ASEG2024 HOBART

Leveraging Machine Learning and Geophysical Data for Automated Detection of Interior Structures of Cratons

Hojat Shirmard
EarthByte Group
School of Geosciences
The University of Sydney
Sydney, Australia
hojat.shirmard@sydney.edu.au

Ehsan Farahbakhsh
EarthByte Group
School of Geosciences
The University of Sydney
Sydney, Australia
e.farahbakhsh@sydney.edu.au

Karol Czarnota
[1]Research School of Earth Sciences,
Australian National University,
Canberra, ACT, Australia
[2]Geoscience Australia,
Canberra, ACT, Australia
karol.czarnota@ga.gov.au

R Dietmar Müller
EarthByte Group
School of Geosciences
The University of Sydney
Sydney, Australia
dietmar.muller@sydney.edu.au

SUMMARY

The internal structures and discontinuities of cratons hold considerable economic value due to their tendency for reactivation and different horizontal stress, serving as conduits for fluid flow and mineral deposition over time. Detecting these structures at various depths is critical for accurately mapping prospective zones of metallic mineralisation. This study demonstrates the effectiveness of integrating signal processing, feature extraction, and clustering on magnetic and gravity data for mapping the internal structures of the Gawler Craton, which has undergone rifting, sedimentation, extension, and orogenic processes. This combined approach results in precise internal structural mapping. Validated by three distinct metrics and geological maps, the resulting clustered maps can serve as foundational tools for further exploration and support decision-making in mineral exploration. Our findings indicate that most known metallic mineral occurrences, including all significant ones, are formed near the boundaries of these clusters. Therefore, mapping and targeting these boundaries can significantly reduce the search area for structurally controlled, extension-related mineral systems. Our proposed framework addresses the challenges of mapping hidden shallow and deep crustal structures, enhancing the capabilities of exploration geophysicists and geologists to investigate geological settings and the interiors of cratons. It provides a rapid, reliable, and cost-efficient method for generating geophysical features, which can be used as input to supervised prospectivity mapping workflows to identify favourable sites for mineralisation at any stage of an exploration program.

Keywords: Unsupervised machine learning, Clustering, Mineral exploration, Feature extraction, Craton structures

1. INTRODUCTION

Cratons are known to have internal structures inherited either from horizontal forces or from the closure of early oceans. Geodynamic models and new seismological observations suggest that mid-lithospheric discontinuities in various cratons may be remnants of deformation structures from craton formation following ancient ocean closure (Cooper & Miller, 2014). It has been demonstrated that the Gawler Craton exhibits the thickest lithospheric layer, contrasting with the thinner lithosphere found in other regions. Mineral compositions delineate two litho-chemical zones within the shallow and deep subcontinental lithospheric mantle, divided by a mid-lithosphere discontinuity (Sudholz et al., 2022).These internal structures of cratons are of significant economic interest because these craton sutures are known to be reactivated through time and represent pathways for fluid flow and mineralisation (Hagemann et al., 2016). Therefore, it is vitally important to identify structures at various depth intervals to delineate areas of particular interest in critical mineral exploration. While plate tectonics has dramatically enhanced our comprehension of processes occurring at plate boundaries, there still needs to be a greater understanding regarding intra-plate deformation dynamics. Despite the predominant occurrence of large-scale tectonic deformation along plate boundaries, it is recognised that substantial deformation can also take place within plate interiors (Gorczyk et al., 2012).

Tectonic interpretations in these old geological contexts are ambiguous because of the relative rarity of exposure, the density of data, and the propensity of fundamental geological information to be obscured during later tectonism (Betts & Giles, 2006). There have been only a few attempts to detect the internal structures of cratons quantitively. In a recent study, Motta et al. (2019) combined satellite gravity, magnetotelluric, aeromagnetic, and seismic data to identify shear zones in the Archean to Proterozoic Gawler Craton in South Australia. The findings suggest that the distribution of mineral provinces in the area is influenced by ancient tectonic events, indicating that earlier crustal boundaries and related structures play a significant role in shaping later geological processes. Furthermore, Heath et al. (2009) used an edge detection technique called “worming” on magnetic and gravity data to map lithospheric discontinuities in the Gawler Craton. However, this technique results in many small structures, most of which do not necessarily correspond to major internal craton boundaries. Therefore, the detailed mapping of internal cratonic terrane boundaries and associated structures as a function of depth remains a knowledge gap.

It is known that the Gawler Craton in South Australia was formed by the accretion of numerous terranes to its core (Betts & Giles, 2006). The deformation record observed in the terrane-scale shear zones within the Gawler Craton suggests a nuanced strain distribution pattern during tectonic events from the late Palaeoproterozoic to the early Mesoproterozoic. This pattern does not indicate a singular episode of terrane assembly or reworking; instead, it underscores the complexity of tectonic processes involving strain partitioning across different periods (Swain et al., 2005). We aim to develop a more robust, automated method for mapping these boundaries.

In a recent paper, Farahbakhsh et al. (2023) proposed a framework that integrates an advanced generative adversarial network with positive and unlabeled learning to address the common challenges associated with conventional mineral prospectivity mapping. They generated prospectivity maps for cobalt, chromium, and nickel within mafic–ultramafic intrusion-hosted mineralisation in the Gawler Craton, South Australia. Given the potential bias introduced by the sparse distribution of known mineral occurrences as training data in supervised methods and the uncertainty associated with generating negative samples, we opted to employ unsupervised methods. This motivates us to develop a new framework for processing geophysical data, which includes feature extraction, integration, and clustering, to map internal cratonic structures and their boundaries with an unsupervised approach. In our study, we use magnetic and gravity data to investigate the relationship between two distinct datasets of mineral occurrences with shallow and deep structures. It is imperative to address potential limitations when using wavelength as a depth proxy in magnetic and gravity signatures, particularly in regions such as the Gawler Craton, where sedimentary cover is prevalent. The presence of sediments introduces long-wavelength signatures that originate from shallow depths, which can confound interpretations based solely on wavelength analysis. To mitigate this issue, robust geological models incorporating detailed sedimentary thickness and distribution are essential. However, such comprehensive models are often lacking, necessitating acknowledgment of this limitation. Awareness of these complexities is crucial as they may significantly impact the interpretation and conclusions drawn from final results in geological and geophysical studies. With knowledge of this issue, we apply two clustering methods, i.e., K-means and SOM, to principal components extracted from the analysis of the input features for comparison. Principal component analysis plays a key role in our study in reducing the dimensionality of the feature space and capturing the most significant information (Shirmard et al., 2020).

2. Materials and Methods

The Total Magnetic Intensity (TMI) grid was created by consolidating open-file aeromagnetic surveys conducted within South Australia, with a resolution of 80 meters (m) per cell. Gravity information has been replaced by the South Australian regional gravity grids and data, 2023. All data from this study are obtained and available from the South Australian Resources Information Gateway[1](Laszlo F Katona & Reid, 2018). Numerous methods are dedicated to identifying the boundaries of potential field anomalies originating from geological formations. Amongst them, edge detection algorithms are critical in interpreting potential field data. In this study, we apply different signal processing filters, including TGA (total gradient amplitude), TDR (tilt angle), HTDR (total horizontal derivative), DX (horizontal derivatives in the x direction), DY (horizontal derivatives in the y direction), and HGA (horizontal gradient amplitude), and upward continuation. We use Gray Level Co-occurrence Matrix (GLCM) filters to extract textural features. GLCM filters are based on the statistical relationship between pairs of pixels in an image. They calculate the frequency of occurrence of pairs of pixel intensity values at a certain distance and angle within an image (Haralick et al., 1973). To reduce the dimensions of all extracted features and capture the most important information, we apply PCA since it identifies a set of linearly uncorrelated components, known as principal components, which represent projections of the original data onto principal axes or eigenvectors. These components are organised based on variance, with the first principal component exhibiting the highest variance and subsequent components in decreasing order of variance (Farahbakhsh et al., 2020).

This paper utilises two clustering methods: self-organising maps (SOMs) and k-means clustering. K-means clustering is a well-known and straightforward machine learning method that is only suitable for areas without any gaps in the coverage of the data sets. In contrast, self-organising maps (SOMs) are unsupervised neural network algorithms designed to analyse, visualise, and interpret multi-dimensional datasets (Carter-McAuslan & Farquharson, 2020). SOMs are artificial neural networks that transform complex, nonlinear statistical relationships among high-dimensional data into simple geometric relationships on a low-dimensional display. We use the elbow curve to determine the optimal number of clusters. The Silhouette Score was used as a criterion to determine the best hyperparameters. It quantifies the consistency of data points within their assigned clusters relative to neighbouring clusters, with higher scores indicating more cohesive and well-separated clusters. This metric provides valuable insight into the quality of clustering results, facilitating the selection of optimal hyperparameters for clustering algorithms (Rousseeuw, 1987).

2.1. Framework

The geophysical data compilation is followed by the pre-processing of magnetic and gravity datasets, which serve as the primary inputs for our proposed framework. We standardise these datasets to ensure consistent treatment across all features. As depicted in Fig. 1, we categorise deep and shallow structures by applying various upward continuation filters to explore the relationship of our deposit datasets with clustered maps. Then, our framework applies seven signal processing filters to four upward continuation maps generated from two different geophysical data types: magnetic and gravity data. In our study, we first focus on detecting shallow structures by creating four upward continuation maps from magnetic and gravity data at heights of 500, 1000, 2000, and 3000 m. To investigate deep anomalies, upward continuation filters are applied at 5, 10, 20, and 30 kilometres (km) heights. Subsequently, seven geophysical signal processing filters, including TGA (total gradient amplitude), TDR (tilt angle), HTDR (total horizontal derivative), DX (horizontal derivatives in x direction), DY (horizontal derivatives in y direction), and HGA (horizontal gradient amplitude), are applied to each upward continuation map of RTP for both shallow and deep analyses, using both magnetic and gravity data. As a result, 28 maps are prepared for analysing shallow and deep structures, respectively. For each of the 28 gravity and 28 magnetic maps, we extracted four features using the GLCM method, resulting in 112 features for magnetic data and an equal number for gravity data. Therefore, the total number of our extracted features is 448, which means 224 for deep and an equal number for shallow structures. To reduce the dimensions while benefiting from all information, we applied PCA since only a small subset of these components involves the majority of the data variance and provides the most relevant information about the data structure (Shirmard et al., 2020, 2022). We then apply

[1] https://map.sarig.sa.gov.au

two clustering methods, k-means and Self-Organizing Maps (SOM) for comparison. Additionally, we use all 224 features derived from the geophysical data sources as inputs for an integrated PCA analysis, combining both gravity and magnetic data. The PCA results are then used to create clustered maps using the abovementioned algorithms. Comparative analyses using tectonic and geological maps and the crustal structure of the Gawler Craton are performed to assess the distribution of mineral deposits in this geological setting. We have developed a framework based on open-source Python tools for implementing the approach that is designed to map hidden crustal structures within cratons, which can be applied universally.

3. Results and Discussion

A comparative analysis of the deep and shallow clustering results, generated using two clustering methods, reveals notable differences in their performance metrics. We used the silhouette score, which measures how similar an object is to its own cluster compared to other clusters, with higher scores indicating better-defined clusters. K-means clustering achieved a Silhouette score of 0.35 for deep-seated structures, indicating relatively good clustering quality. In contrast, the SOM method yielded poorer results, with a Silhouette score of -0.14, suggesting less effective clustering. Similarly, K-means outperformed SOM for the shallow structures, with a Silhouette score of 0.44 compared to SOM's Silhouette score of -0.02. These results indicate that K-means clustering provides more coherent and distinct clusters for both targets than the SOM method. The number of clusters estimated using the elbow curve, as the rate of change in the curve's slope decreases notably from around the eighth number onwards, with subsequent numbers demonstrating minimal variation. The search space for new deposits has been narrowed down to 70 per cent. According to the clustered maps created for deep structures (Fig. 2) and the location of mineral occurrences relative to the cluster boundaries, more than 80% of mineral occurrences are found within a distance less than 7 km of the detected boundaries (Fig. 3). It would be more accurate to consider the tonnage of mineral systems rather than simply the number of them when calculating this figure. However, tonnage information is frequently not readily available. Considering the tonnage instead of the number of deposits, this figure would likely be significantly higher than 80%. This is because the Olympic Dam, one of the largest world-class deposits, and the Cattlegrid, Prominent Hill, and Carrapateena deposits, significant in South Australia, are located within this buffer zone around boundaries. This indicates that many mineral deposits are formed within the interior of this craton along deep structures. Finally, promising areas for discovering mineral deposits, including Co, Cr, Ni, Au, Cu, Fe, and Mn, have been delineated by considering the boundaries within which mineral occurrences exist (Fig. 6). The prospective areas comprise only 30% of the study area, indicating that the search space for new mineral deposits has been effectively reduced by 70%. Additionally, the remaining 20% of mineral occurrences, which are of low significance, are located within 5 km of shallow structures but are relatively insignificant in grade and tonnage. If geophysical data from other types are incorporated into this framework, it would enable the detection of more structures and, consequently, target more deposits. The map generated by the proposed framework can serve as a crucial factor in guiding future exploration programs. This map can be further enhanced by conducting comprehensive geophysical surveys, such as high-resolution (with around the same or comparable line space of magnetic and gravity data) electromagnetic and magnetotelluric surveys with predefined profiles and appropriate spacing, and collecting consistent data. According to the known mineral occurrences, Au, Cu, Fe, and Mn mineralisations have been found along the margins of crustal-scale terranes, accurately detected at the boundaries in the clustered maps. Additionally, mafic-ultramafic intrusive-related cobalt, nickel, and chromium mineralisation have formed along extensional structures within the interior of the Gawler craton. Igneous and hydrothermal-related gold and copper also follow these trends in the eastern part of the study area. Furthermore, the presence of the Olympic Dam, Cattlegrid, Prominent Hill, and Carrapateena mineral systems in potential locations supports the high probability of economic mineralisation in these buffer zones around boundaries. Given that mineral deposits occur in diverse geological settings, we recommend using the extracted features and boundaries as inputs for the proposed framework by Farahbakhsh et al. (2023) and integrating these results with other geological information. This approach can reduce uncertainty, narrow the exploration space, and enhance the efficiency of mineral prospectivity mapping, particularly for buried deposits within cratons.

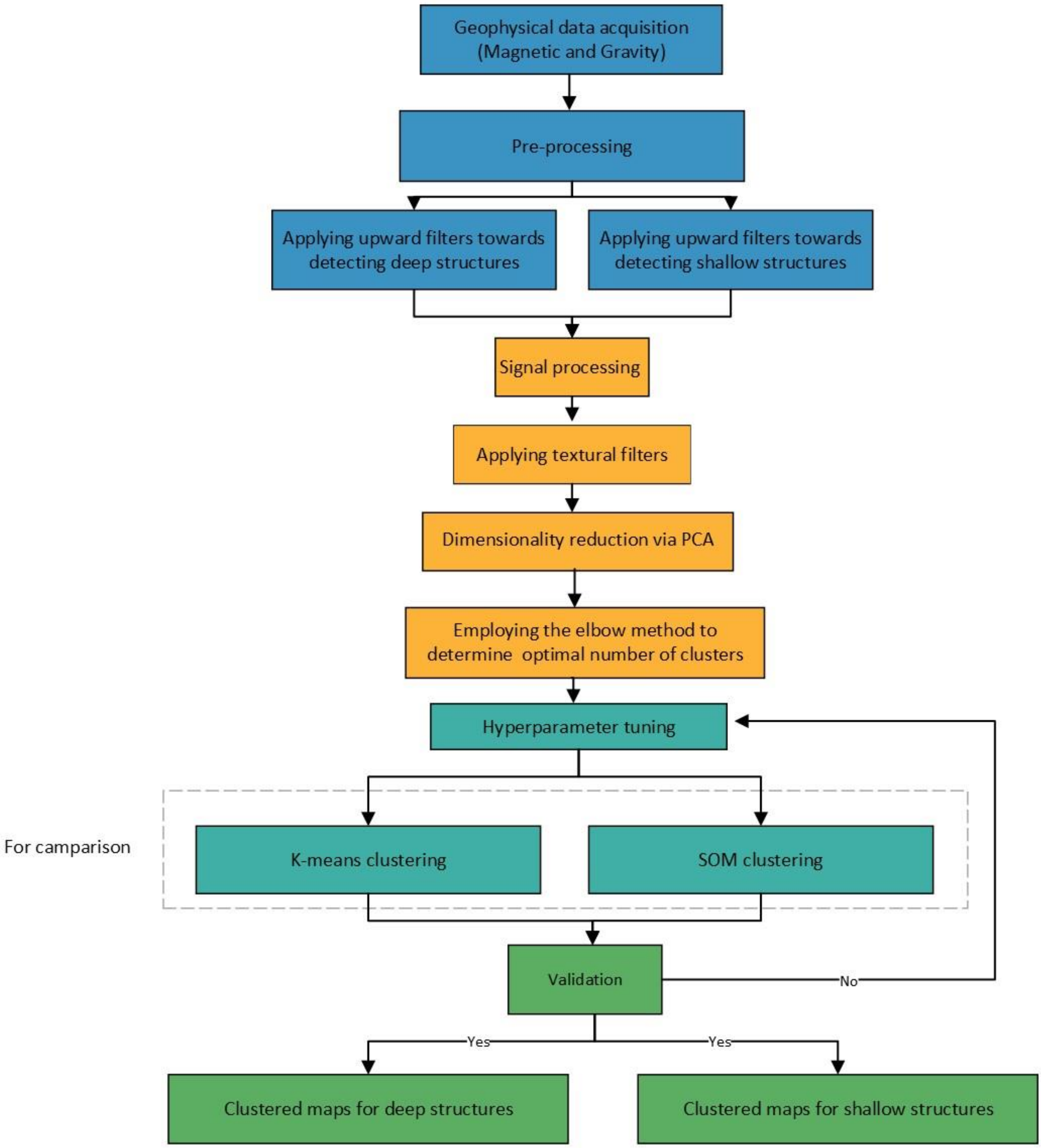

Fig 1. Proposed framework for generating deep (from 5 to 30 km) and shallow (from 500 to 3000 m) structural maps using geophysical data and two clustering methods.

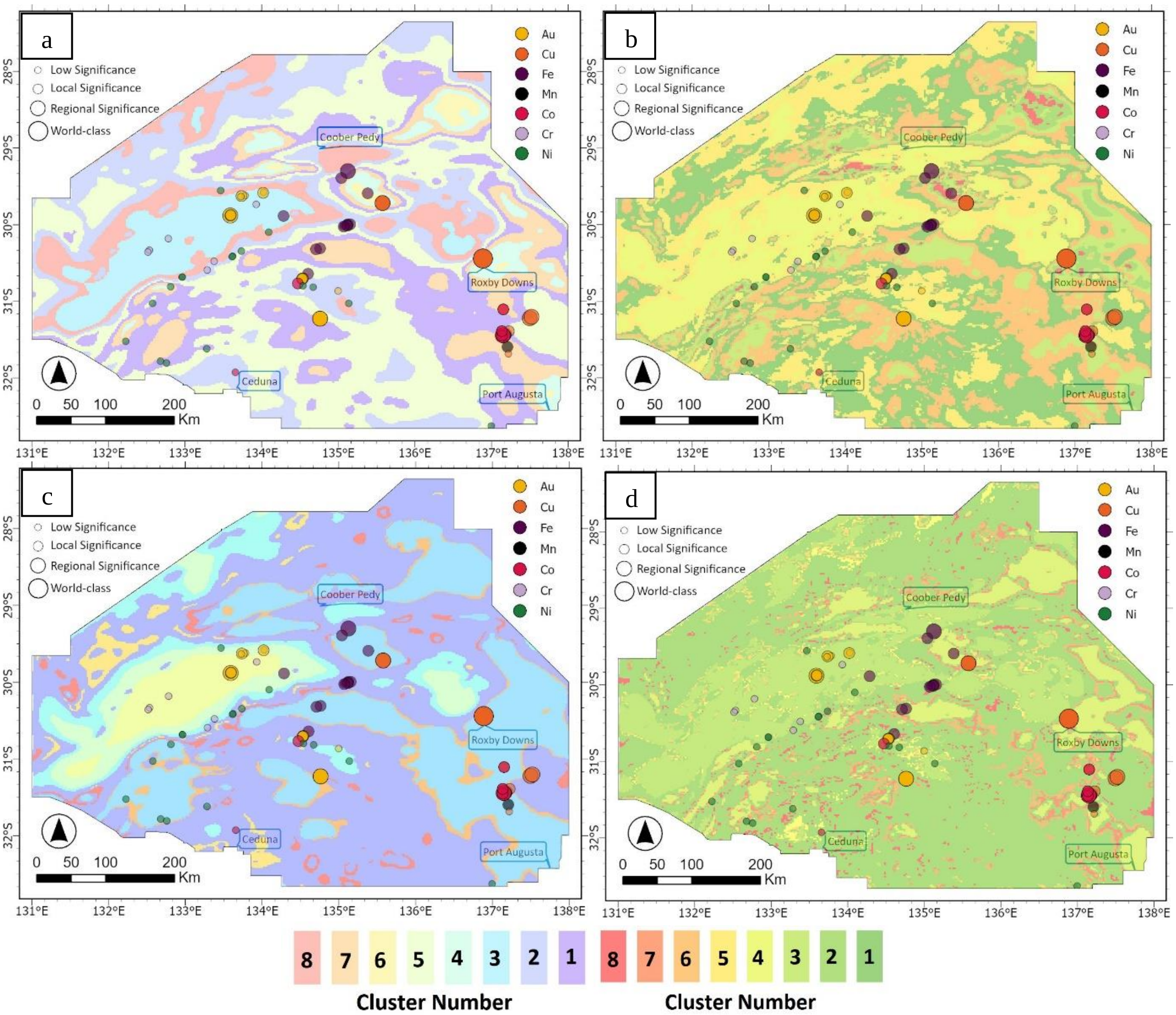

Fig 2. Clustered maps created by applying signal processing filters to magnetic and gravity data, extracting statistical features, and using PCA on 228 features for deep boundaries. a) deep clustered map generated with K-means and overlaying mineral occurrences; b) Shallow clustered map generated with K-means and overlaying mineral occurrences; c) Deep clustered map generated with SOM and overlaying mineral occurrence; d) Shallow clustered map generated with SOM and overlaying mineral occurrences.

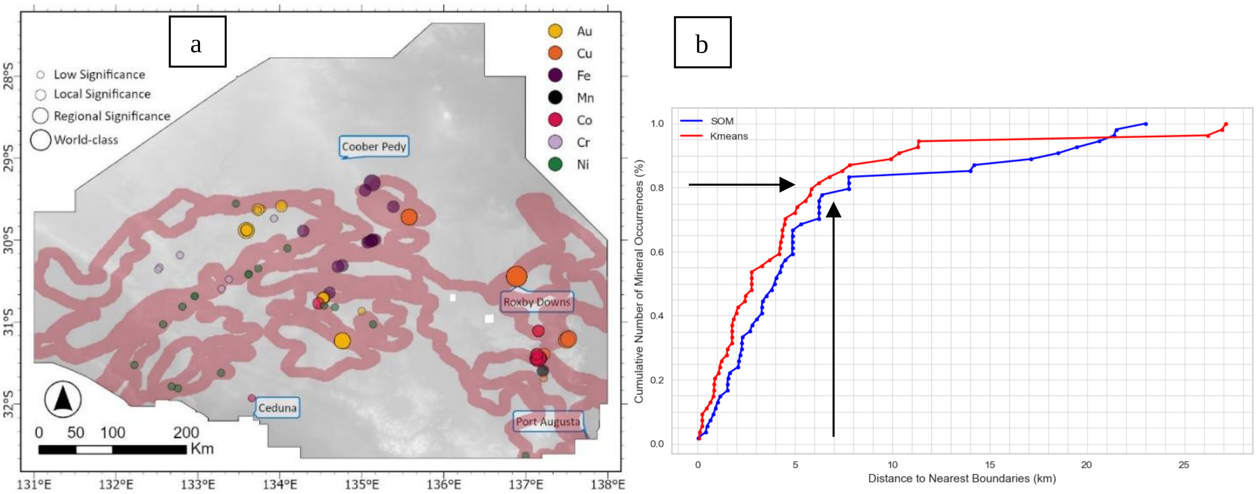

Fig 3. a) Promising areas within the 7 km buffer zone, identified through K-means clustering of deep structures, which encompass more than 80% of all mineral occurrences of cobalt (Co), chromium (Cr), nickel (Ni), gold (Au), copper (Cu), iron (Fe), and manganese (Mn), b) Distance from deep clustered maps boundaries generated with Kmeans and SOM mineral occurrences.

4. CONCLUSIONS

This study introduces an innovative exploration framework that integrates advanced signal processing techniques to extract numerous features, followed by dimensionality reduction to utilise all available information and applies unsupervised machine learning algorithms to address common challenges in exploring critical mineral deposits in covered regions with limited training data. The models developed for targeting critical minerals in the Gawler Craton demonstrated high stability and applicability. The resulting clustered maps revealed a strong spatial correlation between buffer zones around boundaries and known mineral occurrences, identifying several promising structurally controlled greenfield areas. This framework not only improves the accuracy of detecting internal craton structures but also accelerates the exploration process, potentially leading to significant economic benefits. Our study sets the stage for a comprehensive exploration of how an automated machine learning framework can revolutionise the detection of craton structures using geophysical data, ultimately aiding in discovering mineral deposits. By combining clustering methods such as SOM and K-means with magnetic and gravity data and employing various filters alongside statistical feature extraction, we effectively mapped the interior structures of the Gawler Craton. Based on known mineral occurrences in the study area, we demonstrate that this approach significantly reduces the search space (up to 70 per cent) for prospectivity mapping associated with internal craton structures, especially those covered by regolith. The unsupervised machine learning approach proves efficient for mapping potential mineralised areas, with the advantage of not requiring training data compared to supervised methods. Our findings indicate that over 80% of mineral occurrences, including all significant deposits, are within 7 km of the identified boundaries, while 20% of less significant mineral occurrences follow shallow structures. In conclusion, we have reduced the search space for exploration associated with internal crustal sutures in the Gawler Craton by more than two-thirds. The results may vary in other regions depending on the extent and geological characteristics of the target area, such as dominant rock types and the geophysical characteristics of the features.

REFERENCES

Betts, P., & Giles, D. (2006). The 1800–1100Ma tectonic evolution of Australia. *Precambrian Research*, *144*(1–2), 92–125. https://doi.org/10.1016/j.precamres.2005.11.006

Carter-McAuslan, A., & Farquharson, C. (2020). Application of SOMs and k-means clustering to geophysical mapping: Lessons learned. *SEG Technical Program Expanded Abstracts 2020*, 3843–3846. https://doi.org/10.1190/segam2020-w10-01.1

Cooper, C. M., & Miller, M. S. (2014). Craton formation: Internal structure inherited from closing of the early oceans. *Lithosphere*, *6*(1), 35–42. https://doi.org/10.1130/L321.1

Farahbakhsh, E., Chandra, R., Olierook, H. K. H., Scalzo, R., Clark, C., Reddy, S. M., & Müller, R. D. (2020). Computer vision-based framework for extracting tectonic lineaments from optical remote sensing data. *International Journal of Remote Sensing*, *41*(5), 1760–1787. https://doi.org/10.1080/01431161.2019.1674462

Farahbakhsh, E., Maughan, J., & Müller, R. D. (2023). Prospectivity modelling of critical mineral deposits using a generative adversarial network with oversampling and positive-unlabelled bagging. *Ore Geology Reviews*, *162*, 105665. https://doi.org/10.1016/j.oregeorev.2023.105665

Gorczyk, W., Hobbs, B., & Gerya, T. (2012). Initiation of Rayleigh–Taylor instabilities in intra-cratonic settings. *Tectonophysics*, *514–517*, 146–155. https://doi.org/10.1016/j.tecto.2011.10.016

Hagemann, S. G., Lisitsin, V. A., & Huston, D. L. (2016). Mineral system analysis: Quo vadis. *Ore Geology Reviews*, *76*, 504–522. https://doi.org/10.1016/j.oregeorev.2015.12.012

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-3*(6), 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Heath, P., Dhu, T., Reed, G., & Fairclough, M. (2009). Geophysical modelling of the Gawler Craton, SA - interpreting geophysics with geology. *ASEG Extended Abstracts*, *2009*(1), 1. https://doi.org/10.1071/ASEG2009ab127

Laszlo F Katona, T. W. and, & Reid, A. (2018). *Vectorisation of residual gravity and TMI data in the northern Gawler Craton: Implications for exploration targeting and GIS analysis*.

Motta, J. G., Betts, P. G., de Souza Filho, C. R., Thiel, S., Curtis, S., & Armit, R. J. (2019). Proxies for Basement Structure and Its Implications for Mesoproterozoic Metallogenic Provinces in the Gawler Craton. *Journal of Geophysical Research: Solid Earth*, *124*(3), 3088–3104. https://doi.org/10.1029/2018JB016829

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Shirmard, H., Farahbakhsh, E., Beiranvand Pour, A., Muslim, A. M., Müller, R. D., & Chandra, R. (2020). Integration of Selective Dimensionality Reduction Techniques for Mineral Exploration Using ASTER Satellite Data. *Remote Sensing*, *12*(8), 1261. https://doi.org/10.3390/rs12081261

Shirmard, H., Farahbakhsh, E., Heidari, E., Beiranvand Pour, A., Pradhan, B., Müller, D., & Chandra, R. (2022). A Comparative Study of Convolutional Neural Networks and Conventional Machine Learning Models for Lithological Mapping Using Remote Sensing Data. *Remote Sensing*, *14*(4), 819. https://doi.org/10.3390/rs14040819

Sudholz, Z. J., Yaxley, G. M., Jaques, A. L., Cooper, S. A., Czarnota, K., Taylor, W. R., Chen, J., & Knowles, B. M. (2022). Multi-Stage Evolution of the South Australian Craton: Petrological Constraints on the Architecture, Lithology, and Geochemistry of the Lithospheric Mantle. *Geochemistry, Geophysics, Geosystems*, *23*(11). https://doi.org/10.1029/2022GC010558

Swain, G., Hand, M., Teasdale, J., Rutherford, L., & Clark, C. (2005). Age constraints on terrane-scale shear zones in the Gawler Craton, southern Australia. *Precambrian Research*, *139*(3–4), 164–180. https://doi.org/10.1016/j.precamres.2005.06.007