemical landscapes of the conterminous United States - new map presentations for 22 elements. United States Geological Survey Professional Paper 1648 38 pp. Available at:. http://pubs.usgs.gov/pp/1648/report.pdf.
Jenny, H., 1941. Factors of Soil Formation – A System of Quantitative Pedology. McGraw Hill, New York 0-486-68128-9 281 pp.
Johnson, J., 2006. Onshore Energy Security Program underway. AusGeo News 84 Available at:. http://www.ga.gov.au/ausgeonews/ausgeonews200612/onshore.jsp.
Korsch, R.J., Doublier, M.P., 2015. Major Crustal Boundaries of Australia. Scale 1:2 500 000. second ed. Geoscience Australia, Canberra Available at:. http://www.ga.gov.au/metadata-gateway/metadata/record/83223.
Korsch, R.J., Doublier, M.P., 2016. Major crustal boundaries of Australia, and their significance in mineral systems targeting. Ore Geology Reviews 76:211–228. http://dx.doi.org/10.1016/j.oregeorev.2015.05.010.
Lech, M., Caritat, P. de, 2007a. Regional geochemical study paves way for national survey - geochemistry of near-surface regolith points to new resources. AusGeo News 86 Available at:. http://www.ga.gov.au/ausgeonews/ausgeonews200706/geochemical.jsp.
Lech, M.E., Caritat, P. de, 2007b. Baseline geochemical survey of the Riverina region of New South Wales and Victoria, Australia: concentrations and distributions of As, Ba, Br, Cd, Co, Cr, F, Ga, Mo, Sb, U and V compared to national and international guidelines. Geochemistry: Exploration, Environment, Analysis 7:233–247. http://dx.doi.org/10.1144/1467-7873/07-141.
McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117:3–52. http://dx.doi.org/10.1016/S0016-7061(03)00223-4.
Morris, P.A., Pirajno, F., Shevchenko, S., 2003. Proterozoic mineralization identified by integrated regional regolith geochemistry, geophysics and bedrock mapping in Western Australia. Geochemistry: Exploration, Environment, Analysis 3:13–28. http://dx.doi.org/10.1144/1467-787302-041.
Mueller, U.A., Grunsky, E.C., 2016. Multivariate spatial analysis of lake sediment geochemical data; Melville Peninsula, Nunavut, Canada. Applied Geochemistry (in press). 10.1016/j.apgeochem.2016.02.007.
Mueller, U., Lo, J., Caritat, P. de, Grunsky, E., 2014. Structural analysis of the National Geochemical Survey of Australia data. Mathematics of Planet Earth. Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences. Lectures Notes in Earth System Sciences. Springer:pp. 99–102 http://dx.doi.org/10.1007/978-3-642-32408-6_24.
Nakamura, A., Milligan, P.R., 2015. Total Magnetic Intensity (TMI) Colour Composite Image 2015. Geoscience Australia, Canberra Available at: http://www.ga.gov.au/metadata-gateway/metadata/record/82799/.
Ottesen, R.T., Bogen, J., Bølviken, B., Volden, T., 1989. Overbank sediment: a representative sample medium for regional geochemical sampling. Journal of Geochemical Exploration 32:257–277. http://dx.doi.org/10.1016/0375-6742(89)90061-7.
Raymond, O.L., 2012. Surface Geology of Australia Data Package 2012 Edition [Dataset]. Geoscience Australia, Canberra Available at:. https://www.ga.gov.au/products/servlet/controller?event=GEOCAT_DETAILS&catno=74855.
Ryall, W.R., Nicholas, T., 1979. Surface geochemical and biogeochemical expression of base-metal mineralization at Woodlawn, New South Wales, Australia. Journal of the Geological Society of Australia 26, 187–195.
Salminen, R., 2005. Geochemical Atlas of Europe. Part 1 - Background Information, Methodology and Maps. Geological Survey of Finland, Espoo, Finland Available at:. http://weppi.gtk.fi/publ/foregsatlas/index.php.
Shacklette, H.T., Boerngen, J.G., 1984. Element concentrations in soils and other surficial materials of the conterminous United States. United States Geological Survey Professional Paper 1270 Available at:. http://pubs.usgs.gov/pp/1270/pdf/PP1270_508.pdf.
UNCOVER, 2016. http://www.uncoverminerals.org.au/.
Walker, K.L., 1978. Involvement of government organisations in geochemical surveys in Australia. BMR Journal of Australian Geology and Geophysics 3:311–317 Available at:. http://www.ga.gov.au/metadata-gateway/metadata/record/80973/.