https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

An explainable semi-supervised deep learning framework for mineral prospectivity mapping: DEEP-SEAM v1.0

Zijing Luo[1,2], Ehsan Farahbakhsh[2], Stephen Hore[3], R. Dietmar Müller[2]

[1]School of Resources and Environment, Henan Polytechnic University, Jiaozuo, 454003, Henan, China

[2]EarthByte Group, School of Geosciences, The University of Sydney, Sydney, Australia

[3]Geological Survey of South Australia, Adelaide, South Australia

Correspondence to: Zijing Luo (zijingluo@hpu.edu.cn)

Abstract. The clean energy transition demands a significant increase in exploration for critical minerals, particularly rare earth elements (REEs), beyond the well-explored surface deposits. Discovery rates have been declining for decades, escalating the need for new exploration methodologies. Deep learning (DL) utilizes multi-layer neural networks to automatically model high-level abstractions of data, extracting information relevant to the target task, thereby positioning itself as a potentially powerful tool for mineral prediction. But the non-linear and highly heterogeneous characteristics of complex exploration datasets paired with sparse, imbalanced training data and a lack of interpretability represent significant challenges. In this study, we have developed DEEP-SEAM v1.0, a novel explainable semi-supervised DL framework for prospectivity mapping of REE mineralisation in Northern Curnamona Province, South Australia. This framework proposes a comprehensive data preprocessing pipeline and introduces a semi-supervised anomaly detection DL model, termed the Deviation Network (DevNet). DevNet leverages a limited number of positive samples alongside a large number of unlabelled samples to effectively establish the mapping between multi-source exploration data and mineralisation probability. The results indicate that prospective mineralisation areas exhibit a strong spatial coupling with known REE deposits, with high-probability mineralisation areas primarily concentrated in faulted regions, felsic granites, and Mesoproterozoic strata. To address concerns about the poor interpretability of DL models, we incorporate a post-hoc model interpretation technique known as the SHapley Additive exPlanations method. The method facilitates an improved understanding of the decision-making mechanisms and logic underlying DevNet. By comparing DEEP-SEAM's decisions with our understanding of mineral systems, we not only enhance the model's transparency and interpretability but also strengthen the reliability and credibility of the predicted prospective mineralisation areas.

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

1. Introduction

Rare earth elements (REEs) comprise the lanthanides, ranging from lanthanum (La) to lutetium (Lu), and typically include yttrium (Y) and scandium (Sc) due to their similar chemical properties (Connelly et al., 2005). REEs are crucial in modern industries and indispensable components of defence systems, green technologies, and electronic applications (Dushyantha et al., 2020). For instance, rare earth alloys and permanent magnets are extensively used in renewable energy technologies, including electric vehicles, energy storage systems, solar panels, and wind turbines (Zhou et al., 2016). The growing demand for REEs will continue to be driven by advancements in both traditional applications and emerging technologies (Alonso et al., 2012; Dushyantha et al., 2020; Goodenough et al., 2018). Currently, the majority of global REE supply is derived from natural (primary) resources (Goodenough et al., 2018). REE mineral deposits are found in various types of rocks, including igneous, sedimentary, and metamorphic formations (Smith et al., 2016). The concentration and distribution of REEs in these deposits are influenced by various geological processes, ranging from deep magmatic activity to surface weathering (Jaireth et al., 2014; Smith et al., 2016). The interaction of different geological processes and complex formation mechanisms further complicates the exploration of REE deposits. Moreover, overburden layers introduce additional difficulties (Cheng, 2012). Overburden may dilute or alter signals that reflect mineralisation, and masking effects can reduce the accuracy of geological data, thereby further increasing the challenges of exploration (Cheng, 2012; Xiong et al., 2018).

Creating low-cost mineral prospectivity maps is essential for outlining potential ore bodies and guiding further data collection and discovery in mineral exploration. Mineral prospectivity mapping (MPM) is a computer-supported workflow that integrates multiple criteria across several stages to estimate where the target-type mineral deposits are likely to occur within a defined area (Singer, 1993; Zuo et al., 2021). MPM involves integrating information from various geoscientific datasets, including geological mapping, geochemical surveys, geophysical measurements, and satellite-based remote sensing (Brown et al., 2000; Zuo, 2020). These datasets reflect key mineralisation processes and mineral systems components, enabling the recognition of complex spatial distribution patterns of geological features associated with mineral deposits (Carranza, 2009; McCuaig and Hronsky, 2014). By employing data-driven, knowledge-driven, or hybrid models, the intrinsic relationships between these features and mineral deposits can be revealed, leading to the generation of mineral prospectivity maps (Zuo, 2020). In recent years, new theories and data analysis methods, such as various machine learning (ML) techniques, have been introduced into MPM workflows, becoming powerful tools for addressing mineral prediction challenges (Chen and Wu, 2017; Rodriguez-Galiano et al., 2015; Singer and Kouda, 1996; Zuo and Carranza, 2011). ML methods enable more efficient analysis and integration of geographical information from different sources, thereby optimising the exploration process, improving the success rate of mineral prospecting, and reducing costs (Farahbakhsh et al., 2023).

Traditional ML techniques are limited in processing raw natural data and typically rely on manually engineered processors to convert raw observations into suitable internal representations or model-ready feature vectors (LeCun et al., 2015). Deep learning (DL), a subfield of ML, offers a end-to-end representation learning approach that employs multiple levels of abstraction (LeCun et al., 2015; Nguyen et al., 2019). It directly extracts and transforms features through a cascade of multiple layers of non-linear processing units, excelling at uncovering complex structures in high-dimensional data (Bergen et al., 2019; LeCun et al., 2015). By utilising multiple layers of abstraction, DL can effectively capture the nonlinear relationships between multi-source heterogeneous exploration data and mineralisation probability (Sun et al., 2020; Zuo and Xu, 2023; Zuo et al., 2023). A common challenge in applying DL to MPM is insufficient positive samples (known mineral occurrences) to effectively generalise the model's attributes (Granek and Haber, 2015). Additionally, mineralisation processes vary significantly across geological periods, regions, and environments. Rare mineralisation events can cause supervised DL models to struggle when identifying mineralisation-related information (Cheng, 2007; Granek and Haber, 2015). These models often tend to favour the majority classes while overlooking the minority classes of interest, which can hinder their ability to detect rare but critical mineralisation signals (Farahbakhsh et al., 2023; Leevy et al., 2018; Li et al., 2021; Yang et al., 2022). Conversely, unsupervised DL models dispense with labeled training sets, circumventing the challenge of acquiring large-scale positive samples. As a result, they have been widely applied in MPM, offering a promising solution to these limitations (Xiong et al., 2018; Zuo et al., 2022). The core idea of popular unsupervised DL models for MPM is to learn latent embeddings of the data and then compare the reconstructed data with the original (Xiong et al., 2018). Anomalous samples, such as mineralized regions, are difficult to effectively encode into the latent representations learned from the abundant background data, leading to significant reconstruction errors (Ruff et al. 2021). These unsupervised DL models separate feature extraction and anomaly scoring processes, which may result in suboptimal data representations and low-quality reconstruction errors (Pang et al. 2019; Gao et al. 2021). A critical issue of their application in MPM is that unsupervised DL models lack prior knowledge regarding the spatial distribution of actual mineral deposits, which is essential to guide the optimization process. This absence of domain-specific knowledge makes the model prone to misclassifying noisy data or other non-mineralisation-related information as anomalies related to the target mineralisation task (Pang et al. 2023; Luo et al., 2024).

In a response to these challenges, semi-supervised learning emerges as a promising approach that integrates the benefits of supervised and unsupervised paradigms (Ruff et al. 2019). This is particularly beneficial for anomaly detection (AD), where the identification of rare events is crucial. Semi-supervised AD algorithms work by utilizing a limited number of labeled anomaly samples in conjunction with a large quantity of unlabelled samples to learn decision boundaries that isolate outliers from expected or normal data (Zhang et al. 2018). Studies have demonstrated that with as little as 1% labeled anomalies, appropriate semi-supervised methods can achieve effective performance (Pang et al. 2019). However, some semi-supervised AD algorithms assume that anomalous samples share similar attributes (Liu et al. 2002; 2003). In contrast, the formation processes of different types of mineral deposits vary, leading to differences in the exploration data features that reflect these distinct geological processes. Pang et al. (2019) proposed a semi-supervised neural architecture, known as Deviation Network (DevNet), which directly optimizes anomaly scores through an end-to-end network. This method leverages a few labeled outliers as prior information to accommodate anomalies exhibiting different abnormal behaviors. The approach has demonstrated excellent performance in fields such as network security intrusion detection, gear pitting fault detection, and geochemical anomaly identification (Alper et al. 2023; Zheng et al. 2024; Luo et al. 2024).

In addition, although DL models have demonstrated excellent performance in MPM, their complexity and “black-box” nature make it difficult to interpret the model’s predictions (Mou et al., 2023; Zuo et al., 2023). In the field of mineral prediction, which is characterized by high risks and substantial economic value, the issue of transparency is of particular concern. Decision-making in MPM requires not only accurate predictions but also a comprehensive understanding of how these predictions are derived (Hronsky and Kreuzer, 2019). This involves addressing questions such as, 'What factors drive the model's predictions?' Such insights enable geologists and experts from related fields to integrate their knowledge, validate the credibility of the predictions, and ultimately help mitigate exploration risks. Thus, improving the interpretability of DL models has become a key focus in research on mineral resource prediction.

Building on a foundational work of Luo et al. (2024), this study presents DEEP-SEAM v1.0, an explainable semi-supervised DL framework for MPM that integrates multi-source exploration data. By addressing the limitations inherent in the previous study that focused exclusively on single-source geochemical data, this approach seeks to overcome challenges associated with inadequate reflection of potential mineralisation information and insufficient model interpretability. This study targets REE mineralisation in the northern part of the Curnamona Province in South Australia, utilising a sizable open-access exploration dataset to generate heterogeneous evidential layers reflecting the target mineralisation, including geological data layers, geophysical and remote sensing layers, and geochemical data layers. Appropriate preprocessing methods are applied to each type of data to achieve effective data mining. To reduce interference from redundant information and noise in MPM, we employ Spearman correlation analysis to eliminate geophysical and remote sensing features with a high correlation. Additionally, robust principal component analysis

(RPCA) is applied to the geochemical data to extract key geochemical information relevant to geological processes while minimizing the influence of noise and outliers. We employ an end-to-end semi-supervised AD DevNet, leveraging a limited number of labelled positive samples alongside a large set of unlabelled samples to establish a coupling relationship between diverse evidential layers and the spatial distribution of mineral deposits. Additionally, the study incorporates a model interpretability technique, i.e., SHapley Additive exPlanations (SHAP), to explain the model's predictions, thereby enhancing transparency and increasing the credibility of the results.

2. Geological Setting and REE Mineralisation

The Curnamona Province is a large, near-circular geological terrane located in the northeastern part of South Australia (Preiss, 2000). It extends eastward from the Olary region, crossing into New South Wales (Newton et al., 2003). The province comprises metamorphosed sedimentary and igneous rocks from the Paleoproterozoic to Mesoproterozoic eras and is a region of significant mineral potential (Rutherford et al., 2007). It hosts numerous medium- to large-scale hard rock mineral systems, making it one of South Australia's most important base and precious metal-bearing geological formations (Robertson et al., 1998). Among these, the world-class Broken Hill lead-zinc-silver-gold deposit is the most renowned, along with significant copper-gold (molybdenum), uranium-REE, tin-tungsten, and silver-lead deposits (Robertson et al., 1998). Much of the Curnamona Province is overlain by younger Cambrian to Cenozoic sedimentary rocks, particularly in the central region (Williams et al., 2009). The geological evolution of the province has experienced several significant tectonic, magmatic, and metamorphic events, primarily during the Paleoproterozoic and Mesoproterozoic periods. These events have shaped the region's complex geological architecture and endowed it with substantial mineral resource potential.

The REE deposits (Table 1) in the study area are located within the predominantly Mesoproterozoic Mount Painter and Mount Babbage inliers of the Moolawatana Domain located in the north-western Curnamona Province (Fig. 1). The most recent geology map presentation for the study area is the Mount Painter region by Hore (2015). Although having limited research, the REE deposits are considered to be primarily of skarn type but may equally include either hydrothermal, metamorphic or metasomatic origins. For example, previous studies have demonstrated that the Palaeozoic to Early Cretaceous Radium Ridge Breccia (with a Mesoproterozoic precurser), within the southern Mount Painter Inlier, contain significant REE mineralisation represented by 360 Ma monazites of hydrothermal origin (Alley and Hore, 2022; Elburg et al., 2013; Drexel and Major, 1990; Hore et al., 2020a; Hore et al., 2020b; Robertson et al., 1998). Additionally, mid-Paleozoic localised high-grade metamorphism affected older rocks along the central eastern flank of the Mount Painter Inlier, generating monazite-rich biotite schists that exhibit remarkable REE enrichment and localised mobilization (Robertson et al., 1998). Another area of significant REE mineralisation is hosted in the metasomatised Mesoproterozoic Yerila Gneiss of the Mount Babbage Inlier (Wülser, 2009). There are also several smaller isolated pockets of REE-rich lithologies scattered throughout the inliers whose origins require further investigation.

The Mount Painter region also contains substantial primary and secondary uranium resources (Bogacz, 2006; Elburg et al., 2013; Robertson et al., 1998; Wülser et al., 2011), with mineralisation likely originating from uranium-, thorium-, and REE-rich granitic bodies, meta-volcanics and meta-sediments located within the inliers (Robertson et al., 1998). Also, the region holds potential for stratiform and volcanic-associated base metal mineralisation, including Cu, Sn and W, hosted within metamorphosed Meso- to Neo-Proterozoic sedimentary and volcanic rock units (Hore et al., 2020b; Robertson et al., 1998; Sheard et al., 1992). Many of the Mount Painter region granites are extremely enriched in heat producing elements (U, Th, K) with the region being notable for the heat generated by radiogenic decay of these elements. These High Heat Producing Granites (HHPG) initiated convectional sub-surface fluid migration and establish a region of high geothermal gradient metamorphism (Kovacs, 2005; Neumann et al., 2000; Sandiford et al., 2002). The region has experienced a prolonged history of these granitic intrusions and consequential localized radiogenic heating, spanning from the early Mesoproterozoic during the Delamerian Orogeny through to the Late Paleozoic, which led to a series and pulses of magmatic and hydrothermal activity. These geological processes have initiated sodic, potassic or chloritic alteration of the granites, locally altering initial lithologies to gneisses and schists, and also facilitated the formation of hydrothermal vein-type, breccia-hosted, and skarn-type mineral deposits of Au, Cu, U, Sn, and REEs.

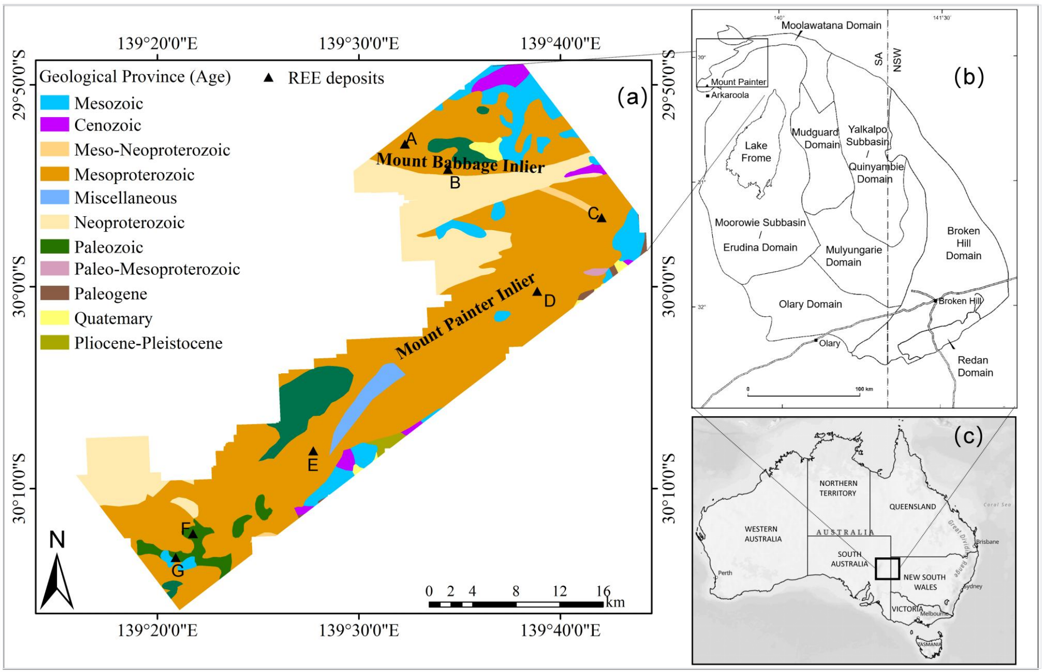


Figure 1: (a) Simplified stratigraphic map and REE mineral deposits in the Mount Painter and Mount Babbage inliers of the north-western Curnamona Province; (b) Geographic location of the study area, after Jagodzinski and Fricke (2010); (c) Outline map of Australia with state borders.

Table 1: Overview of REE mineral deposit characteristics (created by Steve Hore).

Location	Name	Commodity	Suggested mode of formation	Class code
A	Yerila gneiss	Co, REE, Th, U	Metamorphosed and metasomatised sediments	Occurrence
B	Yerila East	REE	Metamorphosed and metasomatised sediments	Occurrence
C	Moolawatana	Cu, REE	Hydrothermal, Pneumatolytic or contact metamorphism	Occurrence
D	Gunsight Prospect	Cu, Co, REE, U	Pneumatolytic and skarn-type	Prospect
E	Four Mile Creek	REE	Metamorphosed and metasomatised sediments	Occurrence
F	Armchair Prospect	U, Cu, REE, U_3O_8	Hydrothermal	Deposit
G	Mount Gee East Deposit	U, REE	Hydrothermal	Occurrence

3. Materials and Methods

3.1. Data Layers and Features

Study data are sourced from the South Australian Resources Information Gateway (https://map.sarig.sa.gov.au). A conceptual model can be established by analysing the geological setting and mineralisation type. Geological, geophysical (magnetic, gravity, radiometric), remote sensing, and geochemical data, along with a digital elevation model, are utilised in this study to enable a mappable representation of the conceptual mineralisation model. This multidimensional data integration reflects geological processes associated with mineralisation from various perspectives.

3.1.1. Geological Data Layers

The geological data layers (Table 2), including faults, granitic rocks, and stratigraphic information closely related to mineralisation, provide critical insights into mineralisation processes, vein distribution, fluid migration, and REE enrichment. Existing research indicates that the REE spatial patterns and elemental concentrations within deposits are largely controlled by the interplay of rock-forming processes (Dushyantha et al., 2020; Jaireth et al., 2014). Faults systems act as conduits for hydrothermal fluid migration and mineral precipitation (Curewitz and Karson, 1997). REE mineralisation associated with skarns is primarily formed from fluids derived from granitoid magmas, and understanding the distribution of granitic rocks aids in REE deposit exploration (Robertson et al., 1998). Australia's REE mineralisation formation was especially active during the Mesoproterozoic era (Spandler et al., 2020), with known REE occurrences in the study area showing a close spatial coupling with Mesoproterozoic strata.

3.1.2. Geophysical Data Layers

The Earth's magnetic field exhibits spatial variations that can be recorded in aeromagnetic surveys. Magnetic measurements highlight the differences in magnetisation levels within near-surface rocks, aiding geological mapping and the interpretation of concealed bedrock structures (Sharma, 1987). Intrusive igneous bodies, particularly carbonatite-alkalic intrusions, alkaline intrusions, and pegmatites, serve as significant sources of heat, materials, and fluids for REE mineralisation (Long et al., 2012). Magnetic survey techniques are regarded as exceptionally efficacious geophysical methodologies for identifying carbonatite-alkalic intrusive bodies, which commonly produce strong positive magnetic anomalies with circular to near-circular, crescent-shaped, or ring-like geometries (Simandl and Paradis, 2018; Thomas et al., 2016). The magnetic data (Table 2) utilised in this study include total magnetic intensity, processed using various methods (see Table 2 for details) to enhance the accuracy, resolution, and interpretability of the analysis. The total magnetic intensity grid is generated from aeromagnetic survey data with a resolution of 80 metres.

Gravity surveys rely on subtle variations in the gravitational field, which are related to differences in the density of subsurface materials (Reynolds, 2011). Alkaline igneous rock systems, such as carbonatites, are often significant parental rocks for skarn-type REE deposits (Goodenough et al., 2021). Gravity surveys are typically employed to study alkaline igneous rock systems, as these rocks generally exhibit high-density characteristics and can be associated with positive gravity anomalies and pronounce density contrasts compared to surrounding rocks (Drenth, 2014). The gravity data (Table 2) for the study area were subjected to Bouguer anomaly correction. Additionally, the dataset includes vertically derived gravity gradient data processed using first vertical derivatives and residual gravity data obtained by subtracting a 1000-metres upward-continued grid. All the gravity data are gridded at a resolution of 100 metres.

Radiometric measurement data play a critical role in REE mineral exploration, as they effectively display, interpret, and help understand underground geological structures and rock characteristics, particularly for rock types that exhibit significant radiation anomalies (Bustillo Revuelta, 2018). Radiometric surveys yield estimates of K, U, and Th concentrations using gamma-ray spectrometric analysis. A widespread positive correlation is observed between REEs

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

and radioactive elements in many known REE deposits (Thomas et al., 2016). This positive correlation primarily arises from the similar geochemical behaviours of U, Th, and REEs, leading to the enrichment of these radioactive elements in REE minerals such as bastnäsite, monazite, thorite, fluorocarbonate cerium, and zircon (Walters and Lusty, 2011). The radiometric data (Table 2) utilised in this study include grids of K (in %K), U (in equivalent ppm), and Th (in equivalent ppm), with these concentrations generated by merging data recorded from various airborne radiometric surveys. Additionally, the database includes ground dose estimates derived from the linear combination of K, Th, and U grids.

3.1.3. Remote Sensing Data Layers

The distribution of REEs in igneous and metamorphic rocks is significantly influenced by hydrothermal alteration and mineralisation processes (Lottermoser, 1992). The study of altered rocks is considered an effective tool for analysing the properties of mineralising fluids, conditions of mineralisation, and the evolution of hydrothermal alteration (Bedini, 2011). Remote sensing is an important method for mineral prospecting, utilising hyperspectral and multispectral data to detect hydrothermally altered minerals with diagnostic spectral absorption characteristics (Pour and Hashim, 2011). Metasomatism plays a crucial geochemical role in the concentration of REEs, U, Th, and other minerals, with Na-metasomatism exemplifying this process, typically manifesting as sodium-bearing minerals replacing primary igneous minerals (Khoshnoodi et al., 2016). In this context, remote sensing products can be used to explore the potential associations between the distribution of metasomatic rocks and the occurrence of REE mineralisation. The suite of ASTER geoscience products (Table 2) used in this study provides fundamental information about the dominant mineral components of the rocks and soils in the study area.

3.1.4. Digital Elevation Model

Digital elevation models (DEMs) are digital representations of Earth’s topography, created using topographic elevation data, effectively representing surface morphology in a numerical array format (Guth et al., 2021). By integrating elevation data with other exploration datasets, a more comprehensive understanding can be gained of how various geological processes and environmental conditions influence mineralisation. The DEM (Table 2) for the study area is based on a 9-second latitude and longitude grid and uses the Geocentric Datum of Australia 1994 coordinate system, with each grid cell representing the approximate elevation at its centre. The elevation errors in the DEM are closely related to terrain complexity: in low-relief areas, the standard error does not exceed 10 metres, while in complex highland regions, the standard error can reach approximately 60 metres.

3.1.5. Geochemical Data

Lithogeochemical data can reveal "in situ" geochemical processes related to the underlying geology (Grunsky and Caritat, 2019). By analysing variations and characteristics of the geochemical composition of rocks, particularly the distribution and enrichment of REEs, as well as potential rock types and hydrothermal influences, important insights for locating REE deposits can be obtained. The geochemical data (Table 2) used in this study encompasses 27 elemental variables, including trace elements, REEs, and major oxide components, which provide crucial information for identifying potential REE mineralisation zones.

Table 2: List of key data layers used for generating features in this study.

Data Type		Data Layer	Resolution
Vector	Polyline	Early Mesoproterozoic and younger (possibly some re-activated) faults	-
	Polygon	Felsic granites	
		Mesoproterozoic strata	
Raster	Magnetic	Total magnetic intensity (TMI)	80 m
		Variable reduction to pole (VRTP) of TMI	
		First vertical derivative of TMI VRTP	
		Automatic gain control filter of TMI VRTP	
		Tilt angle of TMI VRTP	
		Pseudo gravity of TMI VRTP	
		Analytic signal of TMI VRTP	
		Gradient of the magnetic strength in TMI VRTP	
		Amplitude of the total vector of TMI VRTP	
		Third order Cauchy of TMI VRTP	
		Vertical component of TMI VRTP	
		Vertical gradient of TMI VRTP	
	Gravity	Onshore Bouguer gravity anomaly	100 m
		First vertical derivative of onshore gravity	
		Residual onshore gravity	

	Radiometic	Terrestrial radiation dose	100 m
		Potassium concentration	
		Uranium concentration	
		Thorium concentration	
	Remote Sensing	Silica index	30 m
		Quartz index	
		Opaque index	
		MgOH group content	
		MgOH group composition	
		Kaolin group index	
		Gypsum Index	
		Green vegetation content	
		Ferrous iron index	
		Ferrous iron content in MgOH	
		Ferric oxide content	
		Ferric oxide composition	
		FeOH group content	
		AlOH group content	
		AlOH group composition	
	Elevation	Digital elevation model	250m
Tabular	Lithogeochemistry	Ag, As, Au, BaO, Bi, Co, Cr_2O_3, Cs, Cu, La, Mo, Nb, Ni, Pb, Rb, Sb, Sc, Sn, SrO, Ta, Th, U_3O_8, V, W, Y, Zn, Zr	-

This study incorporates vector, raster, and tabular data layers. The geological data layers are in vector format, comprising polylines and polygons. To effectively characterise the controlling influence of geological features such as Early Mesoproterozoic faults, felsic granites, and mesoproterozoic strata on mineralisation, these features are assigned values based on their spatial distribution. Grids covering the study area are first generated, and the distance from each grid point to the nearest geological feature boundary is calculated. The distance data is then normalized and inverted, so that the value assignment adheres to the “distance-weight” principle：the closer to the centre of the geological feature, the higher the assigned value; conversely, the farther away, the lower the value. The value range is defined from 0 to 1. The raster data comprises geophysical and remote sensing layers, as well as a DEM. The following steps are carried out to extract a series of features from these layers for various point sets, including known mineral occurrences and randomly selected samples. First, statistical features for each target point are extracted by calculating the mean and standard deviation of raster values within a circular buffer zone with approximately a one-kilometre (km) radius around each point. These statistics reflect the central tendency and dispersion of the area surrounding the target point. Second, texture features for each target point are computed using a square window of approximately one km side length around each point, based on the calculation of the dissimilarity and correlation of the grey-level co-occurrence matrix (GLCM). This approach characterises the spatial variability within the area surrounding each point. We extract elevation gradients in east-west and north-south directions from the DEM data. We then calculate the mean gradient values within a buffer zone of approximately one km radius around each point to reflect the terrain variation trends. Consequently, higher-dimensional raster features are generated for each point.

To address potential issues arising from feature redundancy and multicollinearity in downstream data mining and analysis, we calculate the Spearman correlation (Hauke and Kossowski, 2011) between the raster features. By comparing different correlation thresholds, removing features with correlations above the threshold, and evaluating the effect of the selected features on model performance, we determined 0.65 to be the optimal threshold. The lithogeochemical data are in tabular format, and to address issues associated with censored data, outliers, and missing values, we employ the Limit Replacement Method (VanTrump and Miesch, 1977), Tukey's boxplot method (Tukey, 1977), and the Random Forest algorithm (Breiman, 2001) for appropriate data processing (Luo et al., 2024). The detailed procedure can be found in Luo et al. (2024). Based on the location of each point, geochemical samples within approximately 1.5 km are extracted. The inverse square of the distance is used as a weight to calculate a weighted average of the geochemical concentrations of these samples, generating the corresponding elemental concentration for each point. In cases where corresponding geochemical samples are unavailable, the median value of the elements is adopted as the elemental value for the point. To address spurious correlations between variables in geochemical data caused by the closure effect, the Isometric Log-Ratio (ILR) transformation (Egozcue et al., 2003) is applied to "open" the geochemical composition data. Subsequently, Robust Principal Component Analysis (RPCA) (Filzmoser et al., 2009) is utilised to extract geochemical information related to geological processes while reducing the effects of noise and outliers. Nevertheless, a limitation of the ILR transformation is that the transformed components lose their direct relationship to the initial input variables. To enable interpretation of the resulting scores and loadings from the RPCA on ILR-transformed data, spatial transformations are employed to back-transform the principal component (PC) scores, facilitating the interpretation of relationships between each PC and the elements (Filzmoser et al., 2009) (Fig. A1). The back-transformed PC scores are then used as geochemical features.

3.2. Semi-Supervised Deep Learning Framework

The DL-based framework developed in this study, DEEP-SEAM (Fig. 2), encompasses data cleaning and preprocessing, redundancy filtering, and noise reduction for multi-source exploration datasets, followed by deep data mining to identify potentially favourable mineralisation areas within the study area. Additionally, the framework employs a post-hoc interpretability technique to analyse the contributions of different exploration data features to account for the obtained mineralisation probabilities. This study provides corresponding processing solutions for each key step and produces reliable and credible results for mineral prospectivity delineation. First, data quality has a significant impact on the performance of DL models. Therefore, this study collects multi-source exploration data and processes it effectively to reflect geological processes and subsurface geological structures associated with mineralisation. Based on this, a set of random samples is created in the study area and prepared for division into training and test sets. From the random sample set, those located over 5 km from known mineral occurrences are selected as negative samples. 70% of the negative and positive samples are randomly chosen to form the training set. The remaining 30% of the samples constituted the testing set. Subsequently, a DevNet, a cutting-edge semi-supervised DL approach, is utilised to learn from the training data. It is important to note that the DevNet used in the framework can function effectively under significant class imbalance between positive and negative samples.

Moreover, the selection of negative samples is not highly sensitive; even if the chosen negative samples contain potential positive samples, the model still performs well, as it learns the general patterns of the negative sample set. The model's performance is evaluated using the testing set to determine the optimal DL model structure. Then, a set of regular samples is generated within the study area to create a prediction set, and the trained DL model is employed to generate a mineral prospectivity map for the prediction set. This framework is able to effectively identify prospectivity areas associated with mineralisation under the guidance of limited positive samples, while avoiding false targets generated by statistical noise or interfering geological processes. This study combines a feature attribution method to analyse the internal mechanisms of model predictions, thereby overcoming the shortcomings of DL approaches regarding interpretability and the resultant credibility issues of predictive outcomes. This analysis is further integrated with existing geological knowledge for a comprehensive evaluation, thereby enhancing the reliability of the predictions.

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

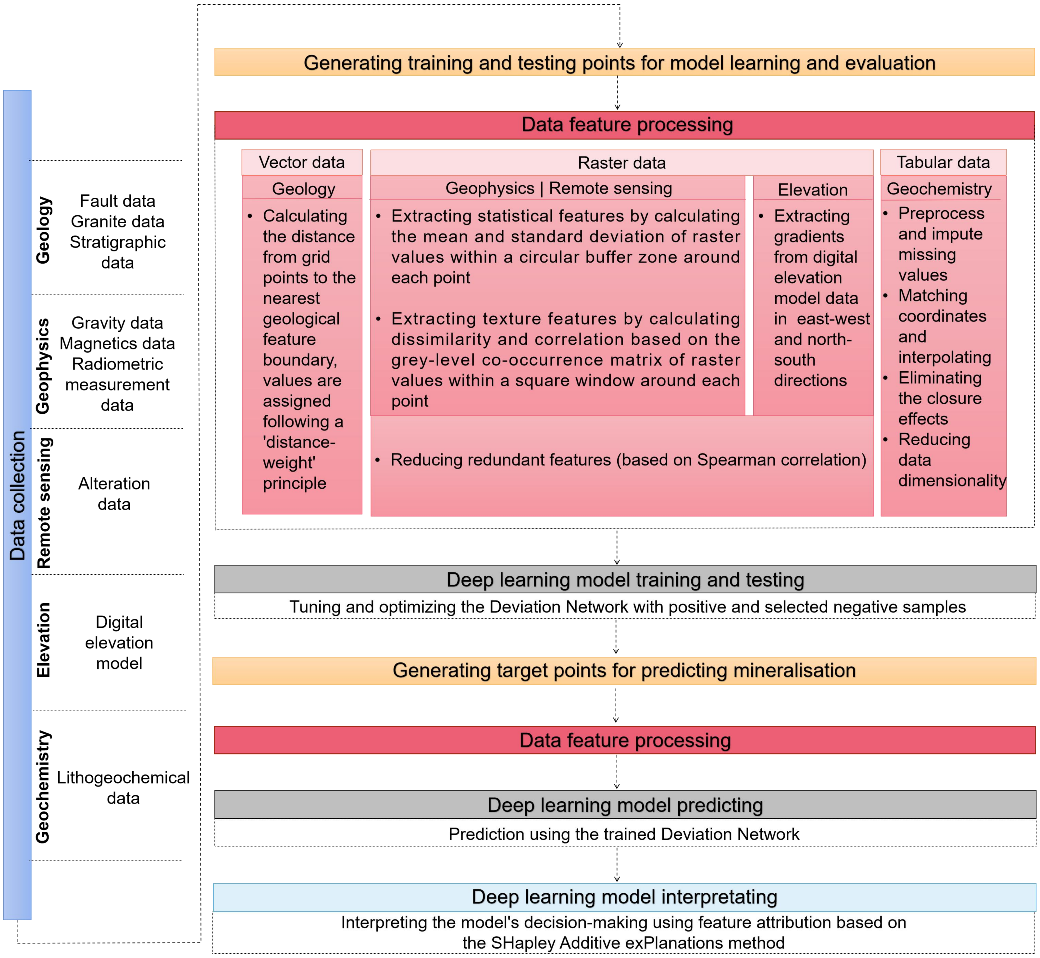


Figure 2: Schematic diagram of the DEEP-SEAM framework for MPM.

3.2.1. Deviation Network

DevNet, proposed by Pang et al. (2019), employs a Gaussian prior and utilises a Z-score-based deviation loss function to facilitate direct anomaly scores optimisation through an end-to-end neural anomaly score learner (Fig. 3). DevNet comprises three primary modules, including an anomaly scoring network, a reference score generator, and a deviation loss.

DevNet differs from the traditional two-step AD methods based on reconstruction error. DevNet establishes a nonlinear functional relationship between the input data and anomaly measurements through an end-to-end anomaly scoring module. Consider a dataset $x=\{x_1,x_2,...,x_N,x_{N+1},x_{N+2},...,x_{N+K}\}$, $x_i \in \mathrm{R}^D$, where $U=\{x_1,x_2,..,x_N\}$ represents the unlabeled data, and $M=\{x_{N+1},x_{N+2},...,x_{N+k}\}$ with $M \ll N$ denotes a small dataset of positive samples, which provides prior knowledge of anomaly patterns for model training. In this study, we utilise a feature representation learner $\psi(\cdot;\Theta_t)$,where the dot $(\cdot)$ represents the input to the function, which is constructed with multiple hidden layers to map the input data x into an intermediate representation space $Q \in \mathrm{R}^L$ (Eq. 1), where L denotes the dimensionality of the space Q resides in. Subsequently, we employ the developed anomaly score learner $\eta(\cdot;\Theta_s):Q \rightarrow \mathrm{R}$ to compute anomaly values from the derived intermediate feature representation (Eq. 2). Combining the above components, the complete anomaly scoring module $\phi(\cdot;\Theta)$ is then formulated as shown in Equation 3.

$$q=\psi(x;\Theta_t), \text{ where } q \in Q, \tag{1}$$

$$\eta(q;\Theta_s)=\sum_{k=1}^{L} \omega_k^o q_k+\omega_{L+1}^o, \text{ where } \Theta_s = \{w^o\},\ w^o \text{ is a weight matrice, } w_{L+1}^o \text{ is a bias term,} \tag{2}$$

$$\phi(x;\Theta)=\eta(\psi(x;\Theta_t);\Theta_s), \tag{3}$$

DevNet integrates a specialized a module designated as a reference score generator, architected to support anomaly scores $\phi(x;\Theta)$ learning processes, and to compute a scalar value referred to as the reference score $\mu_\mathrm{R} \in \mathrm{R}$. The reference score represents the average anomaly scores $\{r_1,r_2,...,r_l\}$ across a randomly chosen subsets of negative samples R. Two primary strategies exist for producing μ_R: data-driven and prior-driven methods. In this study, we adopt the prior-driven approach with μ_R computed according to the Gaussian prior probability.

$$r_1,r_2,\ldots,r_l \sim N(\mu,\sigma^2), \tag{4}$$

$$\mu_\mathrm{R}=\frac{1}{l}\sum_{i=1}^{l} r_i, \tag{5}$$

Where each r_i is derived from $N(\mu,\sigma^2)$ and corresponds to the anomaly score of a randomly selected negative sample. The Gaussian prior-based scores μ_R are used to optimise the anomaly scoring network $\phi(x;\Theta)$. The deviation is specified as a Z-score as follows:

$$dev(x)=\frac{\phi(x;\Theta)-\mu_\mathrm{R}}{\sigma_\mathrm{R}}, \tag{6}$$

Where σ_R indicates the standard deviation of the prior-based anomaly score set. The calculated deviation is subsequently incorporated into the contrastive loss to formulate the deviation loss:

$$L(\phi(x;\Theta),\mu_\mathrm{R},\sigma_\mathrm{R})=(1-y)|dev(x)|+y max(0,a-dev(x)), \tag{7}$$

We assign $y=1$ for an positive (anomalous) sample x, and $y=0$ for a negative (normal) sample, while a corresponds to the Z-score confidence interval parameter.

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

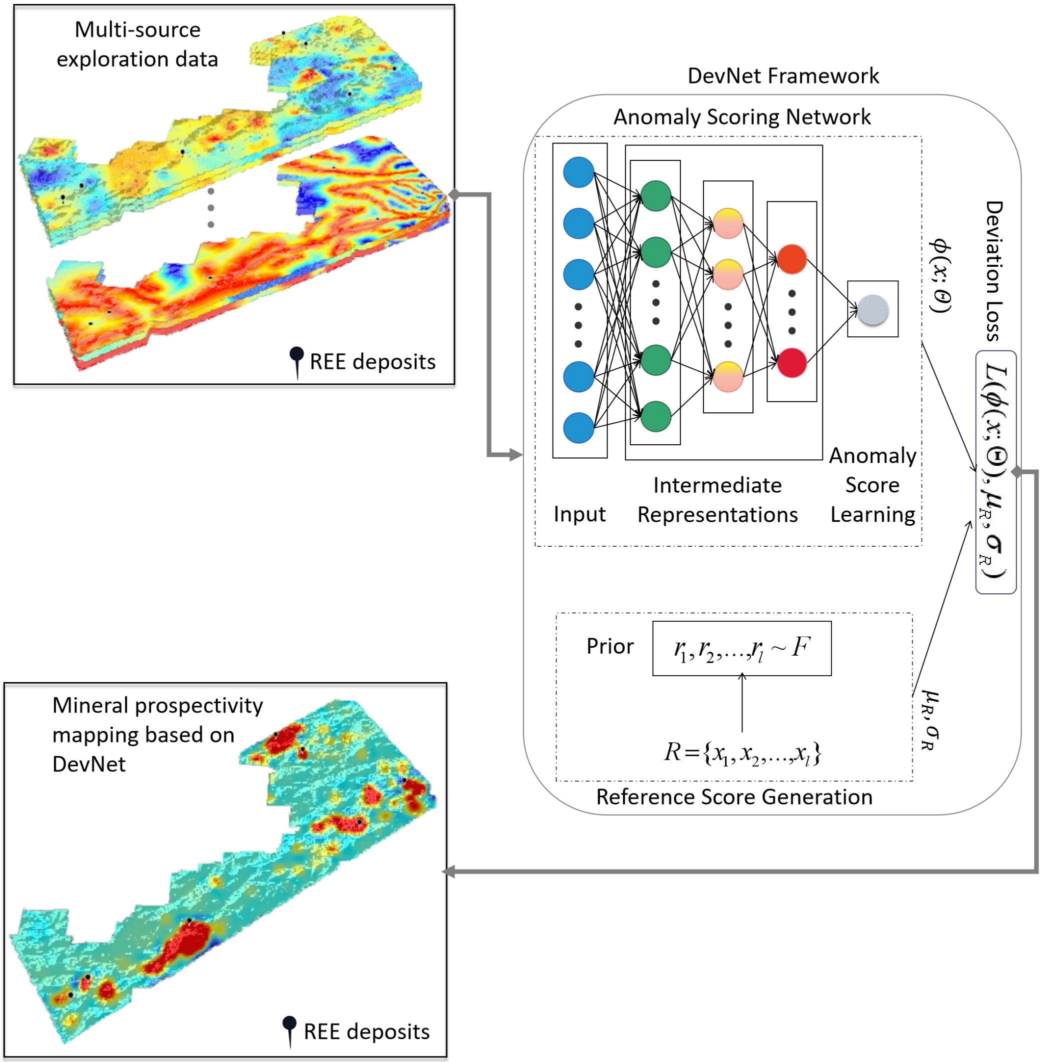


Figure 3: Architecture of the DevNet for probability score calculation.

3.2.2. Post-Hoc Interpretability

Additive feature attribution methods represent a significant category within the field of explainable DL. These methods explain individual predictions of DL models by representing the model output as a sum of contributions from individual input features. The SHAP framework (Lundberg and Lee, 2017), a classical additive feature attribution approach, integrates previous explanation methods such as local interpretable model-agnostic explanations (LIME; Ribeiro et al., 2016) and DL important features (DeepLIFT; Shrikumar et al., 2017), while incorporating Shapley values originating from game theory, thus providing both local and global perspectives for model interpretation (Fig. 4). Furthermore, compared to the aforementioned methods, SHAP is recognised for possessing three ideal properties: local accuracy—where the explainer exactly matches the original model output when approximating it for a specific input; missingness—where features absent from the original input must have no contribution; and consistency—where if the model is revised to increase the influence of a particular feature, the attribution assigned to that feature must not decrease, regardless of other features. These three properties ensure the reliability of the explanations.

The Shapley value represents an equitable distribution mechanism for apportioning benefits among participants according to their individual contributions, originating from economist Lloyd Shapley's foundational work. This allocation principle can be mathematically expressed as:

$$\Phi(x_i)=\sum_{S\subseteq\{1,2,\ldots,\mathrm{K}\}\setminus\{i\}}\frac{|S|!(K-|S|-1)!}{K!}[f_x(S\cup\{i\})-f_x(S)], \tag{8}$$

$$f_x(S)=\mathrm{E}[f(x)|x_s], \tag{9}$$

Where S corresponds to a feature subset within the input space. K indicates the set of all inputs. $\mathrm{E}[f(x)|x_s]$ refers to the expected value of the function on subset S. Consider a training dataset $\{x^i,y^i\}_{i=1,\ldots,n_{train}}$ used to train a predictive model $f(x)$ (specifically a DevNet in this study). This model aims to establish a mapping as accurately as possible between the response value y and the input training data x. SHAP provides an interpretable model $g(x)$ to reveal the impact of input features on the model's output.

$$f(x)=g(x^{'})=\phi_0+\sum_{i=1}^{M}\phi_i x_i^{'}, \tag{10}$$

Where $x^{'}$ denotes the simplified input, $x=\mathrm{h}_x(x^{'})$ serving as a function that transform $x^{'}$ to the original x, and $\phi_0=f(\mathrm{h}_x(0))$ represents the model output with all simplified inputs toggled off. Several versions of SHAP, such as DeepSHAP, Kernel SHAP, LinearSHAP, and TreeSHAP, have been proposed to cater to specific categories of ML models. In this study, we adopted Kernel SHAP. Kernel SHAP integrates linear LIME and constructs an interpretable local model by utilising a small background dataset derived from the data, approximating the original model f. This approach enhances the sample efficiency and accuracy of SHAP value estimations without relying on specific model types.

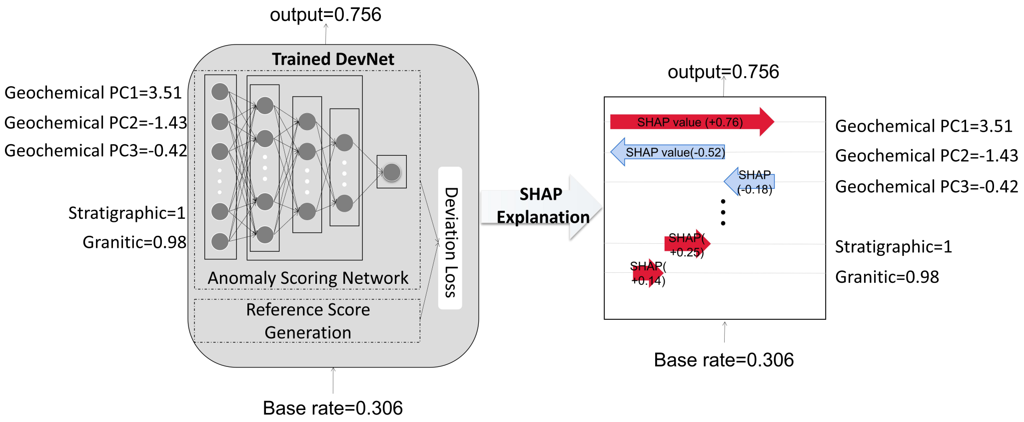


Figure 4: Illustration of SHAP-based interpretation of model predictions via feature contributions.

3.2.3. Implementation

The processing steps of this study include multi-source exploration data preprocessing, feature generation, building a DL model, and interpreting the model, primarily relying on the Python programming language. Additionally, the R programming language is utilised for the ILR transformation and RPCA of the geochemical data. Notably, the construction of DevNet significantly impacts the quality of the mineral prospectivity map. Hyperparameter tuning is conducted to identify the optimal settings, with the goal of improving the performance of the DL model. The architecture of the DevNet model is determined by comparing hyperparameter settings, depth, activation functions, optimisation algorithms, and learning rates of different models based on their receiver operating characteristic (ROC) curves (Fawcett, 2006) on the testing set. The ROC curve is a common technique for classification model performance assessment; it is a two-dimensional plot where the Y-axis represents the true positive rate (TPR) and the X-axis represents the false positive rate (FPR). The model's classification performance is typically quantified by the area under the curve (AUC). The AUC ranges from 0 to 1, summarising the overall performance of the classification model into a single statistical metric. An AUC measurement of 0.5 reflects performance equivalent to random classification, while an AUC value approaching 1 indicates superior classification performance.

Consequently, the architecture of the determined DevNet model comprises an input layer, two hidden layers, and an output layer, with the number of neurons being 57, 24, 12, and 1, respectively, and the weights and biases shown in Figure B1. The hidden layers use the ReLU activation function to introduce non-linearity, enhancing the network's ability to learn complex features. The output layer employs a linear activation function to generate continuous probability scores. The optimisation algorithm used for the model is Nadam, with the initial learning rate set at 0.005. Each batch comprises 128 samples during training, with each epoch consists of 5 batches. A total of 500 epochs are used for model training.

4. Results and Discussion

This study efficiently integrates multi-source exploration data and transforms them into a mineral prospectivity map utilising a semi-supervised DL model. Additionally, it provides interpretability for understanding the model's prediction process. Geological, geophysical, geochemical, remote sensing and DEM data layers are used to construct 57 features for each sample. The DL-based mineral prospectivity model, DevNet, trained with optimised parameter configurations, effectively captures the complex mapping relationships between multidimensional features and mineralisation probability. The mineralisation probability is obtained by normalizing the anomaly scores output (the original value of the model output) by DevNet. Table 3 presents the performance evaluation results of DevNet based on various metrics, using all positive and negative samples from the randomly selected sample set. In addition to the previously mentioned AUC value, accuracy indicates the percentage of samples correctly labelled by the model out of the total number of samples. Cohen's Kappa Coefficient (Cohen, 1960) is a statistical measure of agreement between categorical outcomes, adjusted for the possibility of random chance. This metric effectively reflects the difference between the classifier and random guessing, making it particularly important in evaluating multi-class or imbalanced datasets. The Matthews Correlation Coefficient (MCC) (Baldi et al., 2000) is another robust metric that addresses class imbalance issues. MCC is a symmetrical measure, meaning that its value remains unchanged when positive and negative class labels are swapped. The MCC ranges from -1 to 1, where -1 indicates completely incorrect classification (i.e., the model predicts all positive samples as negative and all negative samples as positive), 1 indicates perfect classification, and values near 0 suggest predictions close to random guessing. The calculation of these three metrics relies on the confusion matrix derived from the model's predictions versus the actual labels. Thus, the classification threshold – the point at which the model's predicted probabilities are converted into binary labels (e.g., positive or negative) – plays a critical role in model evaluation. In this study, the classification threshold is determined by evaluating multiple thresholds and selecting the one that maximises the F1 score. Figure 5 shows the mineralisation probability prediction results based on the trained DevNet model for the regularly sampled prediction dataset. The geometrical Interval method divides the predicted probabilities into four intervals, with nearly all REE deposits located within or near the red high-probability area. The exception being Location G – Mount Gee East – which, however, represents an extensive subsurface U and

REE deposit. The high immediately to the west – Location Y – highlights a number of small hydrothermal Fe-rich U and REE prospects.

Additionally, these red high-potential areas are largely located within the distribution range of felsic granite. Granite bodies are typically considered key heat sources for hydrothermal circulation associated with mineralisation, and their hydrothermal activity is closely related to the mineralisation processes (Hoatson et al., 2011). According to previous studies, the REE mineralisation in the study area is primarily driven during initial stages by either introduction of felsic magmatism, including fluid activity or heat originating from the felsic magmatism (Hoatson et al., 2011), or during subsequent events by the radiogenic heat generated by the felsic granites, or the mobilsation of fluids generated by the radiogenic heat. Notably, the yellow medium-potential and red high-potential areas are distributed in regions characterised by fault development and probable fault reactivation within the granitic and metasedimentary Mesoproterozoic strata, which aligns with the REE mineralisation models utilising fluid pathways at varying time intervals within the study area. Success rate curves are highly effective in revealing the model's predictive performance. The success rate curve (Xu et al., 2020) is employed to evaluate the consistency between the anomaly probability scores derived from predictive samples and known mineral occurrences. This curve is generated by displaying the proportion of accurately identified deposits along the ordinate versus the proportion of the study area designated as prospective along the abscissa. Success rate analysis indicates that the top 2% of the study area contains 86% of the known mineral deposits, and 30% of the area delineates all the REE deposits (Fig. 6), indicating that DevNet exhibits high accuracy in identifying high-potential mineralisation areas. The model-defined prospective areas not only closely match the spatial distribution of known mineral occurrences but also align with existing geological knowledge, demonstrating significant spatial correlations with fault zones, felsic granite bodies, and Mesoproterozoic metasedimentary strata. These results provide valuable guidance and decision support for future mineral exploration efforts within the study area.

Table 3: Evaluation metrics for DevNet performance.

Metric	Value
AUC-ROC	0.9926
Accuracy	0.9974
MCC	0.9246
Cohen's Kappa	0.9218
F1 score	0.9231

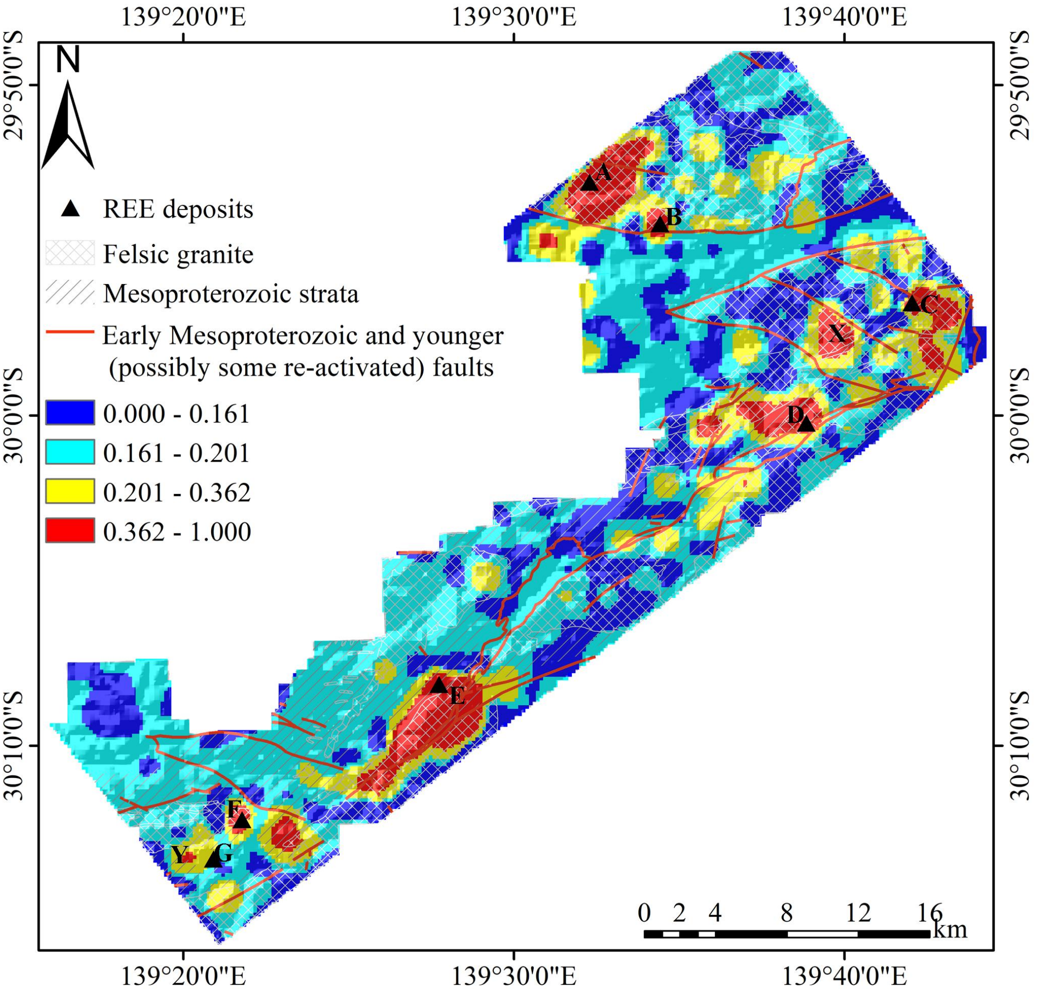


Figure 5: Prospectivity map showing the spatial distribution of predicted values for REE deposits generated by DevNet in the north-western Curnamona Province.

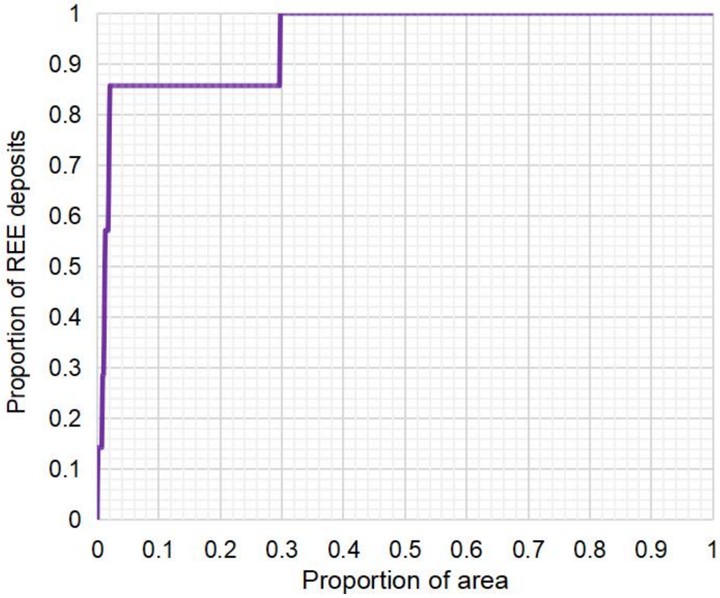
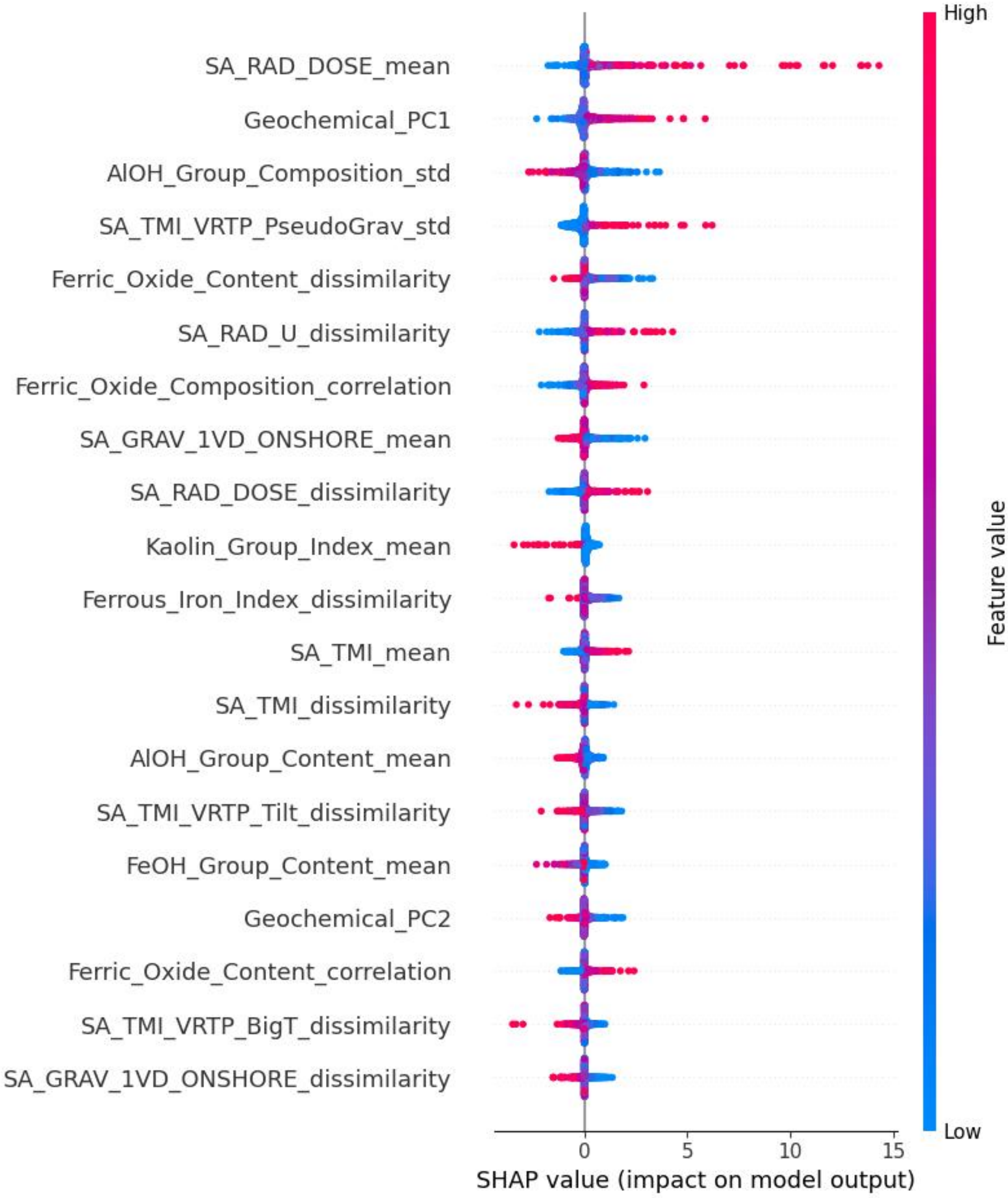


Figure 6: Success-rate curve of DevNet on the predictive dataset.

The “black-box” nature of DL models makes their underlying decision processes difficult to interpret, leading to opacity regarding key driving factors and uncertainty in decision pathways. This lack of transparency can undermine trust in delineating prospective mineralisation zones. To elucidate the internal mechanisms of the DevNet model, this study employs the SHAP technique to analyse the model’s predictive decisions from the perspective of feature contributions. Figures 7 and 8 present SHAP summary plots, which assess each feature's significance in the model’s decision-making framework from a global perspective, revealing the relationship between eature inputs and predicted outputs. In both figures, the top variables contribute more to the model than the bottom ones. In the scatter plot (Fig. 7), each row represents a feature, and the x-axis shows the SHAP values, which can be interpreted as the contribution of that feature to the model’s output. Each point in the plot represents a sample, with colour indicating the feature value—red for high values and blue for low values. The bar plot (Fig. 8) displays the ranking of feature importance along with their corresponding average SHAP values, reflecting the average contribution of each feature to the model’s decision-making. These visualisations provide a clear understanding of the relative importance of different features within the model and their impact on the prediction outcomes. A pronounced positive relationship between the radioactive element concentrations and REEs abundances underscores the importance of radiometric data in REE exploration (Leroy and Turpin, 1988; Shah et al., 2021), while the contribution of ground dose data (Fig. 8) further validates the effectiveness of radiometric measurements in predicting REE prospectivity zones, consistent with existing geological theories.

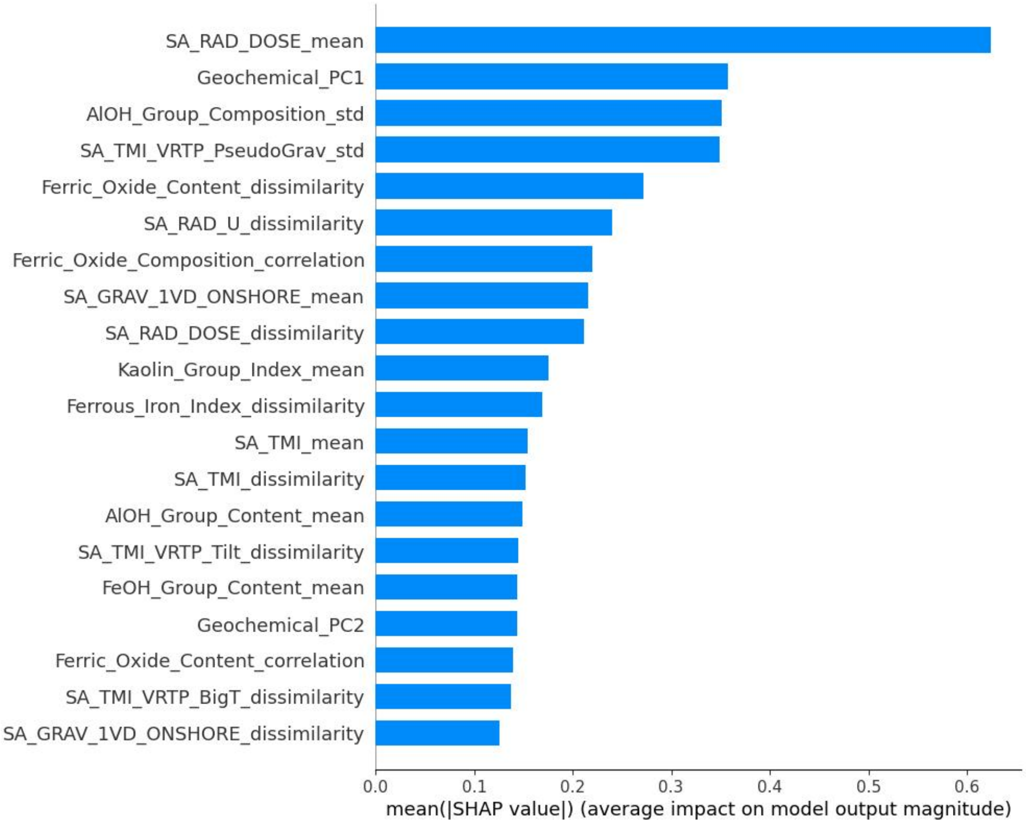
Furthermore, we used the SHAP decision plot (Fig. 9) to provide localised explanations for the set of samples with high probability values. In the plot, each line represents a sample, with the x-axis indicating the model's output value. The results reveal that ground radiation dose is the key factor driving high probability scores, highlighting its crucial role in mineralisation prediction. And the prospectivity map shows that medium- to high-mineralisation potential areas closely align with areas of high radiation dose (Fig. 10). Similarly, the ground radiation of U (Fig. 10) exhibits a positive interaction with mineralisation probability in the SHAP explanation, which is consistent with existing geological understanding (Fig. 7). The pseudo-gravity values derived from total magnetic intensity data adjusted for polarization (SA_TMI_VRTP_PseudoGrav_std) (Fig. 10) effectively reflect the magnetic response of subsurface geological bodies. High values of this feature explain the magnetic anomalies associated with carbonatite bodies linked to skarn-type REE mineralisation (Simandl and Paradis, 2018; Thomas et al., 2016). SHAP analysis of the model's decision-making further validates that the DevNet model successfully learns and captures the intrinsic relationship between this feature and mineralisation probability, confirming its critical role in the prediction process (Fig. 7). Figure 8 reveals that the standard deviation of AlOH group composition (AlOH_Group_Composition_std) (Fig. 10) and the dissimilarity of ferric oxide content (Ferric_Oxid_Content_dissimilarity) (Fig. 10) significantly contribute to the model's predictions. These features show a negative correlation with the model's output anomaly values, indicating that higher values of these features correspond to lower mineralisation probabilities and vice versa (Fig. 7). The SHAP interpretation plots indicate that the first (Geochemical_PC1) and second (Geochemical_PC2) PCs of the geochemical data are more important than other geochemical PCs, with the first PC being particularly significant. Specifically, the first PC scores (Fig. 10) are positively correlated with the model's output anomaly scores, meaning that higher scores in the first PC increase the likelihood of higher mineralisation probability scores in the DevNet model (Fig. 7). The ranked-scaled eigenvector plot of the PCs provides an intuitive visualisation of the loadings of each element across the PCs, including their magnitude and sign. The absolute value of the feature loadings reflects the relative contribution of each element to the respective PC, while the sign (positive or negative) of the loadings indicates the relationship between the element and the PC. Differences in sign imply that elements exert opposite influences on the PC. In the first PC, elements such as Rb, Th, W, Cs, Mo, and Pb exhibit significant contributions (Fig. 11a). Based on the well-established association between Th and REE concentrations in exploration geochemistry, REEs are commonly found coexisting with minerals containing radioactive elements like Th and U, with REE distribution frequently accompanied by Th anomalies (Dhurandhar and Saxena, 1999). Rb and Cs are typical large-ion lithophile elements whose enrichment is commonly associated with late-stage magmatic evolution or specific hydrothermal activities, and they may co-occur with other elements in geological environments related to REE mineralisation (Jowitt et al., 2017). W, Mo, and Pb are common ore-forming elements in skarn-type deposits, with W and Mo being particularly prevalent in hydrothermal systems associated with granite intrusions. The model's decision-making aligns closely with existing geological knowledge, reinforcing the relationship between this combination of elements and mineralisation processes. SHAP analysis reveals that lower scores in the second PC increase the likelihood of high anomaly scores in DevNet. Examining the relationship between scores and loadings, Figure 11b indicates that high concentrations of Sc, Y, Ni, La, Nb, and Th may result in lower second PC scores. In addition to the previously mentioned role of Th in indicating REE mineralisation, Sc, La, and Y are critical REEs, which are a group of seventeen chemically similar metallic elements. The SHAP method's interpretation of the model's decisions indicates that the model effectively learns and captures the logical relationships between these features and mineralisation probability. It not only identifies key features associated with mineralisation but also differentiates their contributions to the mineralising environment. The use of SHAP further validates the model's scientific rationale within the geological context, thereby enhancing the reliability and interpretability of the prediction results.

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025



515 **Figure 7: SHAP summary scatter plot illustrating the association between individual feature values and their corresponding predicted probability scores. Abbreviations: SA–South Australia; RAD–Radiometric; PC–Principal Component; std–Standard Deviation; TMI–Total Magnetic Intensity; VRTP–Variable Reduction to Pole (applied to TMI); PseudoGrav–Pseudo Gravity (applied to TMI VRTP); GRAV–Gravity; 1VD–First Vertical Derivative (applied to onshore gravity); Tilt–Tilt Angle (applied to TMI VRTP); BigT–Amplitude of the Total Vector (applied to TMI VRTP).**

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025



520 **Figure 8: Bar plot of SHAP summary depicting feature importance in DevNet model predictions. Abbreviations: SA–South Australia; RAD–Radiometric; PC–Principal Component; std–Standard Deviation; TMI–Total Magnetic Intensity; VRTP–Variable Reduction to Pole (applied to TMI); PseudoGrav–Pseudo Gravity (applied to TMI VRTP); GRAV–Gravity; 1VD–First Vertical Derivative (applied to onshore gravity); Tilt–Tilt Angle (applied to TMI VRTP); BigT–Amplitude of the Total Vector (applied to TMI VRTP).**

525

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

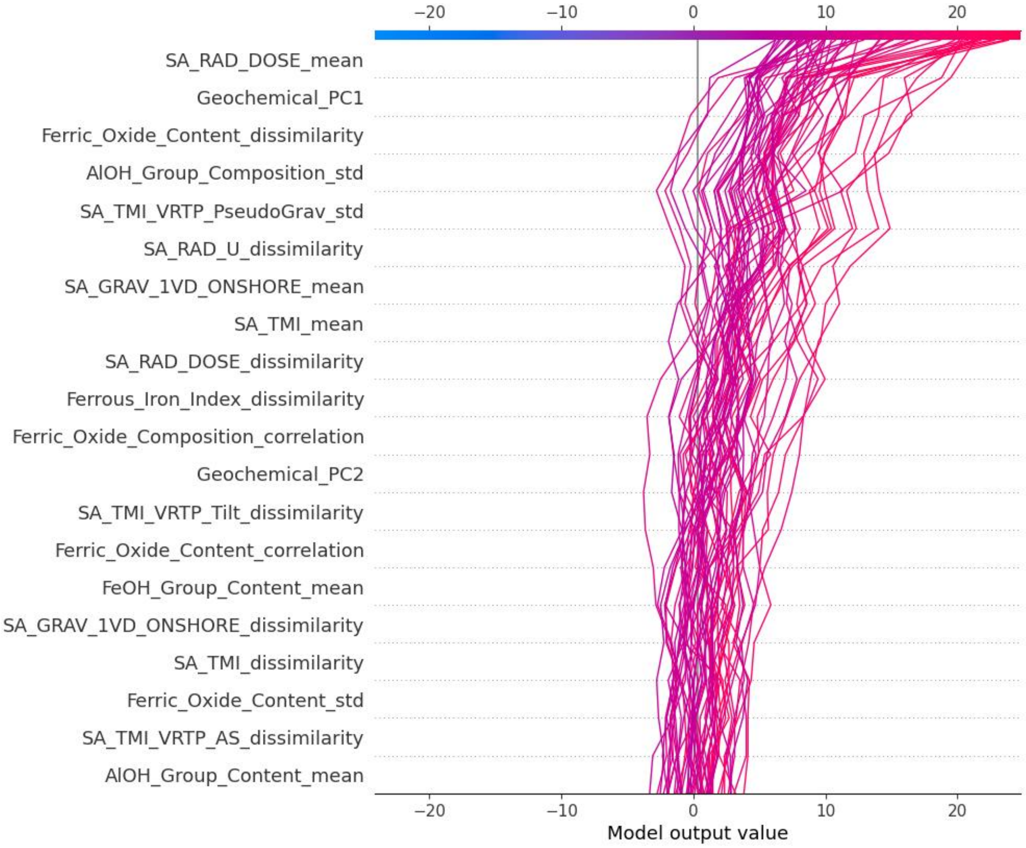


Figure 9: SHAP dependence plot showing the impact of different features on the original value of DevNet model output. Abbreviations: SA–South Australia; RAD–Radiometric; PC–Principal Component; std–Standard Deviation; TMI–Total Magnetic Intensity; VRTP–Variable Reduction to Pole (applied to TMI); 530 PseudoGrav–Pseudo Gravity (applied to TMI VRTP); GRAV–Gravity; 1VD–First Vertical Derivative (applied to onshore gravity); Tilt–Tilt Angle (applied to TMI VRTP); AS–Analytic Signal (applied to TMI VRTP).

https://doi.org/10.5194/egusphere-2025-3283
Preprint. Discussion started: 15 September 2025

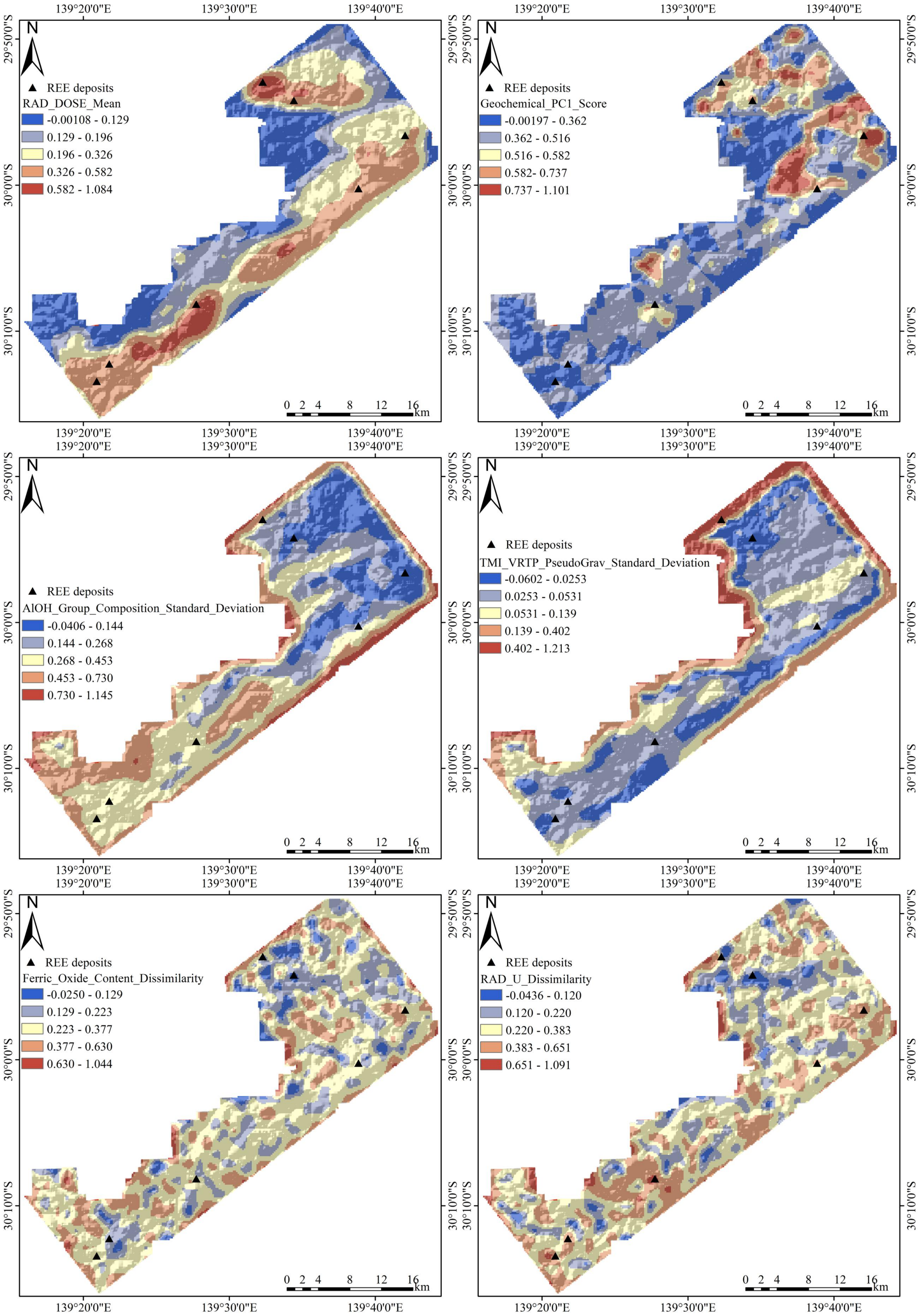
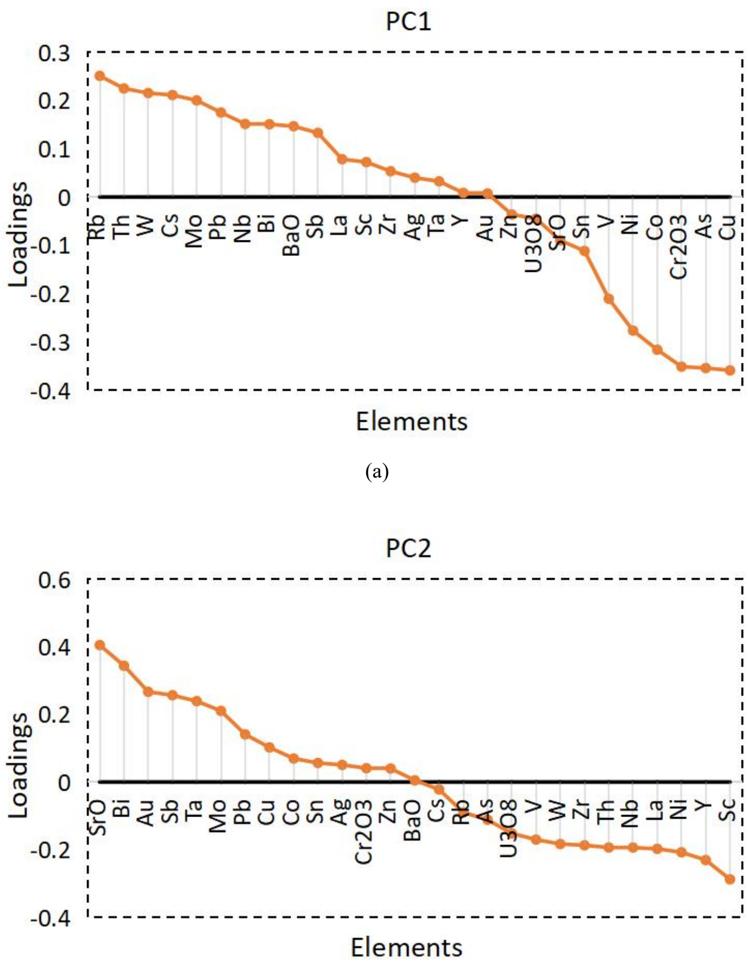


Figure 10: Spatial distribution of the top six features influencing model decisions. Abbreviations: RAD–Radiometric; PC–Principal Component; TMI–Total Magnetic Intensity; VRTP–Variable Reduction to Pole (applied to TMI); PseudoGrav–Pseudo Gravity (applied to TMI VRTP).



(a)

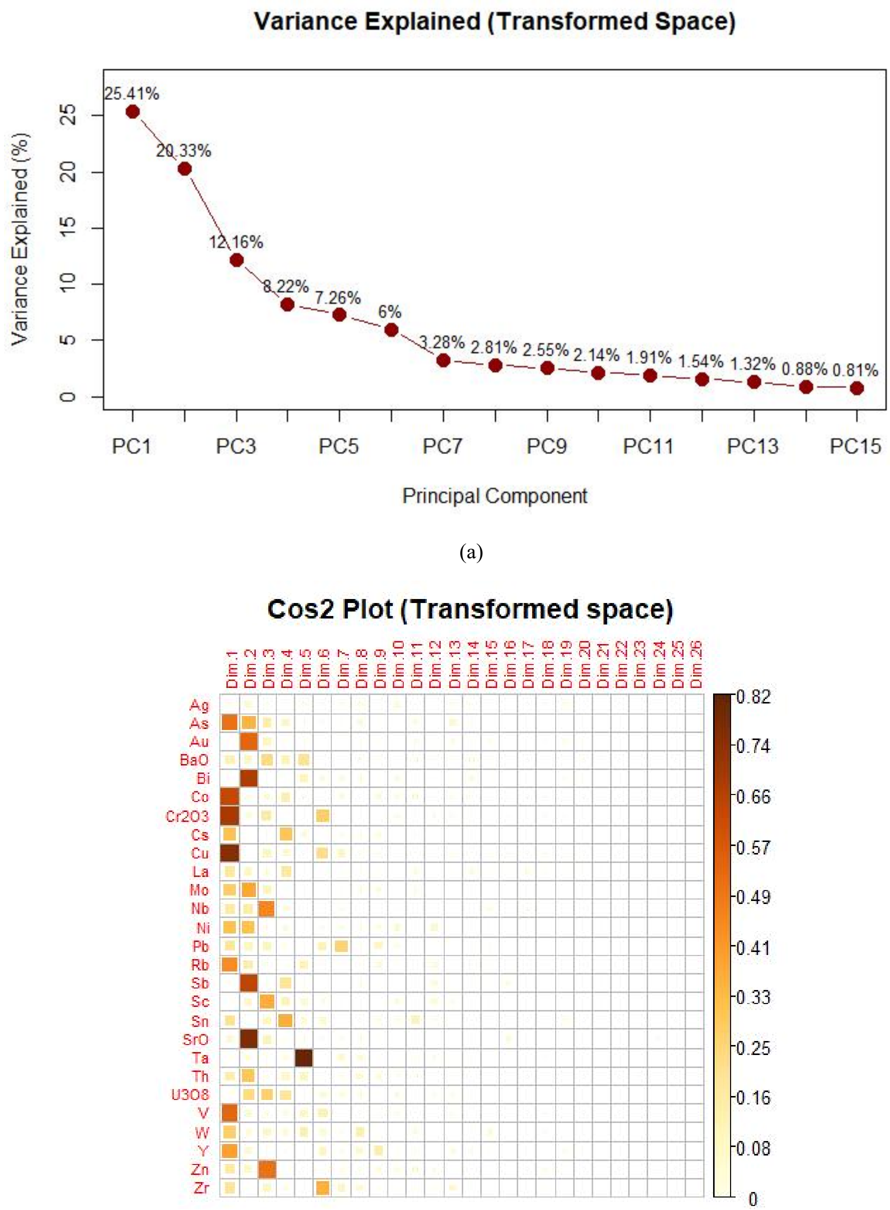
(b)

Figure 11: Ranked-scaled eigenvector plots of (a) first and (b) second principal component.

5. Conclusions

This study establishes the DEEP-SEAM framework for MPM, providing a generalised solution for effectively capturing the non-linear relationships between features extracted from multi-source exploration data and mineralisation processes. The framework alleviates the data noise and redundancy problems caused by complex geological processes and effectively addresses the overfitting phenomena arising from significant class imbalance between positive and negative samples in mineral prediction. Furthermore, the framework delineates potential mineralisation areas associated with specific deposit types using the DevNet model, under the guidance of limited mineralisation information, with prospective areas showing strong spatial association with known mineral occurrences. To enhance the DevNet model's transparency and interpretability, the framework incorporates the SHAP method, which provides insights into the overall decision-making process of the DL model and the local effects of feature contributions. The case study indicates that the DevNet model implemented within the framework demonstrates high predictive accuracy. SHAP analysis of feature roles within the model, coupled with interpretation in the context of existing mineralisation theories, indicates that DevNet can effectively learn potential clues for mineral deposit identification from multi-source exploration data. This further validates the reliability of delineating mineral prospectivity areas. DEEP-SEAM v1.0 proposed in this study demonstrates its capability to effectively predict mineralisation potential in complex geological settings, providing essential decision support for future mineral exploration efforts.

Appendix A. RPCA Analysis Plots

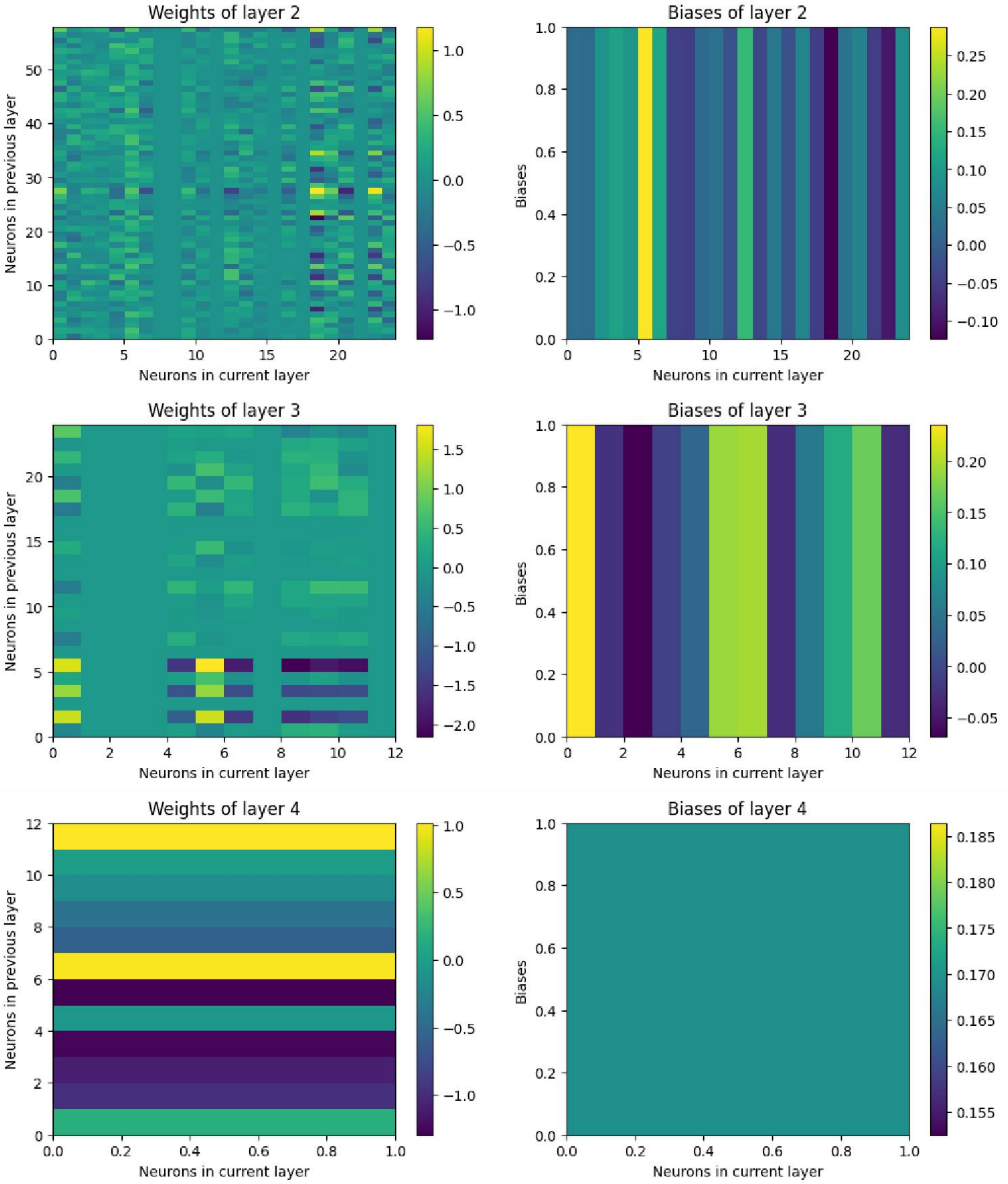


(a)

(b)

Figure A1: RPCA analysis results: (a) scree plot of back-transformed geochemical variables based on variance, (b) representation quality of back-transformed geochemical variables by PCs.

Appendix B. Weights and Biases Heatmap



570 **Figure B1: Heatmap of neuron weights and biases for all layers of the DevNet model except the input layer.**

Code and Data Availability

The code and data used for MPM based on the DEEP-SEAM v1.0 framework are archived on Zenodo (https://doi.org/10.5281/zenodo.17098677; Luo et al., 2025). All data used in this study are publicly available. The raw data can be downloaded from the South Australian Resources Information Gateway (https://map.sarig.sa.gov.au). The DEEP-SEAM v1.0 is also available on GitHub (https://github.com/EarthByte/MPM_Curnamona_REE)

CRediT Authorship Contribution Statement

ZL: Conceptualization, Methodology, Software, Writing – original draft. **EF:** Conceptualization, Resources, Software, Writing - review & editing. **SH:** Conceptualization, Writing - review & editing. **RDM:** Conceptualization, Writing - review & editing, Supervision.

Declaration of Competing Interest

The contact author has declared that none of the authors has any competing interests.

Acknowledgments

ZL was supported by a scholarship from the China Scholarship Council (CSC) under Grant CSC No. 202206410105. EF and RDM were supported by the Australian Research Council grant LP210100173.

References

Alonso, E., Sherman, A. M., Wallington, T. J., Everson, M. P., Field, F. R., Roth, R., and Kirchain, R. E.: Evaluating Rare Earth Element Availability: A Case with Revolutionary Demand from Clean Technologies, Environ. Sci. Technol., 46, 3406–3414, https://doi.org/10.1021/es203518d, 2012.

Alper, O. C., Doğan, H., and Öztürk, H.: Gear pitting fault detection: Leveraging anomaly detection methods. In 2023 14th International Conference on Electrical and Electronics Engineering (ELECO) (pp. 1-5). IEEE, Doi: 10.1109/ELECO60389.2023.10416063, 2023.

Alley N.F., and Hore S.B.: Early Cretaceous sediments reveal a story of prolonged cold climate, glaciations, oscillating sea level and tectonic changes. Geological Survey of South Australia Bulletin 57, 2022. Department for Energy and Mining, South Australia, Adelaide, 2022.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen, H.: Assessing the accuracy of prediction algorithms for classification: an overview, Bioinformatics, 16, 412–424, https://doi.org/10.1093/bioinformatics/16.5.412, 2000.

Bedini, E.: Mineral mapping in the Kap Simpson complex, central East Greenland, using HyMap and ASTER remote sensing data, Advances in Space Research, 47, 60–73, https://doi.org/10.1016/j.asr.2010.08.021, 2011.

Bergen, K. J., Johnson, P. A., de Hoop, M. V., and Beroza, G. C.: Machine learning for data-driven discovery in solid Earth geoscience, Science, 363, eaau0323, https://doi.org/10.1126/science.aau0323, 2019.

Breiman, L.: Random Forests, Machine Learning, 45, 5–32, https://doi.org/10.1023/A:1010933404324, 2001.

Brown, W. M., Gedeon, T. D., Groves, D. I., and Barnes, R. G.: Artificial neural networks: A new method for mineral prospectivity mapping, Australian Journal of Earth Sciences, 47, 757–770, https://doi.org/10.1046/j.1440-0952.2000.00807.x, 2000.

Bogacz W.V.: The tectogenetic explanation of the uranium potential of the Paralana Mineral System in the Mt Painter Inlier, northwest margin of the Curnamona Craton, South Australia. Abstracts for the September 2006 Conference. Compiled by R.J. Korsch & R.G. Barnes. GA Record 2006/21, 2006.

Bustillo Revuelta, M.: Mineral Resource Exploration, in: Mineral Resources: From Exploration to Sustainability Assessment, edited by: Bustillo Revuelta, M., Springer International Publishing, Cham, 121–222, https://doi.org/10.1007/978-3-319-58760-8_3, 2018.

Carranza, E. J. M.: Controls on mineral deposit occurrence inferred from analysis of their spatial pattern and spatial association with geological features, Ore Geology Reviews, 35, 383–400, https://doi.org/10.1016/j.oregeorev.2009.01.001, 2009.

Chen, Y. and Wu, W.: Mapping mineral prospectivity using an extreme learning machine regression, Ore Geology Reviews, 80, 200–213, https://doi.org/10.1016/j.oregeorev.2016.06.033, 2017.

Cheng, Q.: Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China, Ore Geology Reviews, 32, 314–324, https://doi.org/10.1016/j.oregeorev.2006.10.002, 2007.

Cheng, Q.: Singularity theory and methods for mapping geochemical anomalies caused by buried sources and for predicting undiscovered mineral deposits in covered areas, Journal of Geochemical Exploration, 122, 55–70, https://doi.org/10.1016/j.gexplo.2012.07.007, 2012.

Cohen, J.: A Coefficient of Agreement for Nominal Scales, Educational and Psychological Measurement, 20, 37–46, https://doi.org/10.1177/001316446002000104, 1960.

Connelly, N. G., Damhus, T., Hartshorn, R. M., and Hutton, A. T. (Eds.): Nomenclature of Inorganic Chemistry: IUPAC

Recommendations 2005, Royal Society of Chemistry, 2005.

Curewitz, D. and Karson, J. A.: Structural settings of hydrothermal outflow: Fracture permeability maintained by fault propagation and interaction, Journal of Volcanology and Geothermal Research, 79, 149–168, https://doi.org/10.1016/S0377-0273(97)00027-9, 1997.

Drexel J.F., and Major R.B.: Mount Painter uranium – rare earth deposits. In F.E. Hughes (ed), Geology of the mineral deposits of Australia and Papua New Guinea. Australasian Institute of Mining and Metallurgy (AusIMM). Monograph Series 14, 993–998, 1990.

Dhurandhar, A. P. and Saxena, D. N.: Intergrated Airborne Gammaray Spectral and Satellite Data Analysis for U and REE Mineralisation — A Case Sudy from North Sagobandh Area, District Sonbhadra, Uttar Pradesh, India, J Indian Soc Remote Sens, 27, 43–57, https://doi.org/10.1007/BF02990774, 1999.

Drenth, B. J.: Geophysical expression of a buried niobium and rare earth element deposit: The Elk Creek carbonatite, Nebraska, USA, Interpretation, 2, SJ23–SJ33, https://doi.org/10.1190/INT-2014-0002.1, 2014.

Dushyantha, N., Batapola, N., Ilankoon, I. M. S. K., Rohitha, S., Premasiri, R., Abeysinghe, B., Ratnayake, N., and Dissanayake, K.: The story of rare earth elements (REEs): Occurrences, global distribution, genesis, geology, mineralogy and global production, Ore Geology Reviews, 122, 103521, https://doi.org/10.1016/j.oregeorev.2020.103521, 2020.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C.: Isometric Logratio Transformations for Compositional Data Analysis, Mathematical Geology, 35, 279–300, https://doi.org/10.1023/A:1023818214614, 2003.

Elburg, M. A., Andersen, T., Bons, P. D., Simonsen, S. L., and Weisheit, A.: New constraints on Phanerozoic magmatic and hydrothermal events in the Mt Painter Province, South Australia, Gondwana Research, 24, 700–712, https://doi.org/10.1016/j.gr.2012.12.017, 2013.

Farahbakhsh, E., Maughan, J., and Müller, R. D.: 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, 2023.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognition Letters, 27, 861–874, https://doi.org/10.1016/j.patrec.2005.10.010, 2006.

Filzmoser, P., Hron, K., and Reimann, C.: Principal component analysis for compositional data with outliers, Environmetrics, 20, 621–632, https://doi.org/10.1002/env.966, 2009.

Gao, F., Li, J., Cheng, R., Zhou, Y., and Ye, Y.: Connet: Deep semi-supervised anomaly detection based on sparse positive samples. IEEE Access, 9, 67249-67258, Doi: 10.1109/ACCESS.2021.3077014, 2021.

Goodenough, K. M., Wall, F., and Merriman, D.: The Rare Earth Elements: Demand, Global Resources, and Challenges for Resourcing Future Generations, Nat Resour Res, 27, 201–216, https://doi.org/10.1007/s11053-017-9336-5, 2018.

Goodenough, K. M., Deady, E. A., Beard, C. D., Broom-Fendley, S., Elliott, H. A. L., van den Berg, F., and Öztürk, H.: Carbonatites and Alkaline Igneous Rocks in Post-Collisional Settings: Storehouses of Rare Earth Elements, J. Earth Sci., 32, 1332–1358, https://doi.org/10.1007/s12583-021-1500-5, 2021.

Granek, J. and Haber, E.: Data mining for real mining: A robust algorithm for prospectivity mapping with uncertainties, in: Proceedings of the 2015 SIAM International Conference on Data Mining (SDM), Society for Industrial and Applied Mathematics, 145–153, https://doi.org/10.1137/1.9781611974010.17, 2015.

Grunsky, E. C. and Caritat, P. de: State-of-the-art analysis of geochemical data for mineral exploration, Geochemistry: Exploration, Environment, Analysis, 20, 217–232, https://doi.org/10.1144/geochem2019-031, 2019.

Guth, P. L., Van Niekerk, A., Grohmann, C. H., Muller, J.-P., Hawker, L., Florinsky, I. V., Gesch, D., Reuter, H. I., Herrera-Cruz, V., Riazanoff, S., López-Vázquez, C., Carabajal, C. C., Albinet, C., and Strobl, P.: Digital Elevation Models: Terminology and Definitions, Remote Sensing, 13, 3581, https://doi.org/10.3390/rs13183581, 2021.

Hauke, J. and Kossowski, T.: Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data, Quaestiones Geographicae, 30, 87–93, https://doi.org/10.2478/v10117-011-0021-1, 2011.

Hoatson, D. M., Jaireth, S., and Miezitis, Y.: The major rare-earth-element deposits of Australia: geological setting, exploration, and resources, Geoscience Australia, 2011.

Hore, S. B.: Mount Painter region, South Australia 1:100 000 Geological Atlas Special Series Map, DIGIMAP 00005. Geological Survey of South Australia (GSSA), Adelaide. https://dsd-gdp.s3.amazonaws.com/2037106/DIGIMAP00005.zip, 2015.

Hore, S. B., Hill, S. M., Reid, A., Wade, B., Alley, N. F., and Mason, D. R.: U–Pb geochronology reveals evidence of a Late Devonian hydrothermal event, and protracted hydrothermal–epithermal system, within the Mount Painter Inlier, northern Flinders Ranges, South Australia, Australian Journal of Earth Sciences, 67, 1009–1044, https://doi.org/10.1080/08120099.2020.1793383, 2020a.

Hore, S.B., Hill, S.M., Reid, A., Wade, B., Alley, N.F., and Mason, D.R.: U–Pb geochronology reveals evidence of a Late Devonian hydrothermal event, and protracted hydrothermal–epithermal system, within the Mount Painter Inlier, northern Flinders Ranges, South Australia. Aust. J. Earth Sci. 67(7), 1009–1044.

https://doi.org/10.1080/08120099.2020.1793383, 2020b.

Hronsky, J. M., and Kreuzer, O. P.: Applying spatial prospectivity mapping to exploration targeting: Fundamental practical issues and suggested solutions for the future. Ore Geol. Rev, 107, 647-653, https://doi.org/10.1016/j.oregeorev.2019.03.016, 2019.

Jagodzinski, E. and C.E.Fricke: Compilation of new SHRIMP U-Pb geochronological data for the southern Curnamona Province, South Australia, 2010., 2010.

Jaireth, S., Hoatson, D. M., and Miezitis, Y.: Geological setting and resources of the major rare-earth-element deposits in Australia, Ore Geology Reviews, 62, 72–128, https://doi.org/10.1016/j.oregeorev.2014.02.008, 2014.

Jowitt, S. M., Medlin, C. C., and Cas, R. A. F.: The rare earth element (REE) mineralisation potential of highly fractionated rhyolites: A potential low-grade, bulk tonnage source of critical metals, Ore Geology Reviews, 86, 548–562, https://doi.org/10.1016/j.oregeorev.2017.02.027, 2017.

Khoshnoodi, K., Yazdi, M., Behzadi, M., and Gannadi-Maragheh, M.: Using of ASTER, ETM+ and gamma spectrometry airborne data to find the relationship between the distribution of alkali metasomatism and REE mineralization in the Bafq area, Central Iran., Journal of Sciences, Islamic Republic of Iran, 27, 65–77, 2016.

Kovacs, I.: Origin of the South Australian Heat Flow Anomaly. In: (Eds.) Amos Aikman, Katherine Lilly, Julien Celerier, Istvan Kovacs, and Giselle Estermann, An excursion guide to the Flinders Ranges, South Australia, Journal of the Virtual Explorer, Electronic Edition, ISSN 1441-8142, volume 20, paper 14, doi:10.3809/jvirtex.2005.00137, 2005.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015.

Leevy, J. L., Khoshgoftaar, T. M., Bauder, R. A., and Seliya, N.: A survey on addressing high-class imbalance in big data, J Big Data, 5, 42, https://doi.org/10.1186/s40537-018-0151-6, 2018.

Leroy, J. L., and Turpin, L.: REE, Th and U behaviour during hydrothermal and supergene processes in a granitic environment. Chem. Geol, 68(3-4), 239-251, https://doi.org/10.1016/0009-2541(88)90024-1, 1988.

Li, T., Zuo, R., Xiong, Y., and Peng, Y.: Random-drop data augmentation of deep convolutional neural network for mineral prospectivity mapping. Nat. Resour. Res. 30(1), 27–38. https://doi.org/10.1007/s11053-020-09742-z, 2021.

Liu, B., Dai, Y., Li, X., Lee, W. S., and Yu, P. S.: November. Building text classifiers using positive and unlabeled examples. In Third IEEE international conference on data mining. 179-186. IEEE, Doi: 10.1109/ICDM.2003.1250918, 2003.

Liu, B., Lee, W. S., Yu, P. S., and Li, X.: Partially supervised classification of text documents. In ICML 02(485), 387-394, 2002.

Long, K. R., Van Gosen, B. S., Foley, N. K., and Cordier, D.: The Principal Rare Earth Elements Deposits of the United States: A Summary of Domestic Deposits and a Global Perspective, in: Non-Renewable Resource Issues: Geoscientific and Societal Challenges, edited by: Sinding-Larsen, R. and Wellmer, F.-W., Springer Netherlands, Dordrecht, 131–155, https://doi.org/10.1007/978-90-481-8679-2_7, 2012.

Lottermoser, B. G.: Rare earth elements and hydrothermal ore formation processes, Ore Geology Reviews, 7, 25–41, https://doi.org/10.1016/0169-1368(92)90017-F, 1992.

Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, in: Advances in Neural Information Processing Systems, 2017.

Luo, Z., Farahbakhsh, E., Müller, R. D., and Zuo, R.: Multivariate statistical analysis and bespoke deviation network modeling for geochemical anomaly detection of rare earth elements, Applied Geochemistry, 174, 106146, https://doi.org/10.1016/j.apgeochem.2024.106146, 2024.

Luo, Z., Farahbakhsh, E., Hore, S., and Müller, R. D.: An explainable semi-supervised deep learning framework for mineral prospectivity mapping, GitHub [code], https://github.com/EarthByte/MPM_Curnamona_REE, 2025.

McCuaig, T. C. and Hronsky, J. M. A.: The Mineral System Concept: The Key to Exploration Targeting, in: Building Exploration Capability for the 21st Century, vol. 18, edited by: Kelley, K. D. and Golden, H. C., Society of Economic Geologists, 0, https://doi.org/10.5382/SP.18.08, 2014.

Mou, N., Carranza, E. J. M., Wang, G., and Sun, X.: A Framework for Data-Driven Mineral Prospectivity Mapping with Interpretable Machine Learning and Modulated Predictive Modeling, Nat Resour Res, 32, 2439–2462, https://doi.org/10.1007/s11053-023-10272-7, 2023.

Newton, W., Daly, S., Robertson, S., Preiss, W., conor, C., and Burtt, A.: Overview of geology and mineralisation in South Australia, ASEG Extended Abstracts, 2003, 1–18, https://doi.org/10.1071/ASEGSpec12_01, 2003.

Neumann, N., Sandiford, M., and Foden, J.: Regional geochemistry and continental heat flow: implications for the origin of the South Australian heat flow anomaly. Earth Planet. Sci. Lett. 183 (2000), 107-120, https://doi.org/10.1016/S0012-821X(00)00268-5, 2000.

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López García, Á., Heredia, I., Malík, P., and Hluchý, L.: Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey, Artif Intell Rev, 52, 77–124, https://doi.org/10.1007/s10462-018-09679-z, 2019.

Pang, G., Shen, C., and van den Hengel, A.: Deep Anomaly Detection with Deviation Networks, in: Proceedings of the

25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York, NY, USA, 353–362, https://doi.org/10.1145/3292500.3330871, 2019.

Pang, G., Shen, C., Jin, H., and van den Hengel, A.: Deep weakly-supervised anomaly detection. in: Proceedings of the 29th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York, NY, USA, 1795-1807, https://doi.org/10.1145/3580305.3599302, 2023.

Pour, A. B. and Hashim, M.: Identification of hydrothermal alteration minerals for exploring of porphyry copper deposit using ASTER data, SE Iran, Journal of Asian Earth Sciences, 42, 1309–1323, https://doi.org/10.1016/j.jseaes.2011.07.017, 2011.

Preiss, W. V.: The Adelaide Geosyncline of South Australia and its significance in Neoproterozoic continental reconstruction, Precambrian Research, 100, 21–63, https://doi.org/10.1016/S0301-9268(99)00068-6, 2000.

Reynolds, J. M.: An Introduction to Applied and Environmental Geophysics, John Wiley & Sons, 1249 pp., 2011.

Ribeiro, M. T., Singh, S., and Guestrin, C.: "Why Should I Trust You?": Explaining the Predictions of Any Classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 1135–1144, https://doi.org/10.1145/2939672.2939778, 2016.

Robertson, R. S., Preiss, W. V., Crooks, A. F., Hill, P. W., and Sheard, M. J.: Review of the Proterozoic geology and mineral potential of the Curnamona Province in South Australia., AGSO Journal of Australian Geology and Geophysics, 17, 169–182, 1998.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., and Chica-Rivas, M.: Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines, Ore Geology Reviews, 71, 804–818, https://doi.org/10.1016/j.oregeorev.2015.01.001, 2015.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., and Müller, K. R.: A unifying review of deep and shallow anomaly detection. Proceedings of the IEEE, 109(5), 756-795, https://doi.org/10.1109/JPROC.2021.3052449, 2021.

Ruff, L., Vandermeulen, R. A., Görnitz, N., Binder, A., Müller, E., Müller, K. R., and Kloft, M.: Deep semi-supervised anomaly detection. arxiv preprint arxiv:1906.02694, 2019.

Rutherford, L., Hand, M., and Barovich, K.: Timing of Proterozoic metamorphism in the southern Curnamona Province: implications for tectonic models and continental reconstructions * , Australian Journal of Earth Sciences, 54, 65–81, https://doi.org/10.1080/08120090600981459, 2007.

Sandiford, M., McLaren, S. and Neumann, N.: Long-term thermal consequences of the redistribution of heat-producing elements associated with large-scale granitic complexes. J. metamorphic Geol., 20(1), 87-98. https://doi.org/10.1046/j.0263-4929.2001.00359.x, 2002.

Shah, A. K., Taylor, R. D., Walsh, G. J., and Phillips, J. D.: Integrated geophysical imaging of rare earth element-bearing iron oxide-apatite deposits in the Eastern Adirondack Highlands, New York. Geophysics, 86(1), B37-B54, https://doi.org/10.1190/geo2019-0783.1, 2021.

Sharma, P. V.: Magnetic method applied to mineral exploration, Ore Geology Reviews, 2, 323–357, https://doi.org/10.1016/0169-1368(87)90010-2, 1987.

Sheard, M.J., Fanning, C.M., and Flint, R.B.: Geochronology and definition of Mesoproterozoic volcanics and granitoids of the Mount Babbage Inlier, South Australia. South Australia. Geological Survey. Quarterly Geological Notes, 123, 18-32, 1992.

Shrikumar, A., Greenside, P., and Kundaje, A.: Learning Important Features Through Propagating Activation Differences, in: Proceedings of the 34th International Conference on Machine Learning, International Conference on Machine Learning, 3145–3153, 2017.

Simandl, G. J. and Paradis, S.: Carbonatites: related ore deposits, resources, footprint, and exploration methods, Applied Earth Science, 127, 123–152, https://doi.org/10.1080/25726838.2018.1516935, 2018.

Singer, D. A.: Basic concepts in three-part quantitative assessments of undiscovered mineral resources, Nat Resour Res, 2, 69–81, https://doi.org/10.1007/BF02272804, 1993.

Singer, D. A. and Kouda, R.: Application of a feedforward neural network in the search for Kuroko deposits in the Hokuroku district, Japan, Math Geol, 28, 1017–1023, https://doi.org/10.1007/BF02068587, 1996.

Smith, M. P., Moore, K., Kavecsánszki, D., Finch, A. A., Kynicky, J., and Wall, F.: From mantle to critical zone: A review of large and giant sized deposits of the rare earth elements, Geoscience Frontiers, 7, 315–334, https://doi.org/10.1016/j.gsf.2015.12.006, 2016.

Spandler, C., Slezak, P., and Nazari-Dehkordi, T.: Tectonic significance of Australian rare earth element deposits, Earth-Science Reviews, 207, 103219, https://doi.org/10.1016/j.earscirev.2020.103219, 2020.

Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., and Hu, Z.: Data-driven predictive modelling of mineral prospectivity using machine learning and deep learning methods: A case study from Southern Jiangxi Province, China. Minerals 10(2), 102. https://doi.org/10.3390/min10020102, 2020.

Thomas, M. d., Ford, K. l., and Keating, P.: Review paper: Exploration geophysics for intrusion-hosted rare metals, Geophysical Prospecting, 64, 1275–1304, https://doi.org/10.1111/1365-2478.12352, 2016.

Tukey, J. W.: Exploratory data analysis, Pearson, London, UK, 1977.

VanTrump, G. and Miesch, A. T.: The U.S. geological survey rass-statpac system for management and statistical reduction of geochemical data, Computers & Geosciences, 3, 475–488, https://doi.org/10.1016/0098-3004(77)90025-5, 1977.

Walters, A., and Lusty, P.: Rare earth elements, British Geological Survey, 54pp, 2011.

Wülser, P. A.: Uranium metallogeny in the North Flinders Ranges region of South Australia. PhD Thesis, University of Adelaide, Adelaide, 250 pp, 2009.

Wülser, P. A., Brugger, J., Foden, J., and Pfeifer, H. R.: The Sandstone-Hosted Beverley Uranium Deposit, Lake Frome Basin, South Australia:Mineralogy, Geochemistry, and a Time-Constrained Model for Its Genesis. Econ. Geol. 106, 835–867, https://doi.org/10.2113/econgeo.106.5.835, 2011.

Williams, H. A., Betts, P. G., and Ailleres, L.: Constrained 3D modeling of the Mesoproterozoic Benagerie Volcanics, Australia, Physics of the Earth and Planetary Interiors, 173, 233–253, https://doi.org/10.1016/j.pepi.2009.01.002, 2009.

Xiong, Y., Zuo, R., and Carranza, E. J. M.: Mapping mineral prospectivity through big data analytics and a deep learning algorithm, Ore Geology Reviews, 102, 811–817, https://doi.org/10.1016/j.oregeorev.2018.10.006, 2018.

Xu, S., Hu, X., Carranza, E. J. M., and Wang, G.: Multi-parameter analysis of local singularity mapping and its application to identify geochemical anomalies in the Xishan Gold Deposit, North China. Nat. Resour. Res. 29, 3425-3442, https://doi.org/10.1007/s11053-020-09669-5, 2020.

Yang, N., Zhang, Z., Yang, J., and Hong, Z.: Applications of data augmentation in mineral prospectivity prediction based on convolutional neural networks. Computers & Geosci. 161, 105075. https://doi.org/10.1016/j.cageo.2022.105075, 2022.

Zhang, Y. L., Li, L., Zhou, J., Li, X., and Zhou, Z. H.: Anomaly detection with partially observed anomalies. In Companion Proceedings of the The Web Conference 2018. 639-646, https://doi.org/10.1145/3184558.318658, 2018.

Zheng, H., Sun, D., Han, X., Zhang, X., and Zhao, Y.: Research on Network Security Intrusion Detection Based on Devnet. In Proceedings of the 3rd International Conference on Signal Processing, Computer Networks and Communications. 251-256, https://doi.org/10.1145/3712335.3712379, 2024.

Zhou, B., Li, Z., Zhao, Y., Zhang, C., and Wei, Y.: Rare Earth Elements supply vs. clean energy technologies: new problems to be solve, Gospodarka Surowcami Mineralnymi, 32, 29–44, https://doi.org/10.1515/gospo-2016-0039, 2016.

Zuo, R.: Geodata Science-Based Mineral Prospectivity Mapping: A Review, Nat Resour Res, 29, 3415–3424, https://doi.org/10.1007/s11053-020-09700-9, 2020.

Zuo, R. and Carranza, E. J. M.: Support vector machine: A tool for mapping mineral prospectivity, Computers & Geosciences, 37, 1967–1975, https://doi.org/10.1016/j.cageo.2010.09.014, 2011.

Zuo, R., and Xu, Y.: Graph deep learning model for mapping mineral prospectivity. Math. Geosciences 55(1), 1–21. https://doi.org/10.1007/s11004-022-10015-z, 2023.

Zuo, R., Kreuzer, O. P., Wang, J., Xiong, Y., Zhang, Z., and Wang, Z.: Uncertainties in GIS-Based Mineral Prospectivity Mapping: Key Types, Potential Impacts and Possible Solutions, Nat Resour Res, 30, 3059–3079, https://doi.org/10.1007/s11053-021-09871-z, 2021.

Zuo, R., Luo, Z., Xiong, Y., and Yin, B.: A Geologically Constrained Variational Autoencoder for Mineral Prospectivity Mapping, Nat Resour Res, 31, 1121–1133, https://doi.org/10.1007/s11053-022-10050-x, 2022.

Zuo, R., Xiong, Y., Wang, Z., Wang, J., and Kreuzer, O. P.: A New Generation of Artificial Intelligence Algorithms for Mineral Prospectivity Mapping, Nat Resour Res, 32, 1859–1869, https://doi.org/10.1007/s11053-023-10237-w, 2023.