Natural Resources Research (© 2024)
https://doi.org/10.1007/s11053-024-10387-5

Original Paper

Toward Data-Driven Mineral Prospectivity Mapping from Remote Sensing Data Using Deep Forest Predictive Model

Abdallah M. Mohamed Taha,[1,3] **Gang Liu**,[1,2,3,4,5] **Qiyu Chen**,[1,2,3,4] **Wenyao Fan**,[1,2,3,4] **Zhesi Cui**,[1,2,3,4] **Xuechao Wu**,[1,2,3,4] **and Hongfeng Fang**[1,2,3,4]

Received 24 March 2024; accepted 6 July 2024

Remote sensing data prove to be an effective resource for constructing a data-driven predictive model of mineral prospectivity. Nonetheless, existing deep learning models predominantly rely on neural networks that necessitate a substantial number of samples, posing a challenge during the early stages of exploration. In order to predict mineral prospectivity using remotely sensed data, this study introduced deep forest (DF), a non-neural network deep learning model. Mainly based on ASTER multispectral imagery supplemented by Sentinel-2 and geological data, gold ore in Hamissana area, NE Sudan was used to test the DF predictive model capability. In addition to four geological-based evidential layers, 20 remote sensing-based evidential layers were generated using remote sensing enhancing techniques, forming the predictor variables of the proposed model. The applicability of the DF was thoroughly examined including its accuracy for delineating prospective areas, sensitivity to amount of training samples, and adjustment of hyperparameters. The results demonstrate that DF model outperformed conventional machine learning models (i.e., support vector machine, artificial neural network, and random forest) with AUC of 0.964 and classification accuracy of 93.3%. Moreover, the sensitivity analysis demonstrated that the DF model can be trained with a limited number (i.e., < 15) of mineral occurrences. Therefore, the DF algorithm has great potential and proves to be a viable solution for data-driven prospectivity mapping, particularly in scenarios with data availability constraints.

KEY WORDS: Deep forest, Mineral prospectivity mapping, Remote sensing, Deep learning, Gold mineralization.

INTRODUCTION

Mineral prospectivity mapping (MPM) is a crucial practice in the comprehensive assessment of mineral resources. It is the approach that aids in ranking and delineating target areas of mineral deposits by analyzing and synthesizing various layers of spatial evidence that represents ore-forming factors (Zuo & Carranza, 2011; Carranza, 2017; Zuo, 2020). Defining the targeting criteria, which is the most critical step of MPM, guides the selection of relevant

[1] School of Computer Science, China University of Geosciences, Wuhan 430074, China.

[2] Guizhou Key Laboratory for Strategic Mineral Intelligent Exploration, Guiyang 550081, China.

[3] Key Laboratory of Resource Quantitative Assessment and Geoscience Information, Ministry of Natural Resources, Wuhan 430074, China.

[4] Engineering Research Center of Natural Resource Information Management and Digital Twin Engineering Software, Ministry of Education, Wuhan 430074, China.

[5] To whom correspondence should be addressed; e-mail: liugang@cug.edu.cn

Published online: 23 July 2024

geoscience spatial datasets and the processing techniques to generate the spatial evidence (also called feature variables) (Sun et al., 2019, 2020a; Abedini et al., 2023). Integrating these feature variables and analyzing their spatial association with known deposit locations are carried out by means of numerical methods (Bonham-Carter, 1994a, 1994b; Carranza et al., 2008; Rodriguez-Galiano et al., 2015). Therefore, the selection of numerical methods, such as statistical analyses, machine learning (ML) algorithms, and spatial modeling techniques, is essential for obtaining accurate predictions. Techniques of MPM can generally be categorized into two classes (Carranza & Laborte, 2015a, 2015b, 2016): (i) knowledge-driven methods and (ii) data-driven methods. In common practice, the former and latter categories suit under-explored (also called greenfield) and moderately- to well-explored (also called brownfield) regions, respectively (Parsa, 2021). The knowledge-driven approach employs an expert's deep understanding of mineral deposit indicators to specify how those geological indicators and the targeted mineral deposits are spatially associated (Senanayake et al., 2023). It includes methods such as Boolean logic, fuzzy logic, and binary/multi-class index analysis (Bonham-Carter, 1994a, 1994b; Harris et al., 2001; Brown et al., 2003; Carranza, 2009; Abedi et al., 2013; Kashani et al., 2016). In data-driven predictive modeling, the weighting of particular layers is experimentally determined by utilizing the spatial correlations between evidentiary maps and labeled samples of known mineral deposits (Forson et al., 2022; Fu et al., 2023). Traditionally, data-driven models are either bivariate methods such as weights of evidence and evidential belief (Cheng & Agterberg, 1999; He et al., 2010; Yousefi & Nykänen, 2016) or multivariate probabilistic methods such as logistic regression and discriminant analysis (Bonham-Carter & Chung, 1983; Harris & Pan, 1999; Carranza, 2009; Chen et al., 2011). Unlike knowledge-driven methods, most of the latest advancements in artificial intelligence and cutting-edge research have been dedicated to promoting data-driven techniques. However, those methods can still be limited by data quantity and quality, resulting from data availability issues (Senanayake et al., 2023).

In many non-ideal cases or during the early stage of exploration, remote sensing data might be the only available cost-efficient source of data. These cases may range from large study areas, such as regional exploration, to inaccessible regions where it is challenging to conduct detailed geological investigations or geochemical and geophysical surveys (Ngassam Mbianya et al., 2021; He et al., 2022). Moreover, acquiring geochemical and geophysical big data through conventional methods of collection, processing, and analysis can be challenging, often resulting in the production of expensive and low-velocity data (Zhang et al., 2023). Given that remote sensing data demonstrate big data characteristics and have a close connection with geochemistry, they emerge as a feasible solution for producing evidential layers of several mineral deposit models. These mineral deposits include (Zoheir et al., 2019; Abd El-Wahed et al., 2021; Fu et al., 2023): (i) magmatic-hydrothermal deposits that are related to intrusion (i.e., porphyry or epithermal-vein deposits) and (ii) hydrothermal deposits (i.e., orogenic or volcanic massive sulfide), where the significant mineral indicators can be easily extracted from remotely sensed data. In such instances, remote sensing data offer comprehensive details on lithological units, geological structures, and zones of hydrothermal alteration. (Pour & Hashim, 2011, 2012, 2014; Pour et al., 2016).

The scarcity of studies integrating remote sensing data with data-driven predictive modeling predominantly arises due to several contributing factors, notably (i) the choice of an advanced model that produces an accurate prediction, (ii) the adequacy of the number of training locations, specifically in the stage of early exploration, and (iii) the uncertainty resulting from remote sensing data or the predictive model. Although ML-supervised models such as ANN (artificial neural network), SVM (support vector machine), and RF (random forest) have emerged widely as promising tools for MPM, they have not been investigated comprehensively with remote sensing data. According to Shirmard et al. (2022), the number of publications in the year 2020 that used ''remote sensing'', ''machine learning'', and ''mineral exploration'' as keywords did not exceed eight publications. In view of this, to examine the proficiencies of three multispectral satellite datasets—Landsat-8, Sentinel-2, and ASTER—Mohamed Taha et al. (2023) used a RF prediction model for gold prospectivity mapping in Hamissan area NE Sudan. Although the data synergy of these datasets exhibits promising results, the prediction accuracy of the best single dataset (i.e., ASTER) did not surpass 87.5%. Recently, a subfield of ML known as deep learning (DL) uses deep neural network architecture to learn hierarchical representations of input data to extract higher-level features. (Sun et al., 2020b; Fu et al., 2023). DL models were introduced to MPM and remote sensing data to achieve higher prediction than conventional ML models (Zuo et al., 2019; Li et al., 2020a, 2020b; Xu et al., 2021; Yu et al., 2022). For instance, Fu et al. (2023) reported that convolutional neural network (CNN) outperformed RF and SVM in predicting porphyry copper deposit prospectivity based on geochemical element data and ASTER images combined with hyperspectral imaging. However, to satisfy the needs of big labeled data for DL, data augmentation was implemented.

Despite the big data era advances, many practical tasks still lack enough labeled data because of the high labeling costs. The minimum quantity of documented deposits required for the effective training of data-driven models depends chiefly on an algorithm's robustness and its sensitivity to quantify spatial association with multiple layers of evidential data. For instance, several studies demonstrate that less than 20 positive training locations could be used to train ML models such as RF and ANN (Carranza & Hale, 2003; Magalhães & Souza Filho, 2012; Carranza & Laborte, 2015a, 2015b). Nevertheless, this number remains untested or ambiguous in the case of DL-based MPM. DL models, particularly neural networks, require extensive datasets for effective training due to their intricate structures and their capacity to grasp intricate patterns and representations.

To address the above limitations, this study employed DF model as a novel methodology for MPM. DF is a model proposed by Zhou & Feng (2019) for ensembling different models using the theory behind deep neural network (DNN). It can deal with complex problems with fewer parameters compared with DNN models. Another advantage of the DF is that the training cost can be controlled according to the available computational resources. Unlike the backpropagation procedure of the neural network method, DF has the advantage of automatically terminating the training process by assessing the model performance at each layer. Therefore, the complexity of the DF model depends on the scale of the input datasets, which also makes it suitable for small-scale datasets (Pang et al., 2018; Li et al., 2022; Ma et al., 2022a). Similar to other ensemble learning models, DF demonstrates high generalization capability and quite robust performance even when using default parameter settings (Zhou & Feng, 2019). In general, testing the DF algorithm across different datasets from various fields has shown its higher performance compared to other models (see Zhou & Feng (2019) for details).

The purpose of the current study was to fully exploit the advantages of the DF algorithm to construct a data-driven prospectivity model from remote sensing data, despite the insufficient number of known mineral occurrences. Hence, the study comprehensively investigated the DF capability in terms of mapping accuracy, the model's interpretability of parameters, and sensitivity to the size of the training sample. To achieve these goals, we implemented the same ASTER dataset used by Mohamed Taha et al. (2023) for mapping gold prospectivity in the Hamissana area, NE Sudan.

DEEP FOREST ALGORITHM

As described in the Introduction, DF was initially introduced to overcome the complexity and the burden of explaining the black box architecture of DNN models. As an alternative approach, DF was introduced as a deep ensemble classifier comprising multiple RFs arranged in a structure of cascaded layers. Notably, it diverges from relying on the gradient backpropagation mechanism employed by DNN (Zhang et al., 2020). The primary version of the DF, commonly referred to as gcForest, consists of two essential stages: multi-grained scanning and cascaded forests (Li et al., 2022). The multi-grained scanning enhances representation learning to process sequence or image-style data (Zhao et al., 2021). Because MPM data do not have sequential relationships, this study used a simplified DF by ignoring the multi-grained procedure.

In the cascade structure (layer-by-layer) of the DF, every layer is made up of multiple RFs and completely random forests (CRFs). The number of trees in each forest as well as the number of forests is hyperparameters. Each forest produces feature information (i.e., an estimated class distribution or a regression value) from the input features forming a class vector (Fig. 1a). In subsequent phases, the resulting class vector moves forward to the next level of the cascade. Here, it was concatenated with the original input features, creating a new learning problem different from that in the preceding level (Pang et al., 2018; Su et al., 2019).

The decision to extend a new level of cascading is adaptively determined in response to the overall cascade's performance on the verification set. This makes the DF model's complexity more adaptive

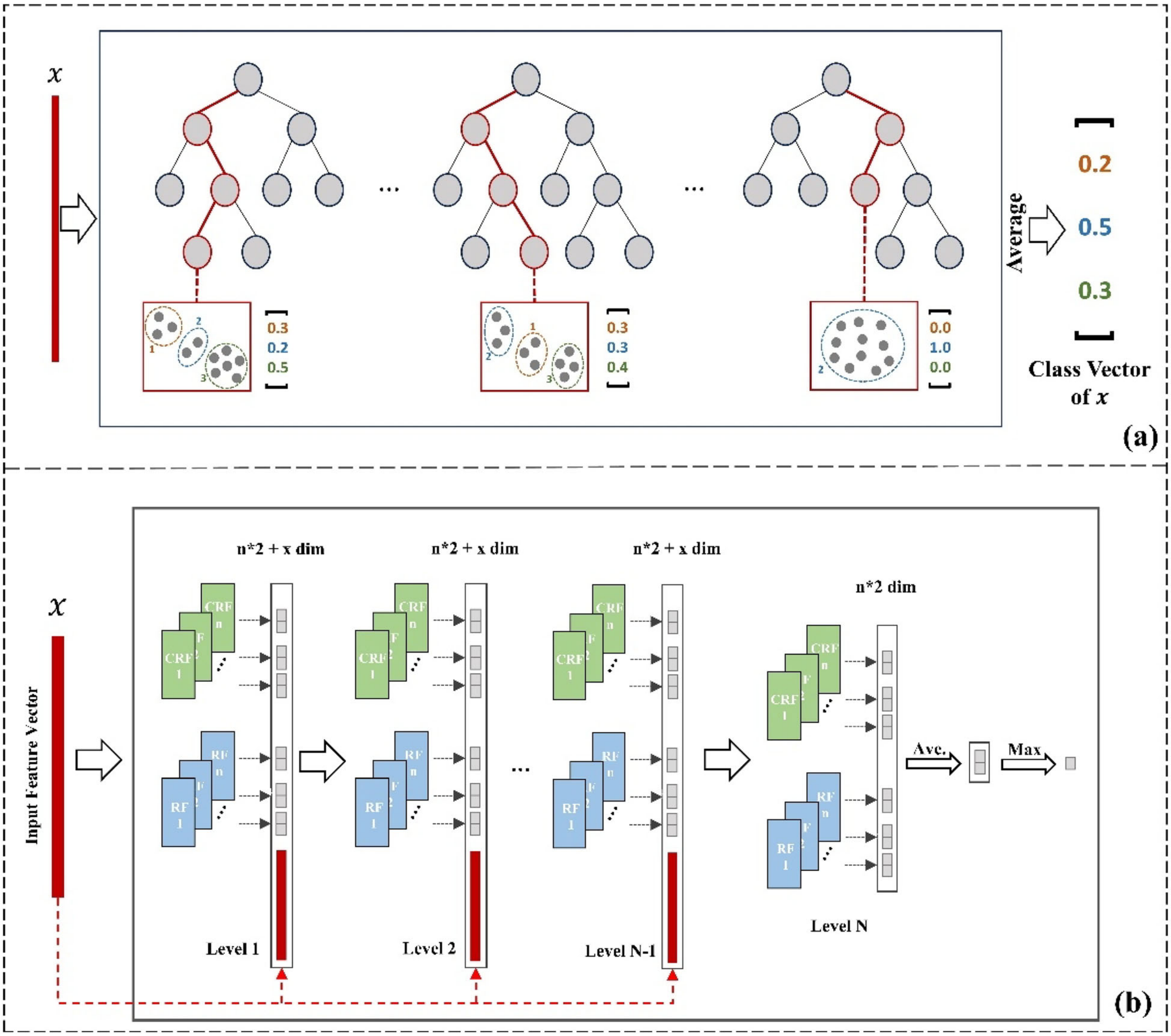

Figure 1. Illustrations of the DF model: (**a**) decision process of a forest for class vector generation and (**b**) cascade forest structures.

than DNN's because it terminates the training process if there is no noticeable performance progress (Zhou & Feng, 2019; Zhang et al., 2020). Hence, the DF's structure is advantageous because it does not rely on the production of copious volumes of data, making it suitable for various scales of training data. Finally, when the cascade layers stop expanding, the average of all the resulting probability vectors is computed, and the prediction result is the label with the maximum probability (Fig. 1b) (Su et al., 2019; Li et al., 2022). As recent developments increase DF complexity, RF or gradient boosting decision tree (GBDT) can be used as a predictor concatenated to the DF.

Referred to as a DF regressor, the DF is structured by employing multiple regression forests, which can then be used for data-driven MPM. Based on the input target variables characterized by binary values (1 and 0 denoting deposit and non-deposit locations, respectively), each hidden forest produces a floating value between 0 and 1 that represents the possibility of mineral deposits (Fig. 2). Hence, the dimensionality of the produced class vector in each cascade layer is ($n \times 1$), where n represents the number of forests (estimators) in the cascade layer.

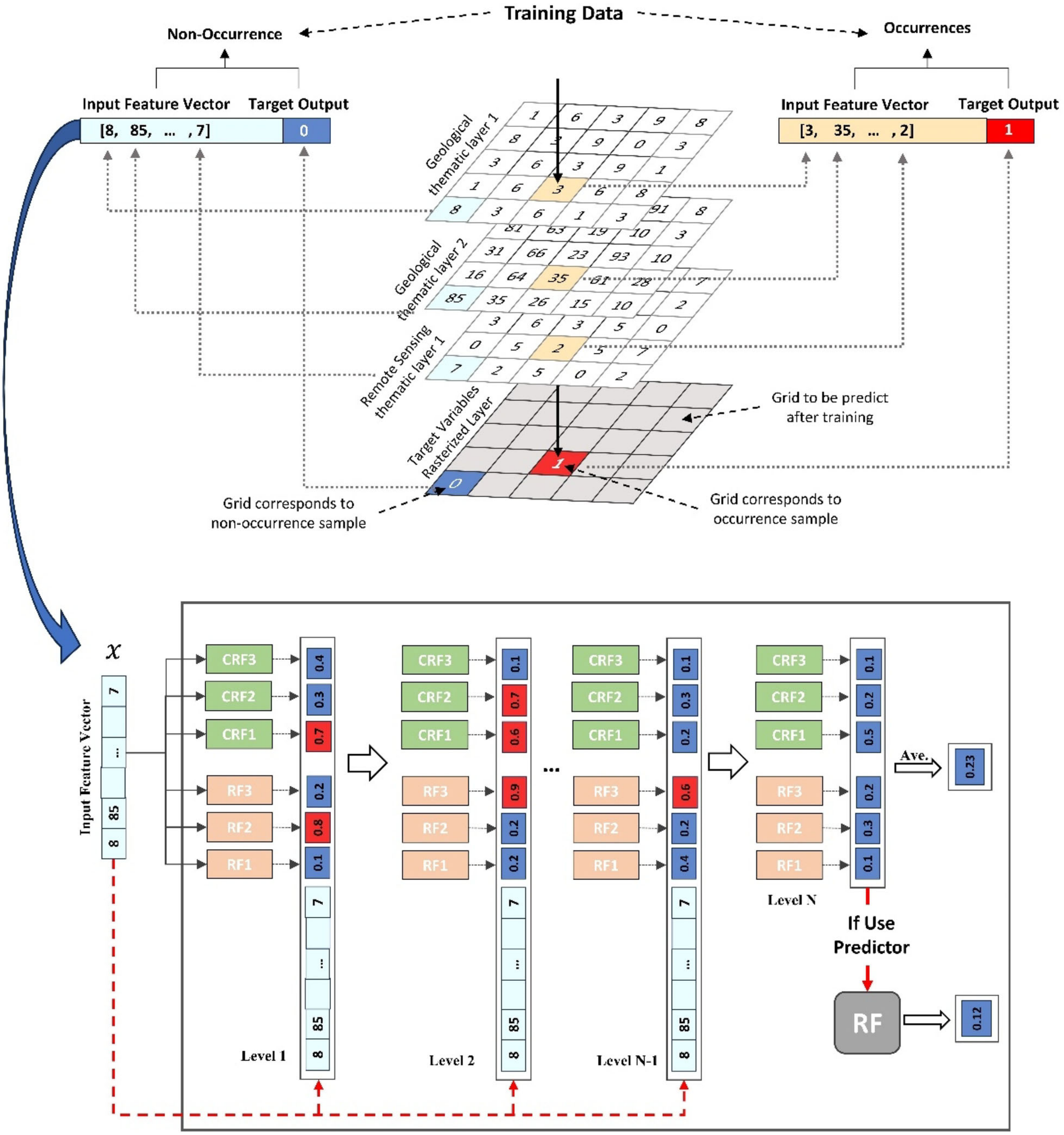
Training Data
Non-Occurrence
Occurrences
Input Feature Vector
Target Output
[8, 85, ... , 7]
0
[3, 35, ... , 2]
1
Geological thematic layer 1
Geological thematic layer 2
Remote Sensing thematic layer 1
Target Variables Rasterized Layer
Grid to be predict after training
Grid corresponds to non-occurrence sample
Grid corresponds to occurrence sample
x
Input Feature Vector
CRF3
CRF2
CRF1
RF3
RF2
RF1
Level 1
Level 2
Level N-1
Level N
Ave.
0.23
If Use Predictor
RF
0.12

Figure 2. DF regressor with structure of three RF and three CRF in each level as an example for MPM.

Therefore, the structure complexity of the DF regressor is determined by the number of estimators, a crucial parameter in the context of predictive modeling for MPM. Other parameters such as the minimum number of samples that are required for internal node splitting and the number of trees are well-known for training any forest because they determine the growth of trees within the forest.

APPLICATION TO MPM IN HAMISSANA AREA (SUDAN)

Study Area and Geological Setting

The selected area for the current study is situated to the west of Hamissana, Wadi Edom, NE Sudan. This area is bounded by latitudes 20° 22′ N to 20° 50′ N and longitudes 34° 00′ E to 34° 45′ E. Geographically, it is positioned on the northwestern flank of the Red Sea Hills, covering an approximate area of 1379 km^2. The study area forms a segment of the Gabgaba terrane, one of the four Arabian Nubian Shield (ANS) terranes in NE Sudan. The Gabgaba terrane extends in a NE–SW direction, running parallel to the Hamissana shear zone. It extends northward to the Eastern Desert Terrane and the Egyptian border, and westward to the city of Atbara. The Gebeit terrane, which is determined by the extent of the Hamissana shear zone, marks the Gabgaba terrane's eastern boundary. The Keraf suture zone, which stretches westward to the Halfa and Bayuda terranes, defines the terrane's western boundary (Fig. 3a).

The late Neoproterozoic rocks exposed in the Hamissana shear zone comprise arc-related low-grade volcano–sedimentary sequences and syn- to post-tectonic intrusive (El Khidir & Babikir, 2013) (Fig. 3b). These rocks are unconformably overlain by immature sediments. The volcano–sedimentary assemblages are related to suturing tectonic and formed during pre- to syn-tectonic events. Metasediment and metavolcanic sequences make up the majority of them, and they are predominant in the study area (Zeinelabdein & Albiely, 2008; Perret et al., 2021). The former, which is composed of marble, schistose turbidites, and E–W-trending quartzite, is the oldest rock unit in the study region. The bending is somehow visible, where the E–W-trending shows its complete or nearly complete rotation from the original orientation to the plane of the deformation. Metavolcanics characterized by low-grade meta-acid volcanic and meta-trachyte crop out well either in surface outcrops or mountain outcrops. In various locations, metasediments are interlayered with metavolcanics and sometimes do not crop out well. In other words, they have a gradual transition that is difficult to map from regional to small scales.

Both pre-to-syn-tectonic (older) and post-tectonic (younger) intrusions are representations of the magmatic assemblage. Older intrusions are composed of meta-diorite to tonalite–trondhjemite–granodiorite (TTG) bodies. They occur as foliated intrusions (i.e., affected by gneissic foliation), being enclosed with the elder sequences (extensive metavolcano–metasedimentary) with sharp contacts (Mohamed et al., 2021). Secondly, the younger intrusions are post-tectonic coarse granitic bodies, which are non-foliated intrusions composed mainly of porphyritic microgranite, granodiorite, and quartz feldspar porphyry. Finally, sedimentary rocks and superficial deposits have the presence of outcrops that are few scattered and low-laying.

Most of the structural features in the area were affected by the N–S Hamissana shear zones with its E-dip. The presence of faults is the most pronounced structure. Based on the rocks, shearing, and slickensides, most faults are strike-slip with horizontal and parallel displacement to the fault's strike. These faults generally trend E–W, NE–SW, and NW–SE (Fig. 3b). Younger granites are cut by several dykes that are oriented ~ N–S with steep dips to the E. Gold mineralization occurs as shear zone-hosted associated with intensely altered zones that are variably zoned outward. Multiple generations of development in quartz veins are present along/within the mineralized intersection. Au-bearing quartz/quartz–carbonate veins and their alteration occurring along the shear zone and cutting/wrapping acid metavolcanic rocks, granite intrusions and/or small felsic intrusions (Zeinelabdein & Albiely, 2008; Zeinelabdein & Nadi, 2014; Bierlein et al., 2015; Johnson et al., 2017; Hamimi et al., 2021; Mohamed et al., 2021; Ahmed, 2022). The sericite and wider halo of carbonate alteration are the main types of alteration of the area. The high-grade parts noticeably contain moderate to weak silicification and occur in the strongly deformed areas with shear zones that are wide and well-developed (Hamimi et al., 2021; Ahmed, 2022).

Spatial Datasets

The spatial dataset of the current study consisted of multispectral remotely sensed data and geological data. All geological maps related to lithology, faults/fractures, and locations of gold occurrences were either acquired or digitized from published literature (Mohamed et al., 2021) and unpublished reports of the Geological Research Authority of the Sudan (GRAS). The geological datasets were stored separately in different shapefile

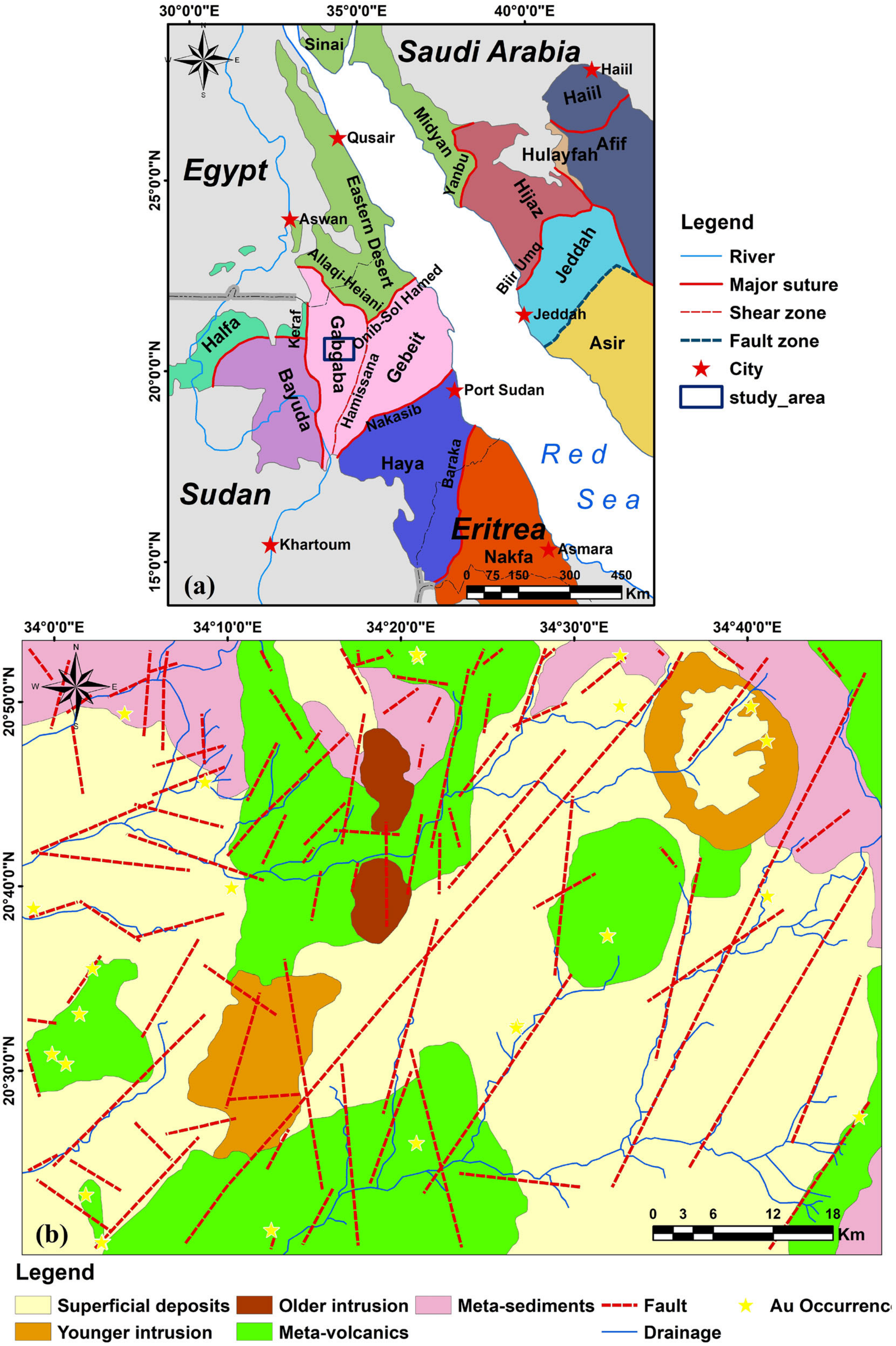

Figure 3. (**a**) Location of the study area and regional structures and (**b**) geological map of the study area (Mohamed Taha et al., 2023).

formats (polygons, lines, and points) to be used in generating geological-based predictor variables. The digitization and preparation of maps were conducted using ArcGIS 10.6.1 software.

We employed the ASTER multispectral remote sensing data, because these demonstrated the highest accuracy among other datasets such as Sentinel-2 and Landsat-8, as noted by Mohamed Taha et al. (2023). Nevertheless, because of its greater spatial resolution, the Sentinel-2 data were utilized to map the study area's lineaments (see *Geological-Based Predictor Variables* section below). The U.S. Geological Survey provides both the ASTER and Sentinel-2 scenes as free data. Four AST_L1T level products were acquired on December 25, 2006 and March 31, 2007. Meanwhile, two scenes of level 1C of Sentinel-2A were acquired on December 3, 2021. These multispectral datasets were free of clouds and terrain correction, meaning that the preprocessing procedures are relatively standardized (Xi et al., 2022; Mohamed Taha et al., 2023). We employed the nearest neighbor method to resample the spatial resolution of ASTER VNIR bands to match that of the SWIR bands (30 m). Also, we utilized a spatial resolution of 30 m for rasterizing the geological data, including lineaments extracted from Sentinel-2, which originally had a resolution of 15 m. Figure 4 shows the experimental process of the current study including data preprocessing, predictor variable generating, training the DF model, and MPM.

Remote Sensing-Based Predictor Variables

One crucial exploration criterion for gold deposits is the presence of hydrothermal alteration zones/minerals (Ali & Pour, 2014; Zhang et al., 2016; Silva dos Santos et al., 2022). Given that the majority of alteration minerals have spectral signatures in the 2.0–2.4-μm wavelength range, significant information about these mineral assemblages can be obtained from remote sensing data (Fu et al., 2023). Argillic, phyllic, and propylitic alteration zones, as well as particular alteration minerals such as hydroxyl-bearing, iron oxides, and clay minerals, have all been effectively mapped in detail using ASTER bands, particularly in the VNIR–SWIR range (Hubbard & Crowley, 2005; Moore et al., 2008; Zhang et al., 2016). To accomplish this, several image processing methods were used, including principal components analysis (PCA), band ratio (BR), mineralogical indices, relative band depth (RBD), and minimum noise fraction (MNF). The successful implementation of these methods facilitates the generation of various thematic layers representing different alteration zones. These layers, after being normalized to the range of [0, 1], were then used as inputs for the predictive model. Table 1 lists all methods that were utilized to produce distinct thematic maps of the targeted alteration minerals and zones.

For minerals and lithology mapping, BR is a very efficient enhancement technique and is one of the most useful image processing techniques (Sabins, 1999; Inzana et al., 2003; Son et al., 2022). This method not only reduces the topographic effect but also enhances the spectral contrast for certain absorption features. Highlighting the spectral contrast of minerals or materials planned to map, BR is achieved by dividing one spectral band by another (Pour et al., 2018; Bolouki et al., 2019). Based on spectral characteristics of ASTER data, five BR images were produced for mapping hydroxyl-bearing, ferric iron, ferrous iron, alunite, and calcite (Fig. 5a–e). For visual aid, the histogram of each generated image was divided into seven intervals using the natural break (Jenks) method, with the first three higher intervals categorized as high, moderate, and low (Fig. 5).

The mineralogical index is another technique that maps relative abundance of specific minerals using mathematical band combinations. This method uses spectral or thermal indices to indicate surface emissivity or reflectance at various wavelengths (Sabins, 1999; Ninomiya, 2003; Zhang et al., 2016; Bolouki et al., 2019; Rajan Girija & Mayappan, 2019). Unlike BR and RBD, mineral indices use different mathematical operations and some constant values, and so it is more complex and fixed to target a specific mineral using specific remote sensing data. Four mineralogical indices were utilized in this study, namely hydroxyl-bearing index (OHI), calcite mineral index (CLI), alunite mineral index (ALI), and kaolinite mineral index (KLI) (Fig. 5f–i).

RBD is the extended version of BR, which includes a three-point ratio formulation. It is calculated by dividing the total number of shoulder bands by the closest band in order to identify the usual absorption signature of a particular mineral or alteration zone. Argillic, phyllic, and propylitic zones were mapped using RBD1 (4 + 6/5), RBD2 (5 + 7/6), and (6 + 9/7 + 8), respectively (Fig. 5j–l).

Unlike the aforementioned methods, PCA and MNF are more complex statistical techniques, which

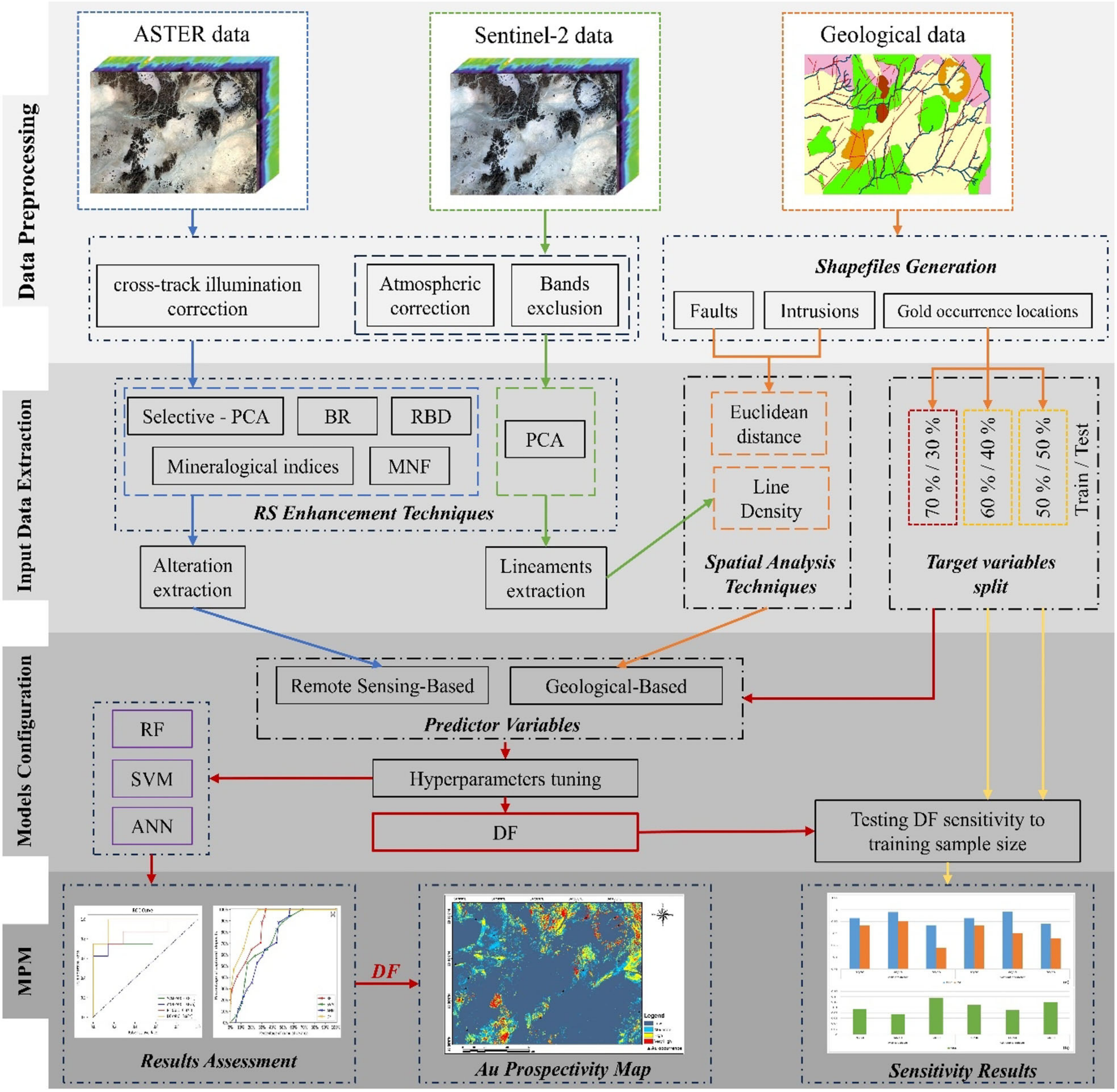

Figure 4. Flowchart of the experiment.

dissolve the reduction in the spectral data. These techniques enhance the remote sensing imagery by transforming information about bands into a new set of data. In other words, they exploit data variability to re-express the original information into a full description of the information in fewer variables. Because each principal component (PC) is formed from uncorrelated linear combinations of values or eigenvector loadings, the resulting dataset (PC components) in the PCA technique shows reduced variance. The computation of these eigenvectors takes place inside a covariance matrix, which represents the statistical connections between each PC. Likewise, the covariance matrix is used by the MNF technique to extract and rescale the noise in the data. The eigenvalue corresponding to each MNF component determines noise reduction and whitening in the transformed dataset.

Table 1. Remote Sensing Enhancement Methods applied to ASTER Data for Mapping the Targeted Minerals

Method	Targeted mineral	Bands used	References
Band ratio (BR)	Hydroxyl-bearing	4/6	Mahdevar et al. (2014), van der Meer et al. (2014), Bolouki et al. (2019), Rajan Girija and Mayappan (2019), Abdelkareem and Al-Arifi (2021)
	Ferric iron	2/1	
	Ferrous iron	(5/3) + (1/2)	
	Alunite	4/7	
	Calcite	4/5	
Relative band depth (RBD)	Argillic (RBD1)	(4 + 6)/5	Zhang et al. (2016), Bolouki et al. (2019), Rajan Girija and Mayappan (2019), Abdelkareem and Al-Arifi (2021)
	Phyllic (RBD2)	(5 + 7)/6	
	Propylitic (RBD3)	(6 + 9)/(7 + 8)	
Mineralogical indices	Hydroxyl-bearing (OHI)	(7/6) × (4/6)	Ninomiya (2003), Rajan Girija and Mayappan (2019), Abdelkareem and Al-Arifi (2021)
	Kaolinite (KLI)	(4/5) × (8/6)	
	Alunite (ALI)	(7/5) × (7/8)	
	Calcite (CLI)	(6/8) × (9/8)	
Minimum noise fraction (MNF)	Related to hydrothermal alteration	The first three that show fair relation	
Principal components analysis (PCA)	Hydroxyl-bearing	1, 3, 4 and 6	Bahrami et al. (2018), Bolouki et al. (2019)
	Iron oxides	1, 2, 3 and 4	
	Argillic	1, 4, 6 and 7	
	Phyllic	1, 3, 5 and 6	
	Propylitic	1, 3, 5 and 8	

In the practice of mineral mapping, PCA appears to offer greater objectivity. Utilizing four selected bands, a modified version of PCA objectively indicates mineral locations through the representation of bright or dark pixels. This technique is referred to as the feature of oriented PC selection, also called the Crosta technique (Loughlin, 1991; Crosta et al., 2003; Bahrami et al., 2018). Based on prior understanding of an object's spectral characteristics, the user chooses these bands (e.g., alteration zone). Two significant loadings with opposing signs indicate the reflectance signature of the targeted mineral/zone that is mapped in one of the PCs. For instance, argillic alteration is mainly recognized by kaolinite, which shows high Al-OH spectral absorption corresponding to ASTER band 6. In this regard, bands 1, 4, 6, and 7 were used to map argillic zones. Hence, PC4 exhibited the typical spectral signature, with strong and oppositely sign loadings in bands 6 and 7. Similarly, phyllic, propylitic, iron oxides, and hydroxyl-bearing zones were mapped using the Crosta technique (Fig. 5m–q).

Conversely, the interpretation of MNF bands is subjective to the visual interpretation and the user's prior knowledge because these bands merely differentiate between areas in the original images. In other words, the MNF images derived from the raw data are just statistics and do not indicate any mineral occurrences. Nine ASTER bands in the VNIR–SWIR spectrum were subjected to MNF transformation in this study. During the screening of resulted images, we focused on the dark and bright pixels that correspond to alteration zones. Subsequently, MNF3, MNF4, and negated MNF2 were selected as predictive variables (Fig. 5r–t).

Geological-Based Predictor Variables

To produce more supplementary thematic maps of ore-controlling factors, geological data were rasterized using spatial analysis methods. Faults represent differential stress changes and deformation resulting from tectonic activity, serving as direct indicators of fluid migration and mineral precipitation. Numerous orogenic gold deposits in the Red Sea Hills are well-documented to be situated in zones characterized by the same linear structures as the shear zone azimuth (Zeinelabdein & Albiely, 2008; Zeinelabdein & Nadi, 2014; Mohamed et al., 2021). Consequently, faults with varying azimuth directions (NE–SW and NW–SE) were utilized to generate two predictor maps using the Euclidean distance method (Fig. 6a, b).

Lineaments are considered the best channels for the movement of hydrothermal solutions either from the source to the deposition or through different depths and lithologies (Abdelkareem & Al-

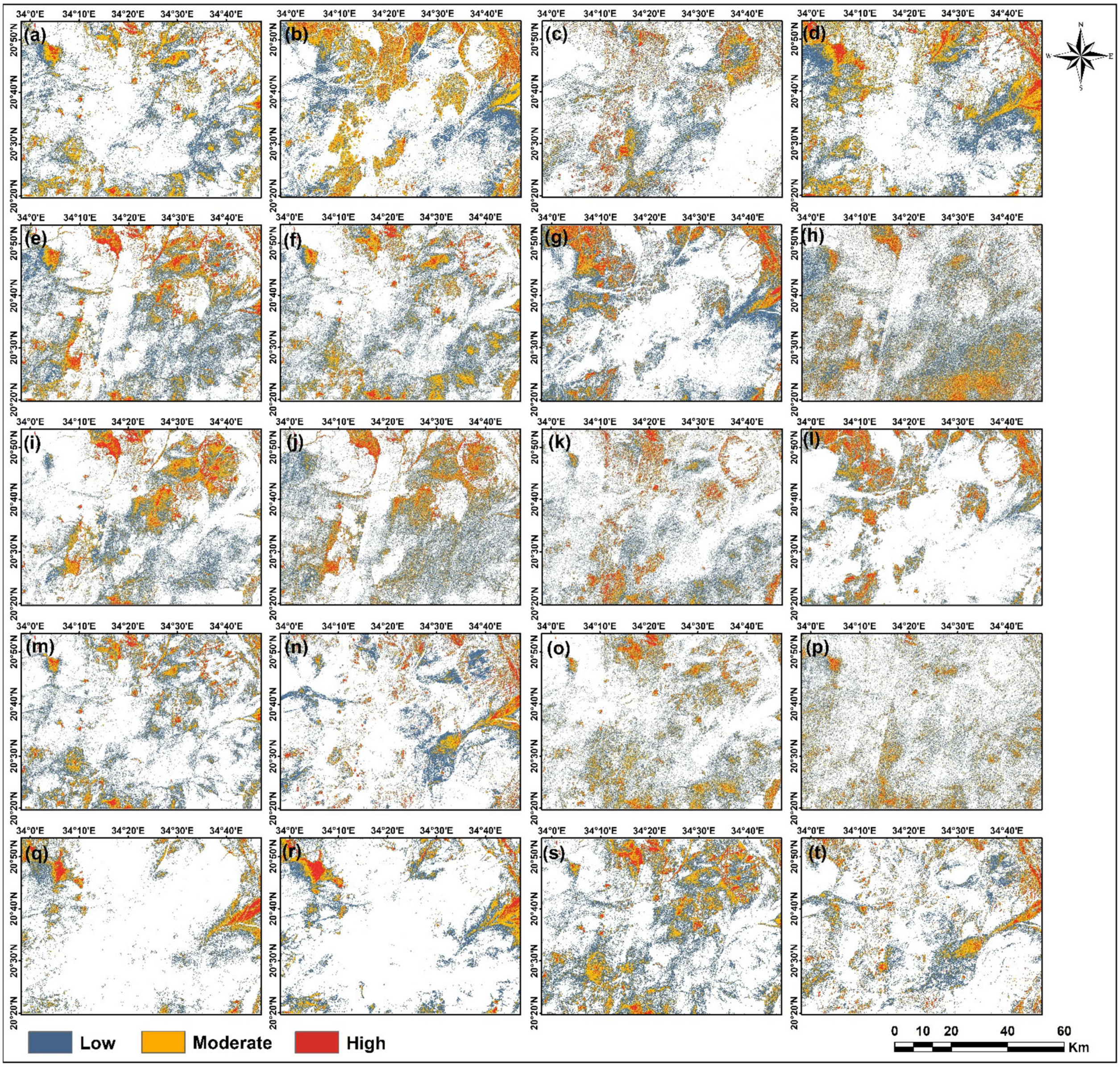

Figure 5. Remote sensing-based predictor variables derived from ASTER data: (**a**) hydroxyl-bearing ''BR 4/6''; (**b**) ferrous iron ''BR (5/3) + (1/2)''; (**c**) ferric iron ''BR 2/1''; (**d**) calcite ''BR 4/7''; (**e**) alunite ''BR 4/5''; (**f**) hydroxyl-bearing ''OHI''; (**g**) calcite ''CLI''; (**h**) alunite ''ALI''; (**i**) kaolinite ''KLI''; (**j**) argillic ''RBD1''; (**k**) phyllic ''RBD2''; (**l**) propylitic ''RBD3''; (**m**) hydroxyl-bearing ''PC4 derived from bands 1, 3, 4, and 6''; (**n**) iron oxides ''PC2 from bands 1, 2, 3, and 4''; (**o**) argillic ''PCA4 from bands 1, 4, 6, and 7''; (**p**) phyllic ''PC4 from bands 1, 3, 5, and 6''; (**q**) propylitic ''PC3 from bands 1, 3, 5, and 8''; (**r**) MNF1; (**s**) MNF2; and (**t**) MNF3.

Arifi, 2021). Hence, they represent structural weaknesses and fracture zones associated with hydrothermal deposits (Pour & Hashim, 2014; Pour et al., 2016; Abd El-Wahed et al., 2021). In the study area, valleys and drainages seem to be structurally controlled. Therefore, we employed Sentinel-2 data to automatically extract lineaments as an indirect indicator for gold deposits. PC bands showed line features better than the original bands. Subsequently, we extracted the lineaments from PC5 using PCI Geomatica software. Thereafter, we generated a density map of lineaments to display the distribution of lineaments in the study area (Fig. 6c).

The contacts of older/younger intrusions and metasediments/metavolcanics became the locus of gold deposits. Moreover, the ring of younger intru-

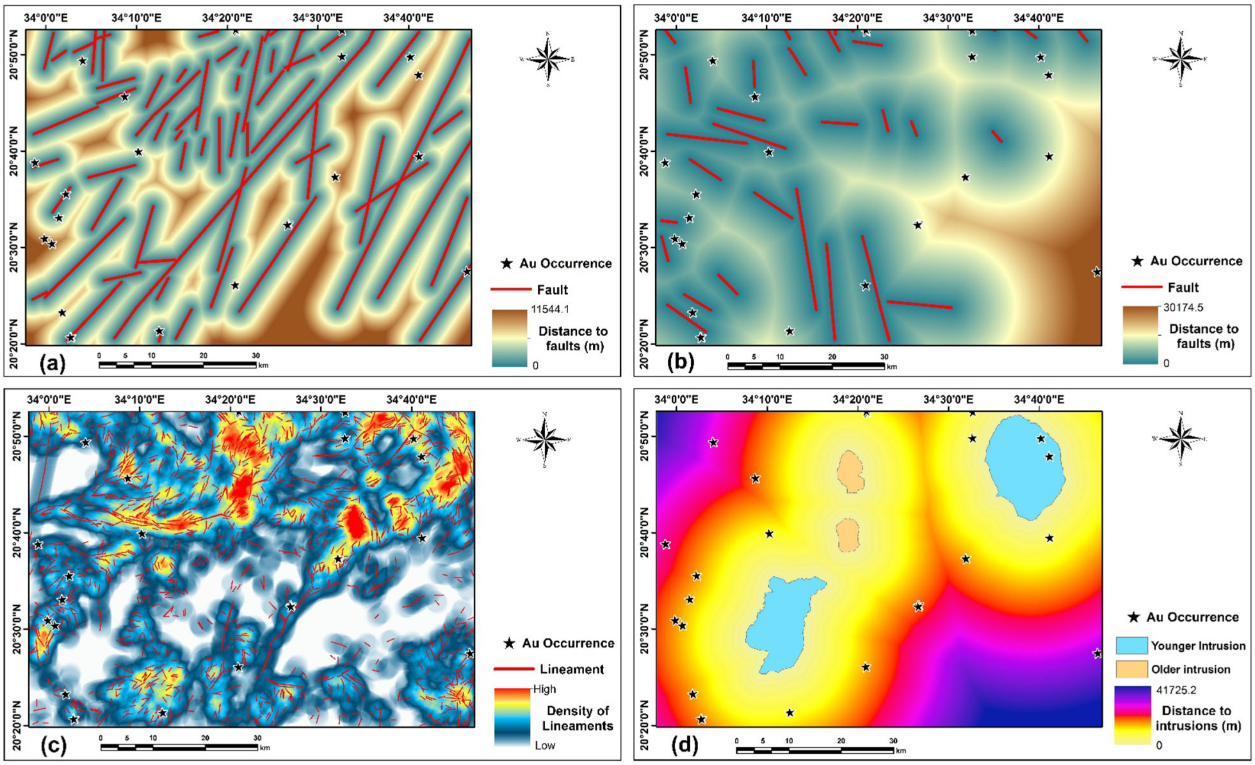

Figure 6. Geological-based predictor variables used in training predictive models: (**a**, **b**) proximity to NE-SW and NW-SE faults, respectively; (**c**) density of lineaments; and (**d**) proximity to intrusions.

sion in the northeastern part is highly sheared and contains several dykes that are associated with different types of alteration minerals (hydroxyl-/iron-bearing minerals). The Euclidean distance method was also employed to produce a predictor map of the intrusive rock contact zone (Fig. 6d).

Target Variables

DF represents a supervised learning model, which requires labeled samples (also called target variables) of the studied phenomena. The target variables in the MPM context are binary values 1 and 0, which represent the locations of mineral occurrences and non-occurrences, respectively. Training and validating the predictive model are carried out using these values. For occurrence locations, we used the locations of 25 Au occurrences. The following processes were used to select 25 samples of non-occurrence sites, which corresponded to a value of 0:

1- Use predictor variables to produce a clustered map, where the locations of non-occurrences are defined given the clusters that do not have agreement with the spatial distribution of gold occurrences (see Mohamed Taha et al. (2023) for details).
2- Apply three selection criteria (Carranza et al., 2008; Carranza & Laborte, 2015a, 2015b): The non-occurrences should be (i) in line with mineral occurrences, (ii) quite distal from the occurrence locations, and (iii) randomly distributed across the study area.

Comparison with the State-of-Art Methods

As demonstrated in a previous MPM study (Mohamed Taha et al., 2023), RF model demonstrated significant performance with ASTER data. In addition to RF results, we also trained two well-known models, namely SVM and ANN, for comparative evaluation against the DF model in terms of

Table 2. Sets of Parameters used for Training ML Predictive Models

Model	Parameter	Range
RF	Number of trees	50, 100, 200, 250, 300, 400, 500
	Number of features	2–12 (at 2 intervals)
SVM	Gamma	0.05–1 (at 0.05 intervals)
	Cost	0.1, 0.5–5 (at 0.5 intervals), 7.5, 10, 20, 25, 30, 40, 50, 75, 100
ANN	Activation function in hidden layers	Sigmoid, ReLU
	Learning rate	0.001, 0.01, 0.1
	Optimizer	Adam, RMSprop
	Batch size	2, 5, 10, 14, 18, 20, 30, 35
	Number of neurons	4, 8, 16, 32, 64, 128
DF	Number of trees	100–1000 (at 100 intervals)
	Number of estimators	2–12 (at 2 intervals)
	Number of bins	2–255 (at 2^n, $n = \ +1$ intervals)
	Minimum sample to split	2–12 (at two intervals)

prediction accuracy. RF is an ensemble learning model that uses the decision tree (DT) as the basic model and outputs the mode of the classes (classification) or the mean prediction (regression) of the individual trees (Breiman, 2001). It utilizes the bootstrapping technique to decorrelate the individual trees and reduce overfitting. SVM is a classifier developed after statistical theory, which aims to define a hyperplane that effectively separates instances in the original feature space into different classes, maximizing the margin between classes (Vapnik, 1999). The parameters of SVM, such as the regularization parameter (C) and the choice of the kernel function, play crucial roles in balancing the trade-off between maximizing the margin and minimizing classification errors. ANN is a network of neurons-based model, where data are processed in a unidirectional flow by interconnected layers of nodes (neurons) (Brown, 2002). The art of designing the architecture of neural networks involves carefully crafting the arrangement of layers, the number of neurons in each layer, and the connections between them, with the aim of optimizing performance for a given task while avoiding overfitting and computational complexity.

Induction of DF Model and Performance Indicators

The DF was developed as a package in the Python programming language (https://deep-forest.readthedocs.io/en/stable). The crucial parameters that affect the DF's performance and must be defined are the number of forests in each cascade layer, the number of trees in each forest, the lowest quantity of samples needed for an internal node splitting, and the number of feature discrete bins. Table 2 illustrates the hyperparameters' range of optimization. Meanwhile, the default setting of 20 was used as the maximum number of cascade levels and 2 for the number of tolerant rounds for the automatic stopping of cascade layers. We trained two DF models, one with a predictor concatenated to the DF, and the other without a predictor. For simplicity, we chose the RF as the model predictor.

To avoid biased comparison resulting from using the parameters' default settings, we sought the optimal parametrization for each model (RF, SVM, and ANN). Each model's optimal value for each of its parameters was assessed by employing a manual grid search procedure. In this step, we used the mean square error (MSE) to evaluate the prediction performance generated from all possible parameter combinations. Based on the suggestion of previous studies (Porwal et al., 2003; Badel et al., 2011; Zuo & Carranza, 2011; Rodriguez-Galiano et al., 2014; Carranza & Laborte, 2015a, 2015b; Rodriguez-Galiano et al., 2015; Sun et al., 2019, 2020b), we defined the values range of key parameters for each model (Table 2). At this stage, the first split ratio (training—70%; testing—30%) was used to train and evaluate different models in terms of prediction and classification accuracies. Utilizing the remaining two split ratios (60–40% and 50–50%), the sensitivity of DF models to the amount of training samples was evaluated.

The performance of MPM models was comprehensively evaluated by the success-rate curve, receiver operating characteristic (ROC) curve, and confusion matrix. The last was employed to evaluate classification accuracy by generating a set of six statistical metrics that accurately elucidate the model results. It is important to mention that the classification accuracy was derived by classifying the prediction result (regression floating value from 0 to 1) to binary values using a threshold value of 0.5, whereby the grids designated as prospective areas were those whose values exceeded the threshold value, while the remaining grids were considered non-prospective. The confusion matrix was calculated by comparing the prospectivity classes between the testing data and the model prediction, which can be categorized into true positive, false positive, true negative, and false negative. Based on these four categories, statistical metrics can be generated including overall accuracy, positive predictive value, negative predictive value, sensitivity, specificity, and Kappa. In the meantime, a thorough assessment of the prediction performance of models was conducted using the ROC curve and success-rate curve. These graphical displays explain the anticipated outcomes in a binary classification system and provide insights into the model's accuracy across a range of discriminating thresholds. By assessing the curves, we gain a nuanced understanding of how the models perform under different thresholds. Details about these performance indicators can be found in Nykänen et al. (2015), Rodriguez-Galiano et al. (2015), Barsi et al. (2018), and Sun et al. (2020b).

RESULTS

Sensitivity of DF Model to Parameter Configuration

Because there is no empirical rule for defining the optimal parameters of any ML model, parameter configuration plays an important role in a model's robustness and error generalization capability. Consequently, training a new model for diverse application backgrounds, such as MPM, becomes more challenging due to the uncertainties and lack of literature defining the acceptable range of values for each key parameter. Figure 7 shows the sensitivity results of MSE obtained by various parameter combinations of the DF, while Table 3 details the significant differences in MSEs among different models. The results indicate that the DF model had robustness that was comparable to RF and outperformed the other models, affirming the error generalization capability of ensemble learning. In general, employing a predictor with the DF model enhanced the accuracy in comparison with DF without predictor, yielding a prediction with an MSE lower than 0.09.

Both DF models, with/without predictor, exhibited similar hyperparameter tuning patterns including the number of bins (4), the number of estimators (2), and the minimum samples to split (2). However, the number of trees differed between the two models: 200 trees in the case of concatenating a predictor and 300 trees in the case of the DF without a predictor (Fig. 7). Although the MSEs during the training of the DF were, on average, lower than those of SVM and ANN, the DF models demonstrated a notable sensitivity to variations in internal configuration. Apart from the optimal values of key parameters, the MSEs exhibited significant fluctuations with different parameter settings. This variability appeared clearly when considering the number of estimators in the case of the DF without predictor (Fig. 7c). Nevertheless, in the context of prospectivity modeling, the DF models in this study showcased high performance with relatively uncomplicated architecture. Increasing those parameters related to the model architecture (e.g., the number of estimators in each cascade layer and the number of trees in each estimator) did not necessarily result in an observable decreasing trend in MSE. This indicates that complex models, with high computing costs, do not conclusively result in more accurate predictions. The observed outcomes can be explained by the relatively modest size of training datasets, which usually comprise a few dozen locations of both non-deposit and deposits. Such limited datasets make it easy to train models and achieve the requisite precision; however, the risk of overfitting increases when employing complex architectures and conducting excessive training.

Performance Assessment of DF Predictive Modeling

Figure 8 illustrates predictive maps accompanied by positive training samples of gold occurrences. These maps display the likelihood of gold prediction generated by ML models trained using the best parameter combinations. Because the likelihood scores of gold are floating values from 0 to 1, prospective areas were identified by ML output values greater than 0.5, while non-prospective tracts are identified by values less than or equal to 0.5. According to such a classification scheme, the clas-

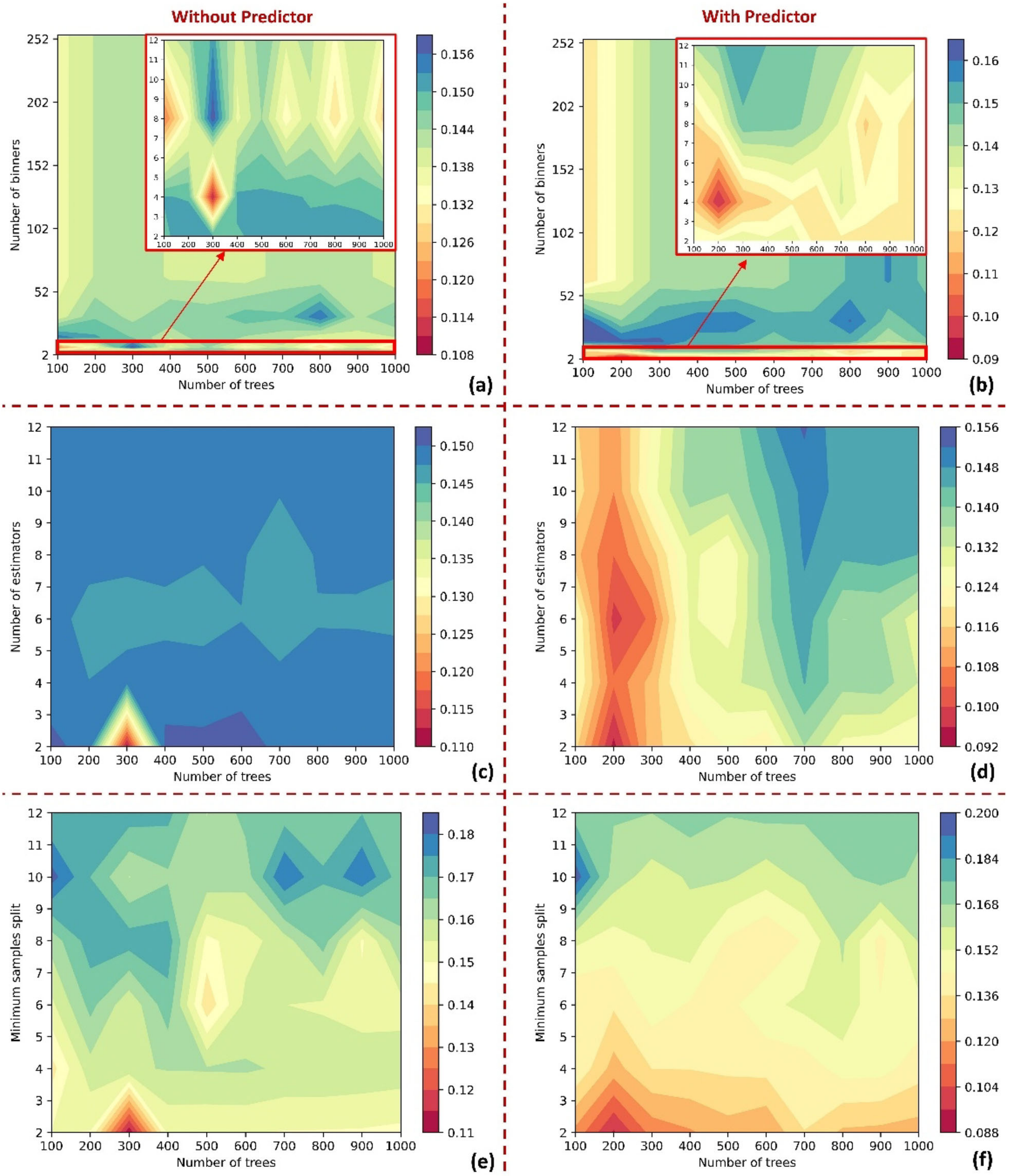

Figure 7. MSE-based mapping accuracy for every combination of parameters used during DF model training: left panel—DF without predictor and right panel—DF with predictor.

sification reports of the confusion matrix derived from the 30% testing dataset are listed for all ML models in Table 4. Based on all six statistic indicators, the DF notably outperformed the rest of the models. Except for the specificity value, which is equivalent to that of the ANN model (85.7%), the other indicators exhibited significantly higher values. Specifically, the overall accuracy of the DF model was 93.3%, approximately 13% higher than the subsequent lower-performing model. The sensitivity and negative prediction value were almost 100%, indicating that the DF model identified all the occurrence locations and the predicted grids as non-occurrence are actually non-occurrence areas. The results of RF and SVM were similar, which yielded the worst classification accuracy.

Table 3. MSEs Showing the Accuracy of Each Predictive Model

	RF	SVM	ANN	DF—with P	DF—without P
Min	0.096	0.134	0.124	0.095	0.110
Max	0.117	0.396	0.301	0.197	0.182
Avg	0.103	0.185	0.189	0.152	0.162

To further evaluate the gold prospectivity map derived by the DF model, both the success-rate curves and the ROC curve efficiently evaluated the prediction accuracy of high-confidence zones. The ideal scenario of the ROC curve is a straight line corresponding to an area under the curve (AUC) of 1, which indicates a probability of the occurrence's

Table 4. Classification Report for Every Predictive Model

Indicator	RF	SVM	ANN	DF
Sensitivity (%)	73.2	73.2	80.4	99.8
Specificity (%)	71.4	71.4	85.7	85.7
Positive predictive values (%)	73.2	73.2	80.4	88.9
Negative predictive values (%)	71.4	71.4	75	99.8
Overall accuracy (%)	73.3	73.3	80	93.3
Kappa	0.646	0.646	0.602	0.865

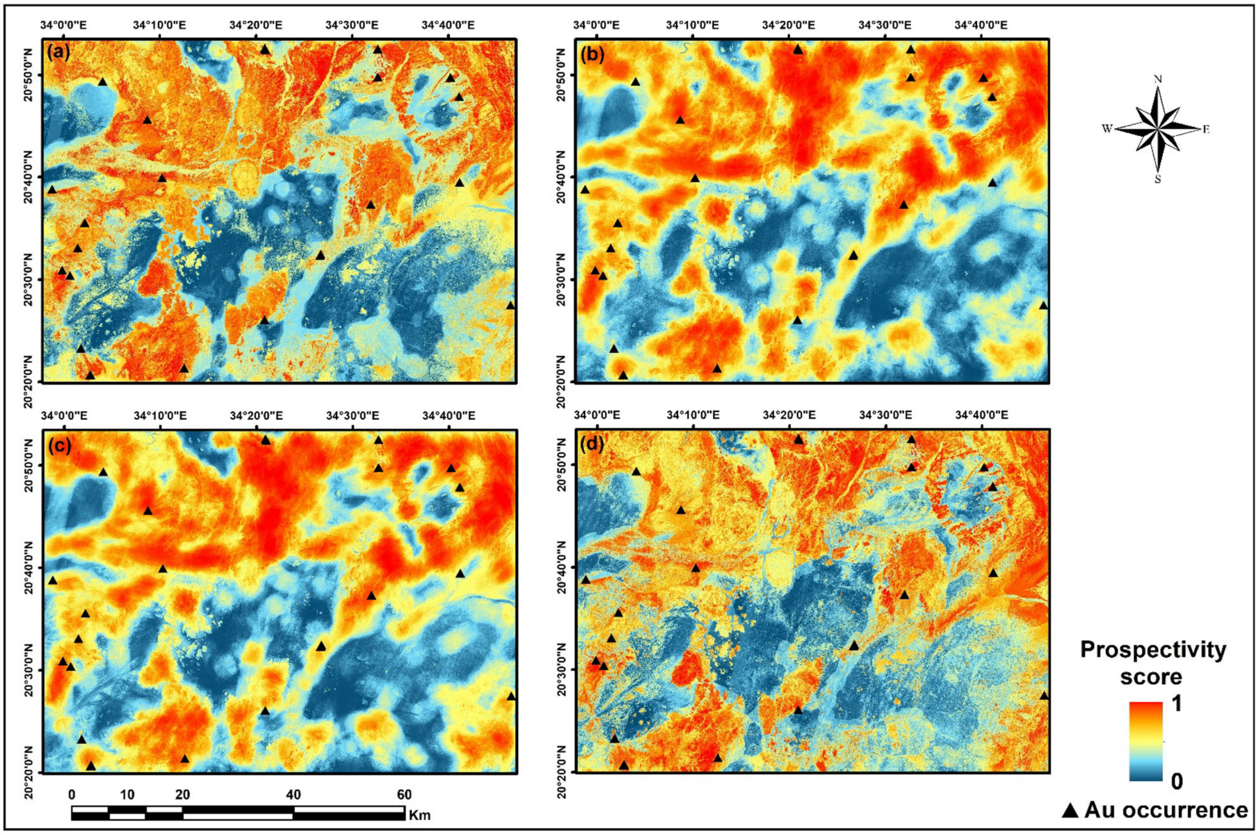

Figure 8. ML-based predictive maps of gold prospectivity obtained using (**a**) RF, (**b**) SVM, (**c**) ANN, and (**d**) DF.

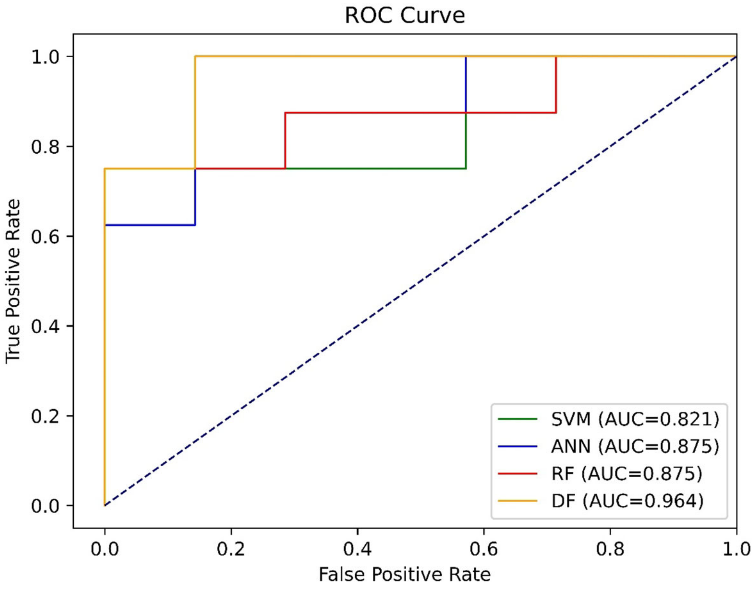

Figure 9. ROC curves and AUC values for every ML model.

samples compared to those of non-occurrence. Contrariwise, the AUC values decrease in situations when there is a high probability of the non-occurrence grid or a low probability of the occurrence grid. Figure 9 shows the ROC curves and the AUC values of the four trained models. There was good performance of RF and ANN with similar AUC values of 0.875, while the SVM was less accurate. However, the ROC results further indicate the superiority of the DF in prediction performance over the rest of the methods with an AUC value reaching up to 0.9647.

The success-rate curve shows how known mineral occurrences are estimated based on various percentages of potential regions created by applying changing threshold values. In this regard, the different slopes of the success-rate curve straightforwardly indicate different prospective zones (high, moderate, and low). Subsequently, the model that captures a higher percentage of mineral occurrence in the smallest possible region has more prediction capability, corresponding to a steeper curve. The success-rate curves of the different ML models and the three regression lines with different slopes that classify the DF prospectivity map are shown in Figure 10. The success-rate curve of the DF map was closer to the upper left corner, which indicates the high ability to capture all the gold occurrences in a smaller area (26.6%) compared to the rest of the models. Thus, other ML models needed larger prospective regions (RF—37.8%, SVM—61.5%, and ANN—60.8%) to reach a success rate similar to of the DF model. It can be observed how the success-rate curves of the DF and RF models started similarly and steeper than the SVM and ANN models, capturing about 60% of gold occurrences in less than 10% prospective area.

Figure 11 shows the DF classified map of different gold prospectivity zones based on the thresholds derived from the three regression lines of the success-rate curve. From a visual point of view, the most favorable tracts revealed by the modeling results are roughly concentrated in the northern region of the study area. These tracts are spatially distributed around the syn-tectonic intrusions in the upper middle part and around the post-tectonic intrusions in the upper right corner and the lower middle part. Beside several distributed bodies (i.e., dykes) with NE–SW trends, other favorable tracts are mainly distributed in the metasediment and metavolcanic units.

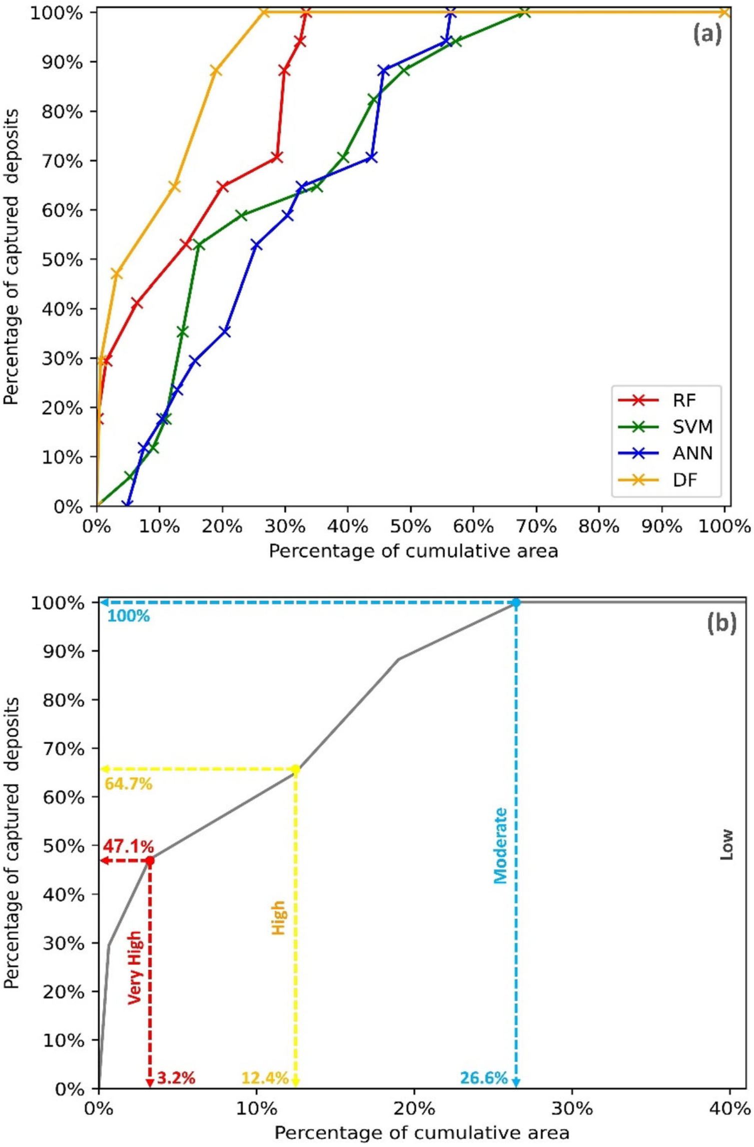

Figure 10. (**a**) Success-rate curves of all predictive models and (**b**) success-rate curve of DF model.

Impact of Training Sample Sizes on DF Model

As mentioned previously, two different split ratios were used to test the sensitivity of the DF models to the reduction in the training sample size. The results are shown in Figure 12. The initial hypothesis was that the models would react gradually to the target variable reduction. However, it can be observed that the initial reduction of 10% unexpectedly increased the accuracy for both models with/without predictor. Although the OA slightly dropped (90%) in the case of the DF without predictor, better results of MSE and AUC took place compared to the original split ratio (70–30%). For the second split ratio (50–50%), corresponding to 12 positive occurrences, both the classification and the prediction accuracies experienced a decrease, reaching 84% OA, 93.5% AUC, and 0.136 MSE in the case of the DF with predictor, which marked the least favorable outcomes among the DF models. Nevertheless, the observed outcomes persisted in surpassing those of other models derived from

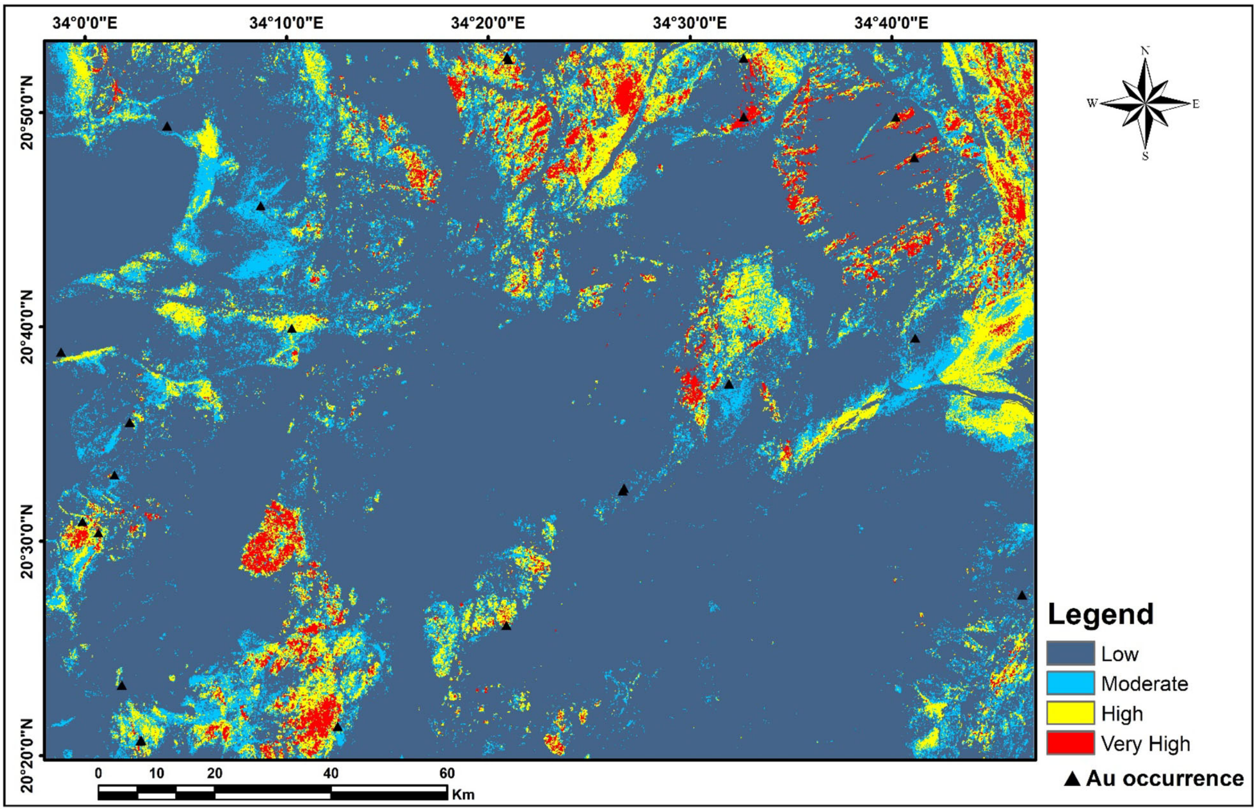

Figure 11. DF classified map of gold prospectivity using success-rate curve threshold values.

training samples without data reduction. These results denote that the DF model can adequately do deep learning with automatic ability to decide model complexity, which makes it more convenient for MPM than classical ML or DNN methods.

DISCUSSION

One of the key applications of remote sensing data is geological investigation in general and mineral mapping in particular. Several factors impede the applicability of utilizing remote sensing data in data-driven MPM: (i) judgment-related: the difficulty of specifying a mineral deposit model or selecting targeting criteria in the early stage of exploration; (ii) data-related: the capability of remote sensing data to produce accurate mappable layers of the ore-forming factors as well as the insufficient number of know deposits/occurrences; and (iii) model-related: the capacity of the algorithm of choice to learn complex relations between the predictor maps and generate an accurate prediction. To comprehensively analyze these factors and discuss the feasibility of the current study, it is important to comprehend the workflows of both the MPM and remote sensing mineral mapping.

Traditionally, the MPM workflow involves defining a conceptual model of a mineral deposit, which therefore controls the selection of geoscience spatial data and the creation of predictor maps. The methods for studying the correlation between the predictor maps and the locations of mineral deposits vary between quantitative analysis (data-driven methods) and expert judgment (knowledge-driven methods). Recently, a workflow of MPM based on Geodata Science (GSMPM) has been proposed, which either mines the original geoscience data or utilizes them to statistically analyze their spatial correlation with mineral deposits. The initial goal of remote sensing mineral mapping is to identify the spatial distribution of dominant minerals (specifically altered minerals), which makes the latter field investigation more convenient (i.e., geoscience in-

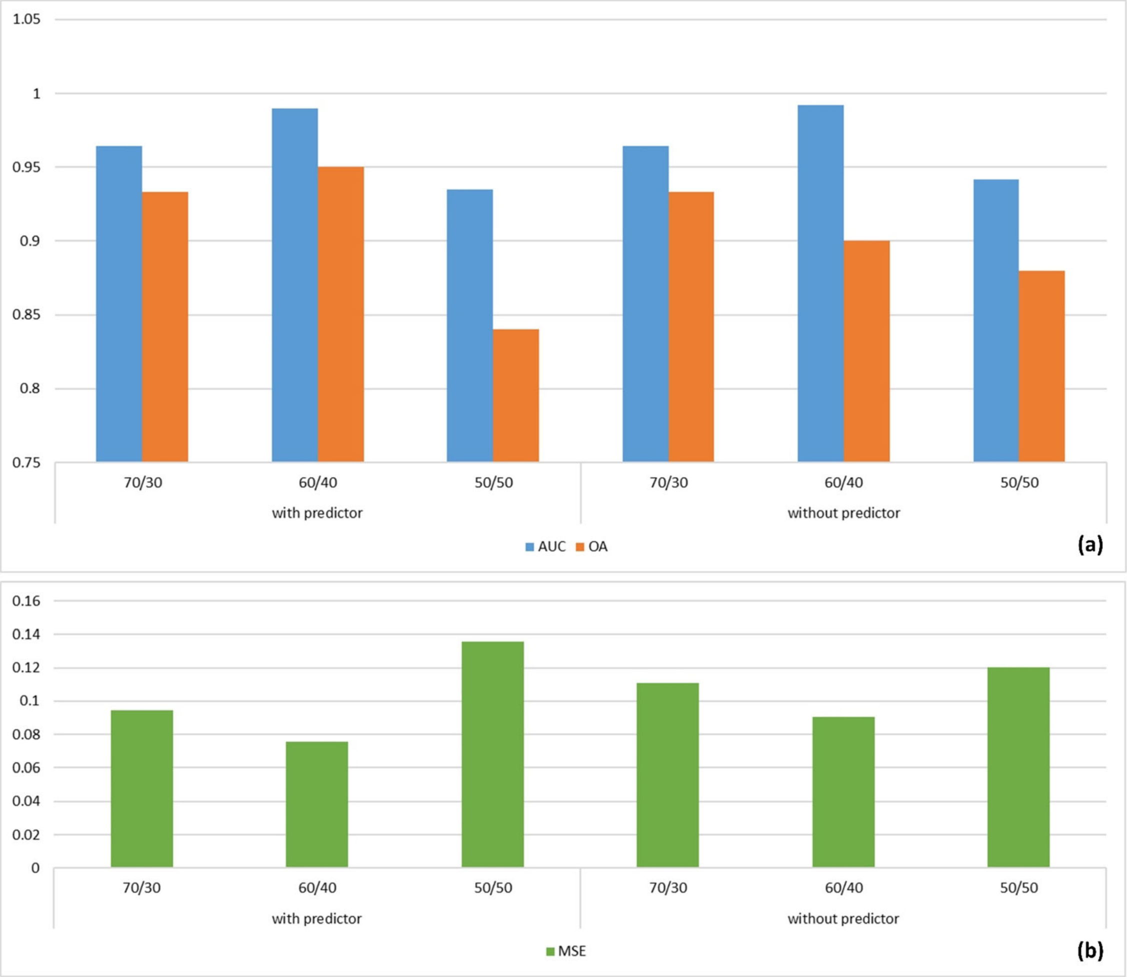

Figure 12. Effect of training sample size on mapping accuracy of DF predictive models.

sight into mineral deposits). However, mineral abundance maps were barely used in further regression analysis (generating continuous values for a specific mineral). The utilization of these maps mostly focuses on classification tasks to generate altered mineral classified maps as the final product. In the meantime, fuzzy logic was the most utilized method with remotely sensed data, which is regarded as knowledge-driven MPM. Because remote sensing data have spectrum characteristics, several ML and DL models were developed to study the relationship between the spectral features and the desired targets. Various DNN models achieved high accuracy in identifying altered minerals using multispectral/hyperspectral satellite data or indoor reflectance spectra (Gewali et al., 2018; Holloway & Mengersen, 2018; Tanaka et al., 2019; Zhang et al., 2022). Such studies can help the future work of GSMPM and geological prospecting big data (GPBD).

In the current study, we attempted to enlighten the feasibility of data-driven MPM based on remote sensing data by addressing some of the aforementioned issues. Despite the advancements in studying mineral deposit models, our comprehension of ore-forming processes remains limited and imperfect. Such comprehension varies across different scales of geological investigation (regional or detailed) and keeps changing based on the updates derived from observations and measures. Remote sensing data can re-express the targeting criteria of several mineral deposit models at different times and scales. Mineral deposit models characterized by hydrothermal alteration or structures as significant targeting criteria can be studied effectively using remote sensing data. Similar to the way of processing geochemical and geophysical data for mapping targeting criteria using interpolation and transformation methods, we also emphasized the need for utilizing remote sensing enhancement techniques rather than using original spectra information. We tried to use objective methods such as BR, RBD, mineral indices, and PCA, or carefully utilized subjective methods such as MNF and lineaments extraction. We successfully generated 20 predictor maps from ASTER data representing different hydrothermal alteration zones (argillic, phyllic, and propylitic) and altered minerals (hydroxyl-bearing and iron oxides), and one map from Sentinel-2 represents tectonic lineaments. Although some of the predictor maps have relatively similar spatial distributions of minerals, the histograms of these maps have different shapes and intensity distributions, which increases the statistical variation, allowing the model to capture distinct information from each map. This clearly appears in the feature importance analysis of RF model, where the importance values vary among these similar maps (see Mohamed Taha et al. (2023), Fig. 16). Overall, data quality and availability determine the successful implementation of data-driven MPM. Compared to other geoscience data, remote sensing data are the most available at different times with the ability to express multiple geological features simultaneously including minerals, lithology, and structures.

The present study fundamentally addressed the ongoing necessity to introduce new robust methods that are adaptive and efficient (model-related issues). To overcome the limited number of known deposits, it is well-known that ensemble classifiers or cost-sensitive learners can be applied (Zhou & Liu, 2005; Zuo, 2020). Following the ensemble learning approach, which posits that multiple predictors usually outperform a single predictor, we proposed the DF ensemble model for data-driven MPM. The DF model has two ensembling ranks (ensemble–ensemble structure), the first rank is the ensemble of decision trees inside the forest and the second is the ensemble of forests that form the deep forest. By introducing diversity (i.e., using two types of forests, CRF, and RF) and cascade layers arrangement, the structure of the DF model ensures the deep learning characteristics. Such a structure brings more advantages than those in DNN models including robustness to overfitting (especially when dealing with small or noisy datasets), scalability, feature learning, efficient training (in both computational resources and training time), and interpretability.

The results of the study show that the simplified version of the DF with hyperparameters tuning had superiority over RF, ANN, and SVM, even though the DF model was trained using 50% of the training samples. Although the simplified DF model achieved an excellent performance, the DF has some limitations and various architectures and techniques can be applied to increase its performance. Several studies reported the difficulty of implementing the DF model when training large datasets because of the high consumption of memory and time (Pang et al., 2018; Ma et al., 2022a). The reason for this is that multi-grain scanning (though not applied in this study) generates a large number of instances, and all these instances should pass all the cascade layers up to the final prediction. Different models were proposed to overcome that shortage such as the DF with different screening techniques (confidence, hashing, and window screening), adaptive weighted DF, or multi-label learning DF (Pang et al., 2018; Sun et al., 2020a; Ma et al., 2022a, 2022b). Although not having a backpropagation procedure is one of the DF advantages, it could be argued that finding a mechanism to adjust the initial parameters of each forest (similar to adjusting weights in a neural network) can lead to better performance. Similar to the state-of-art architectures of DNN, the structure of the DF model can be utilized with different architectures and various basic classifiers (SVM, ANN, RF, and CRF), which could pave the way for deep ensemble learning approaches beyond traditional neural networks, expanding the possibility of non-neural network deep learning.

CONCLUSIONS AND FUTURE WORK

In data-driven predictive MPM, ML techniques, especially deep learning ones, necessitate a large number of training occurrence/deposit locations (e.g., > 15). However, as proposed in this study, the DF algorithm, which is a novel tree-based ensemble model recognized by representation learning and large model capacity, can be used in remote sensing-based data-driven MPM. Based on the experimental results of gold prospectivity in the Hamissana area, NE Sudan, we conclude that our DF model, trained with a few training locations, achieved promising performance that surpassed the results of conventional ML models, including SVM, RF, and ANN.

As ongoing research, the DF can emerge as a potential alternative to the neural networks for MPM studies. Therefore, further validation is essential, involving testing the deep forest model in a standard MPM scenario with access to multisource geoscience big data. In such a case, a comparative analysis will be conducted against a neural network deep learning model, such as a CNN. Moreover, the feasibility of employing the DF algorithm for remote sensing MPM studies needs further evaluation, extending to diverse geographical areas and varying mineral deposit types.

ACKNOWLEDGMENTS

We express our gratitude to the editors and anonymous reviewers for their insightful feedback on how to make our work better. The funding for this study was provided by the National Natural Science Foundation of China and the Science and Technology Strategic Prospecting Project of Guizhou Province.

FUNDING

The National Natural Science Foundation of China (42372345, 42172333) and the Guizhou Province's Science and Technology Strategic Prospecting Project ([2022] ZD003) provided funding for this work.

DECLARATIONS

Conflict of Interest The authors have no known conflicts of interest to declare associated with this publication.

REFERENCES

Abd El-Wahed, M., Zoheir, B., Pour, A. B., & Kamh, S. (2021). Shear-related gold ores in the Wadi Hodein Shear Belt, south eastern desert of Egypt: Analysis of remote sensing. *Field and Structural Data. Minerals, 11*(5), 474.

Abdelkareem, M., & Al-Arifi, N. (2021). Synergy of remote sensing data for exploring hydrothermal mineral resources using GIS-based fuzzy logic approach. *Remote Sensing, 13*(22), 4494.

Abedi, M., Norouzi, G.-H., & Fathianpour, N. (2013). Fuzzy outranking approach: A knowledge-driven method for mineral prospectivity mapping. *International Journal of Applied Earth Observation and Geoinformation, 21*, 556–567.

Abedini, M., Ziaii, M., Timkin, T., & Pour, A. B. (2023). Machine learning (ML)-based copper mineralization prospectivity mapping (MPM) using mining geochemistry method and remote sensing satellite data. *Remote Sensing, 15*(15), 3708.

Ahmed, A. H. (2022). *Mineral deposits and occurrences in the Arabian-Nubian shield*. Springer.

Ali, A., & Pour, A. (2014). Lithological mapping and hydrothermal alteration using Landsat 8 data: A case study in ariab mining district, red sea hills, Sudan. *International Journal of Basic and Applied Sciences, 3*(3), 199–208.

Badel, M., Angorani, S., & Shariat Panahi, M. (2011). The application of median indicator kriging and neural network in modeling mixed population in an iron ore deposit. *Computers & Geosciences, 37*(4), 530–540.

Bahrami, Y., Hassani, H., & Maghsoudi, A. (2018). Investigating the capabilities of multispectral remote sensors data to map alteration zones in the Abhar area, NW Iran. *Geosystem Engineering, 24*(1), 18–30.

Barsi, Á., Kugler, Z., László, I., Szabó, G., & Abdulmutalib, H. M. (2018). Accuracy dimensions in remote sensing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII–3*, 61–67.

Bierlein, F. P., McKeag, S., Reynolds, N., Bargmann, C. J., Bullen, W., Murphy, F. C., Al-Athbah, H., Brauhart, C., Potma, W., Meffre, S., & McKnight, S. (2015). The Jebel Ohier deposit—a newly discovered porphyry copper–gold system in the Neoproterozoic Arabian-Nubian Shield, Red Sea Hills, NE Sudan. *Mineralium Deposita, 51*(6), 713–724.

Bolouki, S. M., Ramazi, H. R., Maghsoudi, A., Beiranvand Pour, A., & Sohrabi, G. (2019). A remote sensing-based application of Bayesian networks for epithermal gold potential mapping in Ahar-Arasbaran Area, NW Iran. *Remote Sensing, 12*(1), 105.

Bonham-Carter, G. (1994a). *Geographic information systems for geoscientists: Modelling with GIS*. Elsevier.

Bonham-Carter, G., & Chung, C. (1983). Integration of mineral resource data for Kasmere Lake area, Northwest Manitoba, with emphasis on uranium. *Journal of the International Association for Mathematical Geology, 15*, 25–45.

Bonham-Carter, G. F. (1994b). *Geographic information systems for geoscientists: Modelling with GIS*. Pergamon.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(5), 5–32.

Brown, W. (2002). Artificial neural network: a new method for mineral prospectivity mapping. In U. o. W. Australia (Ed.).

Brown, W., Groves, D., & Gedeon, T. A. (2003). Use of fuzzy membership input layers to combine subjective geological knowledge and empirical data in a neural network method for mineral-potential mapping. *Natural Resources Research, 12*(3), 183–200.

Carranza, E. J. M. (2009). *Geochemical anomaly and mineral prospectivity mapping in GIS* (Vol. 11). Elsevier.

Carranza, E. J. M. (2017). Natural resources research publications on geochemical anomaly and mineral potential mapping, and introduction to the special issue of papers in these fields. *Natural Resources Research, 26*(4), 379–410.

Carranza, E. J. M., & Hale, M. (2003). Evidential belief functions for data-driven geologically constrained mapping of gold potential, Baguio district, Philippines. *Ore Geology Reviews, 22*(1–2), 117–132.

Carranza, E. J. M., Hale, M., & Faassen, C. (2008). Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. *Ore Geology Reviews, 33*(3–4), 536–558.

Carranza, E. J. M., & Laborte, A. G. (2015a). Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. *Ore Geology Reviews, 71*, 777–787.

Carranza, E. J. M., & Laborte, A. G. (2015b). Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences, 74*, 60–70.

Carranza, E. J. M., & Laborte, A. G. (2016). Data-driven predictive modeling of mineral prospectivity using random forests: a case study in Catanduanes Island (Philippines). *Natural Resources Research, 25*(1), 35–50.

Chen, C., Dai, H., Liu, Y., & He, B. (2011). Mineral prospectivity mapping integrating multisource geology spatial data sets and logistic regression modelling. In *The 2011 IEEE international conference on spatial data mining and geographic knowledge series (ICSDM), Fuzhou, China*.

Cheng, Q. M., & Agterberg, F. P. (1999). Fuzzy weights of evidence method and its application in mineral potential mapping. *Natural Resources Research, 8*, 27–35.

Crosta, A., De Souza Filho, C., Azevedo, F., & Brodie, C. (2003). Targeting key alteration minerals in epithermal deposits in Patagonia, Argentina, using ASTER imagery and principal component analysis. *International Journal of Remote Sensing, 24*(21), 4233–4240.

El Khidir, S. O., & Babikir, I. A. (2013). Digital image processing and geospatial analysis of landsat 7 ETM+ for mineral exploration, Abidiya area, North Sudan. *International Journal of Geomatics and Geosciences, 3*(3), 645–658.

Forson, E. D., Wemegah, D. D., Hagan, G. B., Appiah, D., Addo-Wuver, F., Adjovu, I., Otchere, F. O., Mateso, S., Menyeh, A., & Amponsah, T. (2022). Data-driven multi-index overlay gold prospectivity mapping using geophysical and remote sensing datasets. *Journal of African Earth Sciences, 190*, 104504.

Fu, Y., Cheng, Q., Jing, L., Ye, B., & Fu, H. (2023). Mineral prospectivity mapping of porphyry copper deposits based on remote sensing imagery and geochemical data in the Duolong Ore District, Tibet. *Remote Sensing, 15*(2), 439.

Gewali, U. B., Monteiro, S. T., & Saber, E. (2018). Machine learning based hyperspectral image analysis: a survey. *arXiv preprint* https://arxiv.org/abs/1802.08701.

Hamimi, Z., Fowler, A.-R., Liégeois, J.-P., Collins, A., Abdelsalam, M. G., & Abd El-Wahed, M. (2021). *The geology of the Arabian-Nubian Shield*. Springer.

Harris, D., & Pan, G. (1999). Mineral favorability mapping: a comparison of artificial neural networks, logistic regression, and discriminant analysis. *Natural Resources Research, 8*, 93–109.

Harris, J., Wilkinson, L., Heather, K., Fumerton, S., Bernier, M., Ayer, J., & Dahn, R. (2001). Application of GIS processing techniques for producing mineral prospectivity maps—a case study: mesothermal Au in the Swayze Greenstone Belt, Ontario, Canada. *Natural Resources Research, 10*, 91–124.

He, B., Chen, C., & Liu, Y. (2010). Gold resources potential assessment in eastern Kunlun Mountains of China combining weights-of-evidence model with GIS spatial analysis technique. *Chinese Geographical Science, 20*(5), 461–470.

He, L., Lyu, P., He, Z., Zhou, J., Hui, B., Ye, Y., Hu, H., Zeng, Y., & Xu, L. (2022). Identification of radioactive mineralized lithology and mineral prospectivity mapping based on remote sensing in high-latitude regions: a case study on the Narsaq Region of Greenland. *Minerals, 12*(6), 692.

Holloway, J., & Mengersen, K. (2018). Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sensing, 10*(9), 1365.

Hubbard, B. E., & Crowley, J. K. (2005). Mineral mapping on the Chilean-Bolivian Altiplano using co-orbital ALI, ASTER and Hyperion imagery: Data dimensionality issues and solutions. *Remote Sensing of Environment, 99*(1–2), 173–186.

Inzana, J., Kusky, T., Higgs, G., & Tucker, R. (2003). Supervised classifications of Landsat TM band ratio images and Landsat TM band ratio image with radar for geological interpretations of central Madagascar. *Journal of African Earth Sciences, 37*(1–2), 59–72.

Johnson, P., Zoheir, B., Ghebreab, W., Stern, R., Barrie, C., & Hamer, R. (2017). Gold-bearing volcanogenic massive sulfides and orogenic-gold deposits in the Nubian Shield. *South African Journal of Geology, 120*(1), 63–76.

Kashani, S. B. M., Abedi, M., & Norouzi, G.-H. (2016). Fuzzy logic mineral potential mapping for copper exploration using multi-disciplinary geo-datasets, a case study in seridune deposit, Iran. *Earth Science Informatics, 9*(2), 167–181.

Li, D., Liu, Z., Armaghani, D. J., Xiao, P., & Zhou, J. (2022). Novel ensemble tree solution for rockburst prediction using deep forest. *Mathematics, 10*(5), 787.

Li, H., Li, X., Yuan, F., Jowitt, S. M., Zhang, M., Zhou, J., Zhou, T., Li, X., Ge, C., & Wu, B. (2020a). Convolutional neural network and transfer learning based mineral prospectivity modeling for geochemical exploration of Au mineralization within the Guandian-Zhangbaling area, Anhui Province, China. *Applied Geochemistry, 122*, 104747.

Li, T., Zuo, R., Xiong, Y., & Peng, Y. (2020b). Random-drop data augmentation of deep convolutional neural network for mineral prospectivity mapping. *Natural Resources Research, 30*(1), 27–38.

Loughlin, W. (1991). Principal component analysis for alteration mapping. *Photogrammetric Engineering and Remote Sensing, 57*(9), 1163–1169.

Ma, P., Wu, Y., Li, Y., Guo, L., Jiang, H., Zhu, X., & Wu, X. (2022a). HW-forest: Deep forest with hashing screening and window screening. *ACM Transactions on Knowledge Discovery from Data, 16*(6), 1–24.

Ma, P., Wu, Y., Li, Y., Guo, L., & Li, Z. (2022b). DBC-Forest: Deep forest with binning confidence screening. *Neurocomputing, 475*, 112–122.

Magalhães, L. A., & Souza Filho, C. R. (2012). Targeting of gold deposits in Amazonian exploration frontiers using knowledge-and data-driven spatial modeling of geophysical, geochemical, and geological data. *Surveys in Geophysics, 33*(2), 211–241.

Mahdevar, M. R., Ketabi, P., Saadatkhah, N., Rahnamarad, J., & Mohammadi, S. S. (2014). Application of ASTER SWIR data on detection of alteration zone in the Sheikhabad area, eastern Iran. *Arabian Journal of Geosciences, 8*(8), 5909–5919.

Mohamed, M. T. A., Al-Naimi, L. S., Mgbeojedo, T. I., & Agoha, C. C. (2021). Geological mapping and mineral prospectivity using remote sensing and GIS in parts of Hamissana, Northeast Sudan. *Journal of Petroleum Exploration and Production, 11*, 1123–1138.

Mohamed Taha, A. M., Xi, Y., He, Q., Hu, A., Wang, S., & Liu, X. (2023). Investigating the capabilities of various multispectral remote sensors data to map mineral prospectivity based on random forest predictive model: A case study for gold deposits in Hamissana Area. *NE Sudan. Minerals, 13*(1), 49.

Moore, F., Rastmanesh, F., Asadi, H., & Modabberi, S. (2008). Mapping mineralogical alteration using principal-component analysis and matched filter processing in the Takab area, north-west Iran, from ASTER data. *International Journal of Remote Sensing, 29*(10), 2851–2867.

Ngassam Mbianya, G., Ngnotue, T., Takodjou Wambo, J. D., Ganno, S., Pour, A. B., Ayonta Kenne, P., Fossi, D. H., & Wolf, I. D. (2021). Remote sensing satellite-based structural/alteration mapping for gold exploration in the Ketté goldfield, Eastern Cameroon. *Journal of African Earth Sciences, 184*, 104386.

Ninomiya, Y. (2003). A stabilized vegetation index and several mineralogic indices defined for ASTER VNIR and SWIR data IGARSS 2003. In *2003 IEEE international geoscience and remote sensing symposium. proceedings (IEEE Cat. No.03CH37477)*.

Nykänen, V., Lahti, I., Niiranen, T., & Korhonen, K. (2015). Receiver operating characteristics (ROC) as validation tool for prospectivity models—A magmatic Ni–Cu case study from the Central Lapland Greenstone Belt, Northern Finland. *Ore Geology Reviews, 71*, 853–860.

Pang, M., Ting, K.-M., Zhao, P., & Zhou, Z.-H. (2018). Improving deep forest by confidence screening. In *2018 IEEE international conference on data mining (ICDM)*.

Parsa, M. (2021). A data augmentation approach to XGboost-based mineral potential mapping: An example of carbonate-hosted Zn Pb mineral systems of Western Iran. *Journal of Geochemical Exploration, 228*, 106811.

Perret, J., Feneyrol, J., Eglinger, A., André-Mayer, A.-S., Berthier, C., Ennaciri, A., & Bosc, R. (2021). Tectonic record and gold mineralization in the central part of the Neoproterozoic Keraf suture, Gabgaba district, NE Sudan. *Journal of African Earth Sciences, 181*, 104248.

Porwal, A., Carranza, E. J. M., & Hale, M. (2003). Artificial neural networks for mineral-potential mapping: A case study from Aravalli Province, Western India. *Natural Resources Research, 12*(3), 155–171.

Pour, A. B., & Hashim, M. (2011). Identification of hydrothermal alteration minerals for exploring of porphyry copper deposit using ASTER data, SE Iran. *Journal of Asian Earth Sciences, 42*(6), 1309–1323.

Pour, A. B., & Hashim, M. (2012). Identifying areas of high economic-potential copper mineralization using ASTER data in the Urumieh-Dokhtar Volcanic Belt. *Iran. Advances in Space Research, 49*(4), 753–769.

Pour, A. B., & Hashim, M. (2014). ASTER, ALI and Hyperion sensors data for lithological mapping and ore minerals exploration. *Springerplus, 3*(1), 130.

Pour, A. B., Hashim, M., Makoundi, C., & Zaw, K. (2016). Structural mapping of the Bentong-Raub Suture zone using PALSAR remote sensing data, Peninsular Malaysia: Implications for sediment-hosted/orogenic gold mineral systems exploration. *Resource Geology, 66*(4), 368–385.

Pour, A. B., Park, Y., Park, T.-Y.S., Hong, J. K., Hashim, M., Woo, J., & Ayoobi, I. (2018). Regional geology mapping using satellite-based remote sensing approach in Northern Victoria Land, Antarctica. *Polar Science, 16*, 23–46.

Rajan Girija, R., & Mayappan, S. (2019). Mapping of mineral resources and lithological units: A review of remote sensing techniques. *International Journal of Image and Data Fusion, 10*(2), 79–106.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). 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.

Rodriguez-Galiano, V. F., Chica-Olmo, M., & Chica-Rivas, M. (2014). Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain. *International Journal of Geographical Information Science, 28*(7), 1336–1354.

Sabins, F. F. (1999). Remote sensing for mineral exploration. *Ore Geology Reviews, 14*, 157–183.

Senanayake, I. P., Kiem, A. S., Hancock, G. R., Metelka, V., Folkes, C. B., Blevin, P. L., & Budd, A. R. (2023). A spatial data-driven approach for mineral prospectivity mapping. *Remote Sensing, 15*(16), 4074.

Shirmard, H., Farahbakhsh, E., Müller, R. D., & Chandra, R. (2022). A review of machine learning in processing remote sensing data for mineral exploration. *Remote Sensing of Environment, 268*, 112750.

Silva dos Santos, V., Gloaguen, E., Hector Abud Louro, V., & Blouin, M. (2022). Machine learning methods for quantifying uncertainty in prospectivity mapping of magmatic-hydrothermal gold deposits: A case study from Juruena Mineral Province, Northern Mato Grosso, Brazil. *Minerals, 12*(8), 941.

Son, Y.-S., Lee, G., Lee, B. H., Kim, N., Koh, S.-M., Kim, K.-E., & Cho, S.-J. (2022). Application of ASTER data for differentiating carbonate minerals and evaluating MgO content of magnesite in the Jiao-Liao-Ji Belt, North China Craton. *Remote Sensing, 14*(1), 181.

Su, R., Liu, X., Wei, L., & Zou, Q. (2019). Deep-Resp-Forest: A deep forest model to predict anti-cancer drug response. *Methods, 166*, 91–102.

Sun, L., Mo, Z., Yan, F., Xia, L., Shan, F., Ding, Z., Song, B., Gao, W., Shao, W., Shi, F., Yuan, H., Jiang, H., Wu, D., Wei, Y., Gao, Y., Sui, H., Zhang, D., & Shen, D. (2020a). Adaptive feature selection guided deep forest for COVID-19 classification with chest CT. *IEEE Journal of Biomedical and Health Informatics, 24*(10), 2798–2805.

Sun, T., Chen, F., Zhong, L., Liu, W., & Wang, Y. (2019). GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China. *Ore Geology Reviews, 109*, 26–49.

Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., & Hu, Z. (2020b). 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.

Tanaka, S., Tsuru, H., Someno, K., & Yamaguchi, Y. (2019). Identification of alteration minerals from unstable reflectance spectra using a deep learning method. *Geosciences, 9*(5), 195.

van der Meer, F. D., van der Werff, H. M. A., & van Ruitenbeek, F. J. A. (2014). Potential of ESA's Sentinel-2 for geological applications. *Remote Sensing of Environment, 148*, 124–133.

Vapnik, V. (1999). *The nature of statistical learning theory*. Springer.

Xi, Y., Mohamed Taha, A. M., Hu, A., & Liu, X. (2022). Accuracy comparison of various remote sensing data in lithological classification based on random forest algorithm. *Geocarto International, 37*(26), 14451–14479.

Xu, Y., Li, Z., Xie, Z., Cai, H., Niu, P., & Liu, H. (2021). Mineral prospectivity mapping by deep learning method in Yawan-Daqiao area. *Gansu. Ore Geology Reviews, 138*, 104316.

Yousefi, M., & Nykänen, V. (2016). Data-driven logistic-based weighting of geochemical and geological evidence layers in mineral prospectivity mapping. *Journal of Geochemical Exploration, 164*, 94–106.

Yu, Z., Liu, B., Xie, M., Wu, Y., Kong, Y., Li, C., Chen, G., Gao, Y., Zha, S., Zhang, H., Wang, L., & Tang, R. (2022). 3D mineral prospectivity mapping of Zaozigou Gold Deposit, West Qinling, China: Deep learning-based mineral prediction. *Minerals, 12*(11), 1382.

Zeinelabdein, K. A. E., & Nadi, A. H. H. E. (2014). The use of Landsat 8 OLI image for the delineation of gossanic ridges in the Red Sea Hills of NE Sudan. *American Journal of Earth Sciences, 1*(3), 62–67.

Zeinelabdein, K. E., & Albiely, A. (2008). Ratio image processing techniques: A prospecting tool for mineral deposits, Red Sea Hills, NE Sudan. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37*, 1295–1298.

Zhang, C., Yi, M., Ye, F., Xu, Q., Li, X., & Gan, Q. (2022). Application and evaluation of deep neural networks for airborne hyperspectral remote sensing mineral mapping: A case study of the Baiyanghe Uranium Deposit in Northwestern Xinjiang, China. *Remote Sensing, 14*(20), 5122.

Zhang, L., Sun, H., Rao, Z., & Ji, H. (2020). Hyperspectral imaging technology combined with deep forest model to identify frost-damaged rice seeds. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 229*, 117973.

Zhang, S. E., Nwaila, G. T., Agard, S., Bourdeau, J. E., Carranza, E. J. M., & Ghorbani, Y. (2023). Big geochemical data through remote sensing for dynamic mineral resource monitoring in tailing storage facilities. *Artificial Intelligence in Geosciences, 4*, 137–149.

Zhang, T., Yi, G., Li, H., Wang, Z., Tang, J., Zhong, K., Li, Y., Wang, Q., & Bie, X. (2016). Integrating data of ASTER and Landsat-8 OLI (AO) for hydrothermal alteration mineral mapping in duolong porphyry Cu-Au Deposit, Tibetan Plateau, China. *Remote Sensing, 8*(11), 890.

Zhao, K., Xu, Z., Zhang, T. Z., Tang, Y., & Yan, M. (2021). Simplified deep forest model based just-in-time defect prediction for android mobile apps. *IEEE Transactions on Reliability, 70*(2), 848–859.

Zhou, Z.-H., & Feng, J. (2019). Deep forest. *National Science Review, 6*(1), 74–86.

Zhou, Z.-H., & Liu, X.-Y. (2005). Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on knowledge and data engineering, 18*(1), 63–77.

Zoheir, B., El-Wahed, M. A., Pour, A. B., & Abdelnasser, A. (2019). Orogenic gold in transpression and transtension zones: Field and remote sensing studies of the Barramiya-Mueilha Sector. *Egypt. Remote Sensing, 11*(18), 2122.

Zuo, R. (2020). Geodata science-based mineral prospectivity mapping: A review. *Natural Resources Research, 29*(6), 3415–3424.

Zuo, R., & Carranza, E. J. M. (2011). Support vector machine: A tool for mapping mineral prospectivity. *Computers & Geosciences, 37*(12), 1967–1975.

Zuo, R., Xiong, Y., Wang, J., & Carranza, E. J. M. (2019). Deep learning and its application in geochemical mapping. *Earth-Science Reviews, 192*, 1–14.