Earth Science Informatics
https://doi.org/10.1007/s12145-024-01243-4

RESEARCH

Application of geophysical and multispectral imagery data for predictive mapping of a complex geo-tectonic unit: a case study of the East Vardar Ophiolite Zone, North-Macedonia

Filip Arnaut[1] · Dragana Đurić[2] · Uroš Đurić[3] · Mileva Samardžić-Petrović[3] · Igor Peshevski[4]

Received: 12 July 2023 / Accepted: 28 January 2024

Abstract
The Random Forest (RF) and K nearest neighbors (KNN) machine learning (ML) algorithms were evaluated for their ability to predict ophiolite occurrences, in the East Vardar Zone (EVZ) of central North Macedonia. A predictive map of the investigated area was created using three data sources: geophysical data (digital elevation model, gravity and geomagnetic), multispectral optical satellite images (Landsat 7 ETM + and their derivatives), and geological data (distance to fault map and ophiolite outcrops map). The research included a comparison and discussion on the statistical and geological findings derived from different training dataset class ratios in relation to a testing dataset characterized by significant class imbalance. The results suggest that the precise selection of a suitable class balance for the training dataset is a critical factor in achieving accurate ophiolite prediction with RF and KNN algorithms. The analysis of feature importance revealed that the Bouguer gravity anomaly map, total intensity of the Earth's magnetic field reduced to the pole map, distance to fault map, band ratio BR3 map obtained from multispectral satellite images, and digital elevation model are the most significant features for predicting ophiolites within the EVZ. KNN showed poorer results compared to RF in terms of both the evaluation metrics and visual analysis of prediction maps. The methods applied in this research can be applied for predictive mapping of complex geo-tectonic units covered by dense vegetation, and may indicate the presence of these units even if they were not previously mapped, particularly when geophysical data are used as features.

Keywords Random Forest · K nearest neighbors · Remote Sensing · Geophysical data · Predictive mapping · East Vardar Zone

Introduction

In machine learning algorithms (ML), the automatic inductive approach was used to recognize patterns in data, and the learned pattern relationships were then applied to other similar data or the same datasets but in different domains to generate predictions for data-driven classification and regression problems (Cracknell and Reading 2014). In situations involving the prediction of spatially dispersed categories in extremely complex processes, these algorithms have proven to be immensely useful (Kanevski et al. 2009). The Random Forest (RF) algorithm is widely used for predictive

Communicated by H. Babaie.

✉ Filip Arnaut
filip.arnaut@ipb.ac.rs

Dragana Đurić
dragana.djuric@rgf.bg.ac.rs

Uroš Đurić
udjuric@grf.bg.ac.rs

Mileva Samardžić-Petrović
mimas@grf.bg.ac.rs

Igor Peshevski
pesevski@gf.ukim.edu.mk

[1] University of Belgrade, Institute of Physics Belgrade, Pregrevica 118, 11080 Belgrade, Serbia

[2] University of Belgrade, Faculty of Mining and Geology, Đušina 7, 11000 Belgrade, Serbia

[3] University of Belgrade, Faculty of Civil Engineering, Bulevar Kralja Aleksandra 73/1, 11000 Belgrade, Serbia

[4] Ss. Cyril and Methodius University in Skopje, Faculty of Civil Engineering MK, Bulevar Partizanski Odredi 24, 1000 Skopje, North Macedonia

Published online: 20 February 2024

or geological mapping (Zuo and Carranza 2023). The mineral prospectivity mapping studies conducted in the Philippines demonstrated that the RF model exhibited superior performance compared to other models in predicting the presence of gold deposits (Carranza and Laborte 2015a). This was also observed in cases where there were only a few occurrences of minerals, such as in the predictive mapping of porphyry Cu deposits (Carranza and Laborte 2015b). Furthermore, the RF method offers a notable advantage in the form of feature importance quantification. This quantification, as noted by Rodriguez-Galiano et al. (2014), provides crucial information that aligns with geological expectations. Machine learning methods have demonstrated their efficacy as a valuable tool in geological mapping, because of that, applications in ophiolite mapping could provide satisfactory results.

Remote sensing techniques and satellite imagery, characterized by different spatial and spectral resolutions, have been extensively studied for their application in geological mapping through the analysis and processing of digital images (Cracknell and Reading 2014). Multispectral satellite data, such as Landsat 7 ETM+, Landsat 8, Landsat 9, or Sentinel-2, have been effectively employed for mapping lithology and geological structures in well-exposed geological areas where spectral analysis is optimal (Aliyu et al. 2021; Lorenz 2004; Harris et al. 2005, 2008, 2009, 2014; Schetselaar and Ryan 2008; Leverington 2010; Leverington and Moon 2012; Behnia et al. 2012; Harris and Grunsky 2015; Albert and Ammar 2021). While there were positive outcomes in the application of remote sensing for lithological mapping of surface occurrences, challenges arise in characterizing geological materials in areas covered with dense vegetation or sediments (Leverington and Moon 2012; Harris and Grunsky 2015; Kuhn et al 2018; Ge et al. 2022).

In contrast to prior studies that focused on predicting land cover or vegetation classes using multispectral imagery and machine learning (ML) algorithms (Huang et al. 2002; Foody and Mathur 2004; Ham et al. 2005; Waske and Braun 2009; Cracknell and Reading 2014), this study emphasizes the use of ML methods to map complex geological units, irrespective of whether they are exposed outcrops or not. The approach employed in this study involved utilizing different data sets, considering their level, resolution, and physical properties (Kuhn et al. 2018), including airborne magnetic data, terrestrial gravity data, and multispectral satellite data (Landsat 7 ETM+). These data sets underwent ML algorithm processing to generate continuous surface raster maps of ophiolite complex data, predicting and mapping the complex geological unit—the East Vardar ophiolite zone.

Geophysical data derived from measured values were utilized to assess the spatial distribution of properties, primarily for phenomena not observable at the surface, as commonly done in geophysics (Cracknell and Reading 2014). Physical parameters in geophysics are typically associated with specific lithologies. The data used in this application included multiple remotely sensed data sets, encompassing airborne geophysics (magnetics and elevation), terrestrial geophysics (gravity data), and satellite multispectral data. The primary task involved a high-dimensional input space and a complex, variable set of data relationships. All data were represented as raster data and used in the ML algorithm, regardless of the values they represented.

The evaluated algorithms include RF and K nearest neighbors (KNN). According to Cracknell and Reading (2014) and their cited sources, the architecture of ML algorithms and the statistical distributions of data guide the training of classification models (typically by minimizing a loss function), which are then applied to the same input variables to predict classes present in the training data. The emphasis was on comparing the categorical lithology predictions that were generated by the ML algorithms RF and KNN with various data. In this study, both RF and KNN were used to combine satellite multispectral data with geophysical airborne and terrestrial data, as well as the distance to fault map, as features. Incorporating these data, particularly geophysical data (airborne and terrestrial), into ML predictions has not yet been realized, particularly for mapping complex lithological units as single areas of interest (Cracknell and Reading 2014). This research was focused on the improvement of a set of features (widely accessible remote sensing data supplemented with geophysical data and tectonic settings; distance to fault map) with the RF and KNN algorithms in areas with dense vegetation and complex geological units below the surface. An added element of this study pertains to the utilization of legacy geological maps as the basis for ophiolite test data (Pendžerkovski et al. 1963; Rakikevik et al. 1965, 1969, 1973; Ivanovski and Rakicević 1970; Hristov et al. 1965, 1973; Karajovanovic and Hristov 1976; Dumurdzanov et al. 1981; Karajovanovik and Hadži-Mitrova 1982), which were created using conventional geological mapping methods i.e., surface mapping techniques. The utilization of RF and KNN algorithms presents the potential for predicting ophiolites that may not be noticeable at the terrain surface, specifically those which are situated in the subsurface and that cannot be mapped using conventional geological mapping methods. The study presented a ML task that involved binary classification, which entails the prediction of one of two classes: the non-ophiolite class (designated as 0) and the ophiolite class (designated as 1).

Study area

The study area is in the central part of North Macedonia, spans 5270 km^2 from north to south (Fig. 1), and contains the East Vardar Zone (EVZ). In general, the EVZ traverses

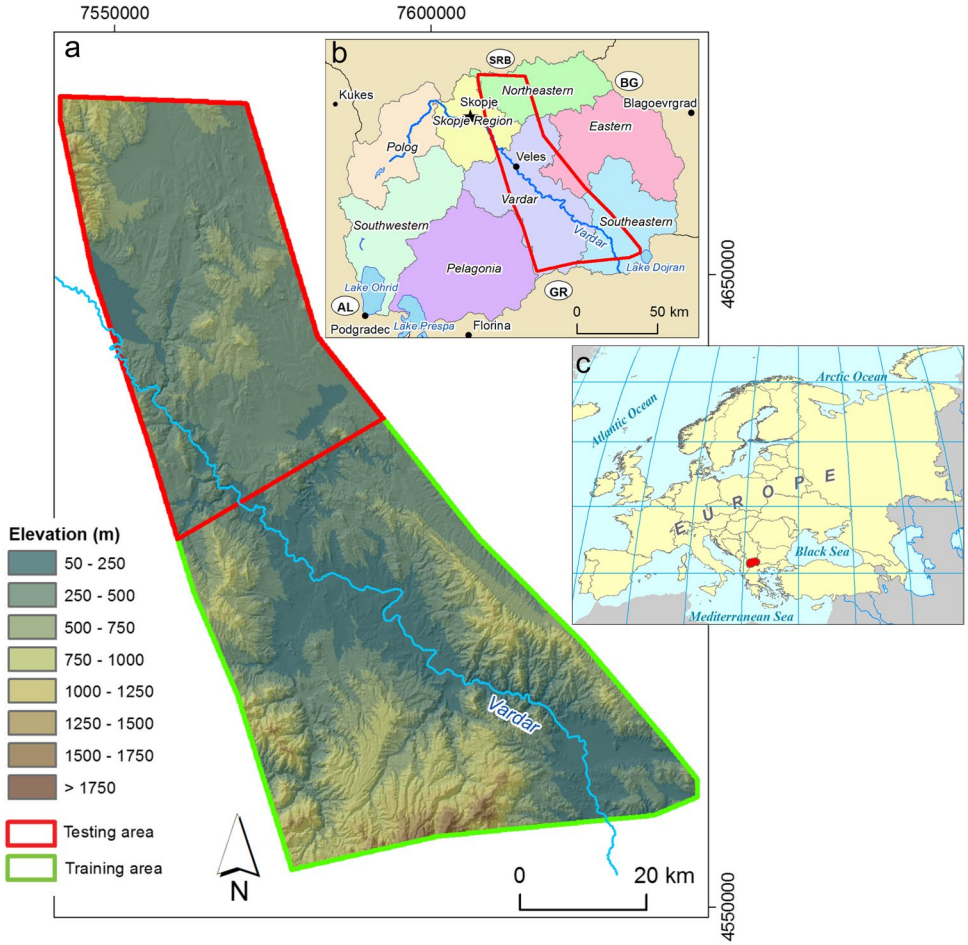

Fig. 1 **a**) Digital elevation model (DEM) of study area with highlighted testing and training area; (**b**) Geographical position of study area (red polygon) within North Macedonia; (**c**) The geographical position of North Macedonia

Romania, Serbia, North Macedonia, Greece, and Turkey (from north to south) (Petrović et al. 2015) and represents the easternmost part of the extremely complex tectonic unit in the central Balkan Peninsula, namely the Vardar Zone (VZ).

As one of several subparallel and NNW- SSE stretching belts that resulted from Mesozoic convergence and subsequent collision, the Vardar Zone is located between the Drina- Ivanjica Unit and the Serbo-Macedonian Massif (SMM). As an NNW- SSE extending belt, EVZ is generally situated between Kopaonik Unit (to the west) and SMM (to the east) (Fig. 2). The presence of the ophiolite complex is a defining feature of EVZ. Ophiolites are portions of the oceanic crust with middle ocean ridge (MOR) or supra-subduction zone (SSZ) affinity that were tectonically deposited on the continental margin during the closure of the ocean or the exhumation process. According to Moore's (1982) classification, EVZ ophiolites belong to the Cordilleran type of ophiolite, which are spatially and temporally associated with island and volcanic arches, pyroclastic rocks, subduction mechanism, and the accretionary mechanism (Petrović et al. 2015). EVZ ophiolites are a complex unit, composed of ingenious rocks (gabbro-dolerites, dolerites, dolerite dykes, and rare basaltic pillow lavas) with geochemical and tectonic affinity to "island arc tholeiite" and MOR. Occasionally, serpentinized harzburgites are found in small numbers at the EVZ margins. All ophiolitic rocks are found in association with intermediate and acid-calcalkaline granitic rocks (Resimić-Šarić et al. 2006). Overlying ophiolites in the study region are Tithonian limestone overstep sequence and/or Cretaceous sedimentary rocks (Dimitrijević 1997).

The western part of the study area corresponds to Kopaonik Unit (KU), also known as Paikon Unit in Greece (Đurić, 2023). The primary lithological components of the KU are shales, phyllites with quartzite, limestone, spilite, and carbonates from the upper Triassic. Part of the eastern portion of the KU is covered by unaltered Lower Cretaceous para-flysch (Schefer et al. 2007; Zelić et al. 2010). The eastern part of the study area consists of SMM metamorphic rocks, which are typically the oldest units in the study area: Precambrian and Paleozoic metamorphites. To obtain a comprehensive synthesis of the geological evolution of the EVZ, it is recommended to refer to the work of Boev et al. (2018).

To simplify the geological analysis and produce a continuous raster, the data used in this study were obtained

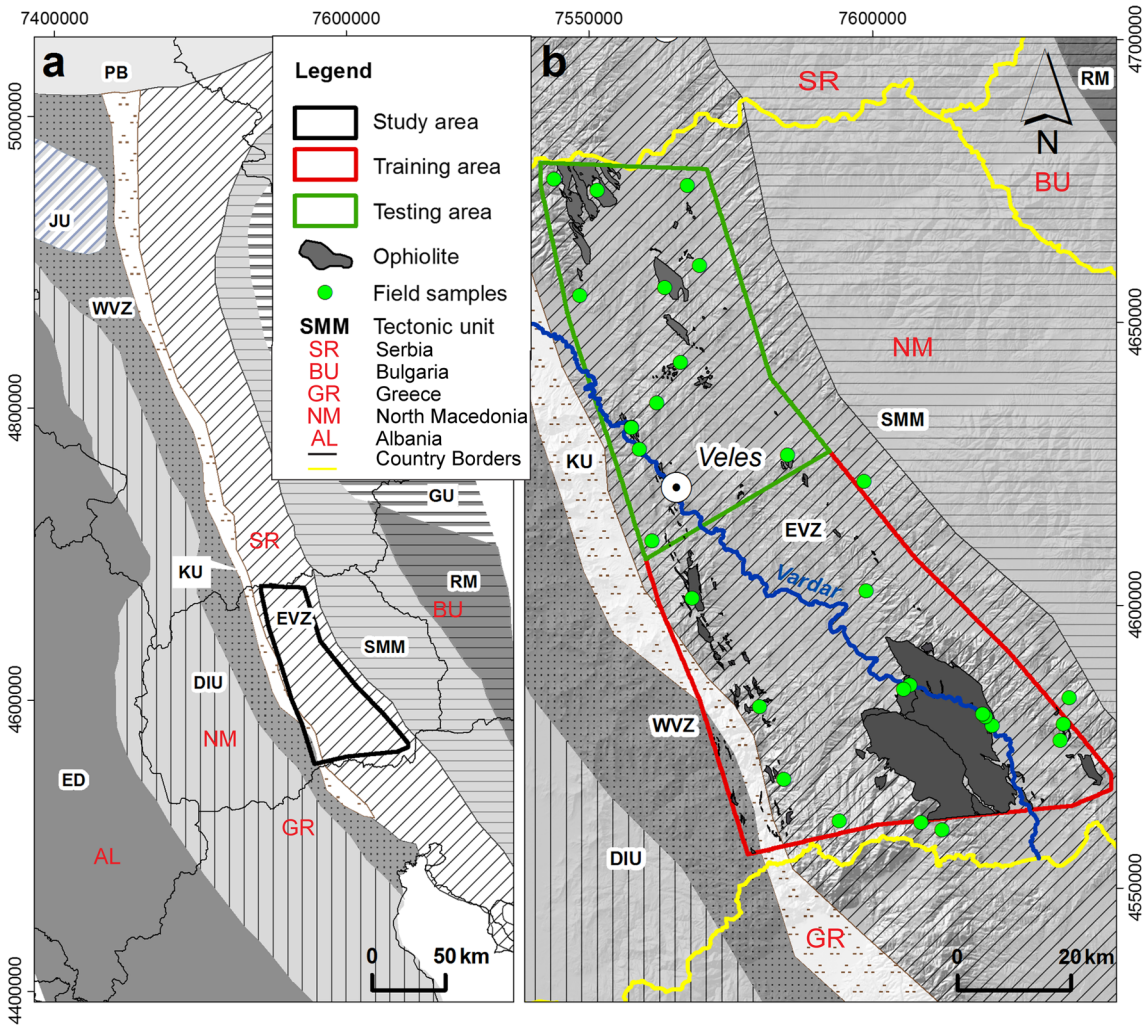

Fig. 2 **a**) Simplified geotectonic map of part of the Balkan Peninsula (modified from Robertson et al. 2009); **b**) Simplified getectonic-geological map of the study area with testing and training areas outlined; Tectonic units: **ED**-External Dinaric Unit, **WVZ**- west Vardar Zone, **KU**- Kopaonik Unit (Kopaonik Block and Ridge Unit), **EVZ**- East Vardar Zone, **JU**- Jadar Unit, **SMM**- Serbo-Macedonian Massif, **RM**- Rhodope Massif, **GU**- Getic Units, **PB**- Pannonian Block

from basic geological maps. The study grouped dolerite, gabbro, gabbro-dolerite, basaltic pillow lavas, and serpentinized harzburgites as a single ophiolite unit (as shown in Fig. 2), while all other geological units were grouped as non-ophiolite units.

Materials and methods

The materials and methods chapter, and the results and discussion chapter, follow Fig. 3, which displays the research workflow. Data sources, including multispectral satellite imaging, geophysical data, and geological data, are described in the materials and methods chapter. Data preprocessing is an essential step in preparing data for ML tasks, and it is considered a distinct result as it involves transforming the original data into data sets for specific purposes in ML modelling (in this case displaying the different results with different class balances) and it is presented in the results section.

Data sources

A predictive map was generated for ophiolites in the study area by utilizing three distinct data sources. These sources include multispectral optical satellite images (Landsat 7 ETM + and their derivatives), geophysical data (including gravity, geomagnetic, and digital elevation model), and geological data (such as the distance to fault map and field mapped ophiolite localities obtained from legacy geological maps).

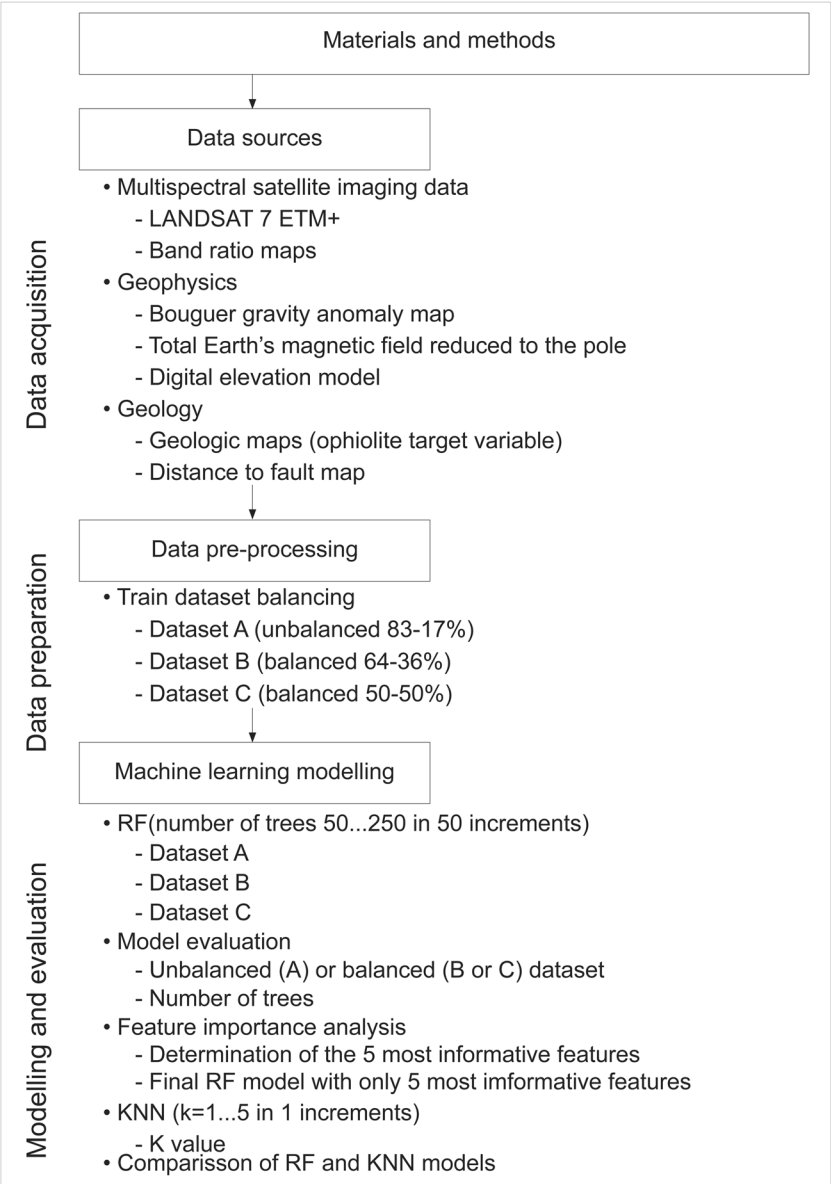

Fig. 3 Data acquisition, preparation, machine learning modelling and evaluation workflow

Multispectral satellite imaging data

The Landsat 7 ETM + satellite images of the investigated area (30 m resolution), bands 1, 2, 3, 4, 5 and 7 were used to predict distinct classes in according due to their established track record of delivering reliable and robust results in similar applications, as evidenced by previous studies (Al-Rawashdeh et al. 2006). The Landsat 7 satellite image underwent essential preprocessing steps, including geometric correction, atmospheric correction, and noise elimination. Additionally, basic processing procedures such as filtering, illumination enhancement, and contrast improvement were applied to enhance the overall quality and interpretability of the image. Band ratios, which are widely used for lithological mapping (Bolt and Bruggenwert 1976; Sposito 1989; Farrand 1997; Longhi et al. 2001; Akhavi et al. 2001; Neville et al. 2003; Al-Rawashdeh et al. 2006), was also derived from Landsat 7 images using the previously mentioned ratios 3/1 (BR1), 5/4 (BR2), and 5/7 (BR3), chosen based on the spectral characteristics of the ophiolite complex. Table 1 provides the corresponding wavelengths for each band of the Landsat 7 ETM + satellite.

Geophysical data

In addition to the DEM (Fig. 1c) that was obtained from the ASTER mission (30 m resolution), the Bouguer i.e., gravity anomaly map (BAM) and the total intensity of the Earth's magnetic field anomaly map, reduced to the pole

(RTP), were selected as geophysical input parameters i.e., features for predictive lithological mapping.

Gravity data (Bilibajkić et al. 1979) were obtained by detailed gravimetric surveys conducted between 1952 and 1984 by the Geophysical Institute in Serbia and Macedonia. The Cassinis Formula (1930) was utilized to calculate normal gravity values. Bullard A, a simple Bouguer correction, and free-air corrections of the first order were applied. The Bouguer anomaly map (Fig. 4a) was generated utilizing a density of 2.67 g/cm^3, whereas the measurement density was approximately 1 point/km^2.

From geomagnetic data (surveys in Serbia and North Macedonia), the total intensity of the Earth's magnetic field anomaly map (Fig. 4b) was calculated. The normal magnetic field was computed using the Geomagnetic Institute's (Grocka) formula for the epoch 1960.0. RTP was applied to these data, resulting in the Total Intensity EMF Anomaly Map being reduced to the pole, considering the Koenigsberger ratio (Q ratio) within the study area (29 localities shown on Fig. 2), which does not exceed the value of 0.001, indicating that the effects of the remanent magnetization can be disregarded (Petrović 2015; Cvetkov et al. 2016).

Geologic data

The published 1:100,000 SFR Yugoslavia Basic Geologic Maps (Pendžerkovski et al. 1963; Rakikevik et al. 1965, 1969, 1973; Ivanovski and Rakicević 1966; Hristov et al. 1965, 1973; Karajovanovic and Hristov 1976; Dumurdzanov et al. 1981; Dimitrijević, 1978; Karajovanovik and Hadži-Mitrova 1982) were used to obtain labelled samples (manually annotated i.e., classified instances of ophiolites) representing ophiolite classes for training and evaluation of RF and KNN predictions.

The strategic placement of ophiolite is intricately linked to the Mesozoic tectonic settings of the region. Moreover, the discernible regional fault pattern plays a pivotal role in this geological context, underscoring its profound significance. Using the same geological map sheets, data obtained by fieldwork, all faults were digitized and subsequently interpreted. Employing a well-established common criterion (Brockmann et al. 1977; Novak and Soulakellis 2000) based on satellite images (Fig. 5a), we determined these faults as regional faults that could be responsible for ophiolite emplacement within the area. The distance-to-fault raster is a map that shows the distance of each pixel in the study area to the nearest fault. In this case, it was created using a fault map and the Euclidian Distance tool (ArcGIS® 10.1). The fault map provides information about the location

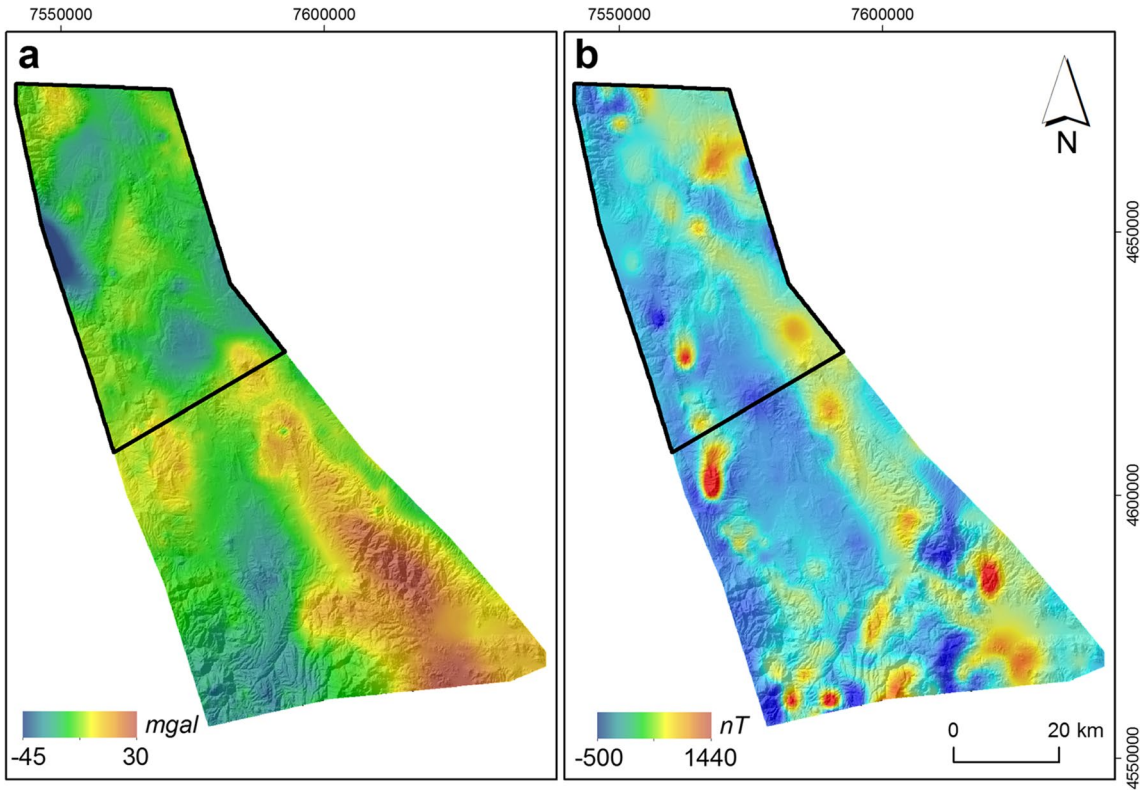

Fig. 4 Geophysical maps; **a**) Bouguer Anomaly Map; **b**) Total Intensity EMF Anomaly Map reduced to the pole; Black outline- testing area

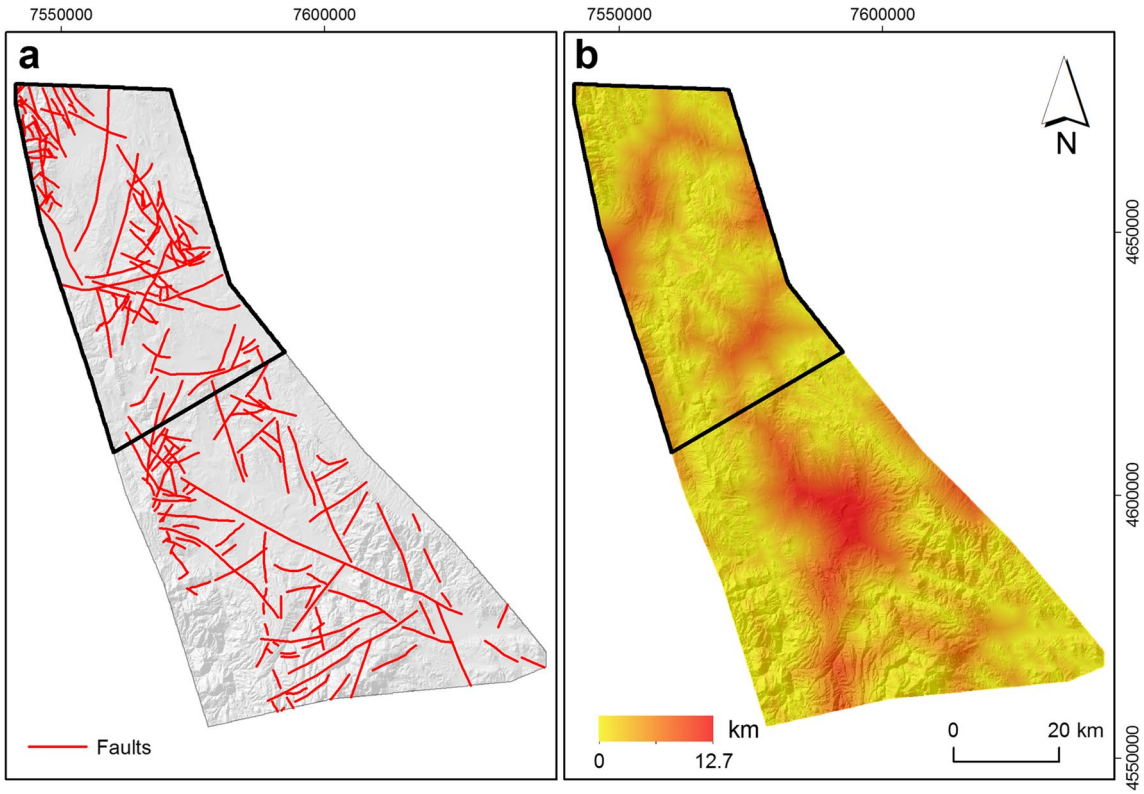

Fig. 5 **a**) Fault map; **b**) Distance to fault map; Black outline- testing area

and extent of known faults in the area, while the Euclidian Distance tool calculates the straight-line distance from each pixel to the nearest fault. The resulting raster map can be used as a feature in ML algorithms to help predict the distribution of geological units or other phenomena in the study area, considering the influence of nearby faults.

All the obtained data was transformed to the MGI Balkans Zone 7 projected coordinate system. The target variable, namely the ophiolite map, was characterized as a binary categorical class with two distinct values (0 for non-ophiolites and 1 for ophiolites). Table 1 provides a brief overview of the data utilized in this study, including the corresponding

Table 1 Utilized data, abbreviations, groups, and types used in this research

Parameter	Symbol	Data group	Data type	Usage	Other
Landsat 7 ETM + channel 1	C1	Satellite	Integer	Feature	0.45–0.52 μm
Landsat 7 ETM + channel 2	C2	Satellite	Integer	Feature	0.52–0.60 μm
Landsat 7 ETM + channel 3	C3	Satellite	Integer	Feature	0.63–0.69 μm
Landsat 7 ETM + channel 4	C4	Satellite	Integer	Feature	0.77–0.90 μm
Landsat 7 ETM + channel 5	C5	Satellite	Integer	Feature	1.55–1.75 μm
Landsat 7 ETM + channel 7	C7	Satellite	Integer	Feature	2.09–2.35 μm
Band ratio 1	BR1	Satellite	Float	Feature	C3/C1
Band ratio 2	BR2	Satellite	Float	Feature	C5/C4
Band ratio 3	BR3	Satellite	Float	Feature	C5/C7
Digital elevation model	DEM	Geophysics	Integer	Feature	/
Bouguer anomaly map	BAM	Geophysics	Float	Feature	/
Total intensity of the EMF reduced to the pole	RTP	Geophysics	Float	Feature	/
Distance to fault map	DF	Geology	Float	Feature	/
Ophiolite map	Ophiolites	Geology	Categorical	Target	/

abbreviations i.e., symbols, data groupings, and data types employed throughout the research.

Machine learning algorithms

The random forest algorithm

Breiman (2001) proposed and developed the Random Forest (RF) ML algorithm for regression and classification, which has widespread applications and popularity. It has been utilized by numerous environmentalists and geoscientists over the past two decades to predict classes and generate spatial data maps. It is a meta-classifier (estimator) that fits several user-defined decision tree classifiers on various subsamples of the training dataset. RF uses averaging to enhance predictive accuracy and prevent overfitting (Kuncheva 2004; Cutler et al. 2007; Youssef et al. 2016).

RF has gained popularity among researchers in various fields both including science and industry for classification and regression tasks. According to Puissant et al. (2014), when compared to other ML techniques, only a few parameters (the number of variables per node and the number of random trees) must be set by the end user, making data preparation easier and saving time during data processing. Since it is a non-parametric method, the data is less dependent on regular statistical (grid) distribution. According to Breiman (2001), it is not susceptible to noise or overfitting (unlike standard random trees) and is relatively fast in comparison to other classification methods, such as boosting techniques. In addition, the RF algorithm is available as a package within most of the popular open-source statistical software packages (R, Weka or JASP). According to Cracknell (2014), RF should be considered a first-choice algorithm for the prediction of classes that represent lithologies (in this case, ophiolite) using commonly available geological and geophysical data, and this was the primary reason for using this algorithm to predict ophiolite.

Tuning and validation are essential components of all ML-based models, including the RF algorithm. According to previous research (Breiman 2001; Catani et al. 2013; Harris and Grunsky 2015), the most effective method for determining the optimal range of the number of trees (n°T) for an algorithm, is to run a certain (user-defined) number of model performance tests with sequentially increasing n°T, until the Out-of-bag (OOB) error stops decreasing and stabilizes at a near value. The number of trees was incrementally increased from 50 to 250 with a step of 50 trees. Subsequently, all results were graphed on the test area and, alongside the evaluation metrics, were interpreted and discussed. The process of conducting ML classification on spatial data involves a crucial step of visual analysis. In this study, the occurrence of false positive predictions may suggest the existence of ophiolites that were not detected using traditional surface mapping methodologies due to their subsurface location. For the false positive values to possess geological plausibility, it is crucial that they are situated in a region that is consistent with existing geological features, such as a previously identified and mapped ophiolite belt. Moreover, the EVZ ophiolites exhibit lithological and tectonic complexity, resulting in spatially varying compositions. All models were plotted and compared to a digitized ophiolite map derived from legacy geologic maps to validate the models visually.

K nearest neighbors

The K nearest neighbors (KNN) algorithm was developed in the middle of the twentieth century (Fix and Hodges 1951; Cover and Hart 1967). Since then, it has become a popular and simple ML algorithm that has been implemented in classification problems across many research domains.

The KNN classification algorithm is a non-parametric algorithm (Ge et al. 2018) that finds the K closest objects in the train data from which it predicts the test data label (Steinbach and Tan 2009). The choice of the K parameter has a significant impact on the KNN classification (James et al. 2013), and a high K value can result in overfitting (Cracknell and Reading 2014). For this study, the JASP software was utilized for both RF and KNN modelling.

Comparative studies employing both RF and KNN on remotely sensed geo-spatial data demonstrate that the RF algorithm outperforms the KNN algorithm (Cracknell and Reading 2014; Ge et al. 2020; Shayeganpour et al. 2021; Bachri et al. 2022).

The classification evaluation metrics used in this study encompassed standard ML classification evaluation metrics, such as accuracy, precision, F1-Score, area under the receiver operating characteristic curve (AUC), rates derived from the confusion matrix (true positive rate—TP, false positive rate—FP, true negative rate—TN, and false negative rate—FN), and statistical parity. Metrics such as accuracy and precision are used to measure the proportion of correctly classified instances out of the total number of instances (accuracy) or the proportion of correct positive predictions out of the total number of positive predictions (precision). The F1-Score is a metric used to evaluate classification performance and it is calculated as the harmonic mean between the true positive rate and precision. The F1-Score is considered a better metric than accuracy in imbalanced ML tasks (Joshi 2002; Hossin and Sulaiman 2015). The AUC value quantifies the model's capacity to differentiate between the classes in a classification ML task, such as its ability to differentiate between the ophiolite and non-ophiolite classes in this research. The statistical parity parameter indicates the distribution of predicted instances. The previously mentioned evaluation metrics (accuracy, precision and F1 score) should for a better model be represented by higher values,

as for the statistical parity metric it should closely resemble the original class distribution in the testing dataset. All the aforementioned evaluation metrics were used individually for each class to meet specific requirements in model evaluation.

The computational cost of the employed algorithms is an additional factor to be considered in classification modeling. Cracknell and Reading (2014) demonstrated that among the five algorithms presented, the support vector machine (SVM) algorithm exhibited the highest computational time, while the RF and artificial neural networks (ANN) algorithms shared the second position. A prevalent strategy for reducing computational time involves the reduction of the number of features utilized, with the least informative ones being disregarded. The present study yields feature selection tables through RF modeling, which are distinguished by two parameters: the mean decrease in node accuracy (MDNA) and the total increase in node purity (TINP). The results of the feature selection process i.e., feature importance analysis can possibly suggest that it is beneficial to construct an alternative RF model that employs a reduced number of features. This approach would yield a decrease in computational time and a comparison of the two models with different number of features utilized.

Data pre- processing

The research area was partitioned into two distinct areas, namely the training and testing areas. The training area spanned 3270 km^2 and encompassed a total of 328 560 sampling data points. On the other hand, the testing area covered 2000 km^2 and comprised 201 254 sampling data points, as indicated in Table 2. Three distinct sets, each comprising different proportions of training samples, were chosen for analysis. These sets were labeled as A (comprising the entire data sample), B (comprising two-thirds of the 0-class sample), and C (comprising an equal number of samples from both the 0 and 1 ophiolite classes). The primary objective was to assess whether a reduction in the quantity of sampling data points, coupled with an increase in the proportion of ophiolites and non-ophiolites in the training dataset, would yield improved predictions in the testing dataset. It is worth noting that the testing dataset is significantly imbalanced, as evidenced by Table 2.

Table 2 Datasets A, B and C with number of points and ratios for training and testing area

Dataset A					
	0	**1**	**Sum**	**0 [%]**	**1 [%]**
Training	273 106	55 454	328 560	83	17
Testing	189 116	12 138	201 254	93	7
Sum	462 222	67 592	529 814	87.2	12.8
Dataset B					
	0	**1**	**Sum**	**0 [%]**	**1 [%]**
Training	100 000	55 454	155 454	64.3	35.7
Testing	189 116	12 138	201 254	93	7
Sum	289 116	67 592	356 708	81	19
Dataset C					
	0	**1**	**Sum**	**0 [%]**	**1 [%]**
Training	55 454	55 454	110 908	50	50
Testing	189 116	12 138	201 254	93	7
Sum	244 570	67 592	312 162	78	22

-non ophiolite class; **1**—ophiolite class;

Sum—total number of points per dataset

Results

Random forest

The RF algorithm was applied to all three datasets (A, B, and C) with varying numbers of data points and distinct ratios of the two binary categorical classes pertaining to ophiolites. The RF modeling process encompassed five distinct models per dataset, each with varying numbers of trees that ranged from 50 to 250, with increments of 50. The RF modelling yielded both evaluation metrics and the predictive maps of ophiolites.

Figure 6 presents a set of ten evaluation metrics that have been chosen to assess the performance of the model across various training data sets and different numbers of trees in the RF model. Regarding accuracy, the models that employed the A dataset exhibit marginally better accuracy for both categories compared to RF models utilizing B or C datasets. However, the difference is negligible, with only a 1.3% variance between the model that demonstrated the highest accuracy (RF_A_100) and the model with the lowest accuracy (RF_C_100). The precision metric pertaining to the non-ophiolite class is analogous to the accuracy metric, as the disparity between the two is negligible. However, for the ophiolite class, the difference is considerably more obvious, with two distinct outliers exhibiting an overall higher precision. Specifically, the RF models that employ the A dataset with 100 and 250 trees demonstrate this trend. The F1 score is a crucial evaluation metric for assessing the overall performance of a model and is on par with the AUC (Area Under the Receiver Operating Characteristic Curve) metric. It displays a categorization of the values that are associated with different datasets for the ophiolite class. The F1 score exhibits the highest values for the balanced dataset (dataset C), reaching up to 0.174. Conversely, dataset A displays the lowest F1 score values, descending as low as 0.02 (RF_A_250). The AUC values demonstrate a distinction in values across various datasets,

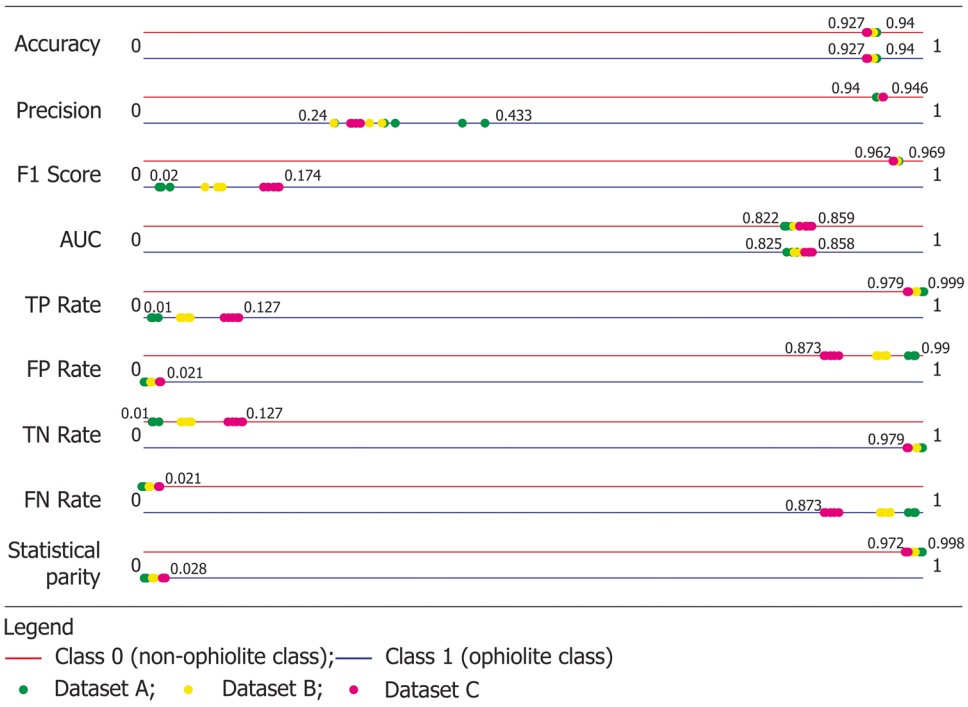

Fig. 6 Selected class evaluation metrics for the RF modelling

with dataset C exhibiting higher values and dataset A exhibiting lower values. The true positive (TP) rate of the ophiolite class indicates that dataset C exhibits the highest rates, while conversely, the dataset C category displays the lowest rates of the true positive rate. The statistical parity, specifically, for dataset C, exhibits peak values of approximately 0.028. This indicates that when the number of data points decreases and the proportion of ophiolites to non-ophiolites approaches equilibrium, the trained model improves in its ability to predict the ophiolite class.

The results of the statistical analysis conducted on the evaluation metrics of the 15 RF models indicate that the most optimal dataset is the C dataset, which is characterized by an even balance of ophiolites to non-ophiolites. In general, the ophiolite class exhibits the highest F1 scores in the C dataset, while the F1 scores for the non-ophiolite class are relatively similar across all three datasets. The dataset denoted as C exhibits the highest AUC values, albeit a slight variance is discernible when compared to other input datasets. Based on the statistical parity metric, it can be observed that the C dataset exhibits the highest likelihood of predicting the ophiolite class. Furthermore, the TP rate for the ophiolite class is also the highest for the C dataset. Overall, within the C dataset, the RF_C_100 model, which employed 100 trees for modeling, was found to be the most optimal. The model exhibits a relatively good AUC of 0.853, with the ophiolite class achieving the highest F1 score of 0.172. Additionally, the model demonstrates a statistical parity value of 0.028, which is also the highest among all models.

In addition to the statistical examination of evaluation metrics, an equally significant analysis pertains to the visual analysis of classification maps, i.e., predictive maps. The predictive maps of the chosen models for all three datasets are illustrated in Fig. 7. The chosen RF models comprise of 50, 100, and 250 trees for every used input dataset.

The RF model, which employed the A dataset and consisted of 50 trees as depicted in Fig. 6, generated 935 predictions for the ophiolite class. Most of these predictions were observed in the NW region of the EVZ. Furthermore, a secondary group of ophiolites is predicted to be present in a limited region situated in the SE part of the EVZ. Conversely, the RF models comprising of 100 and 250 trees exhibited a significantly lower number of predictions, specifically 328 and 302, respectively. The small pocket located in the SE of the EVZ exhibited ophiolite class predictions by both RF_A_100 and RF_A_250. However, it is noteworthy that the A dataset models, overall, exhibit a significant under-prediction of the non-ophiolite class. This can be attributed to the imbalance of the training dataset employed. Furthermore, from a geological perspective, it can be observed that the A dataset models fail to depict the ophiolites' general NE to SW positioning within the EVZ, which is situated at the center of the region.

The initial visual comparison of RF models using dataset A and dataset B reveals that the RF models utilizing dataset B generated a significantly higher number of predictions for the ophiolite class. The RF models that employ a moderately balanced dataset, though not entirely balanced, yield ophiolite class predictions ranging from 2086 to 2391 instances.

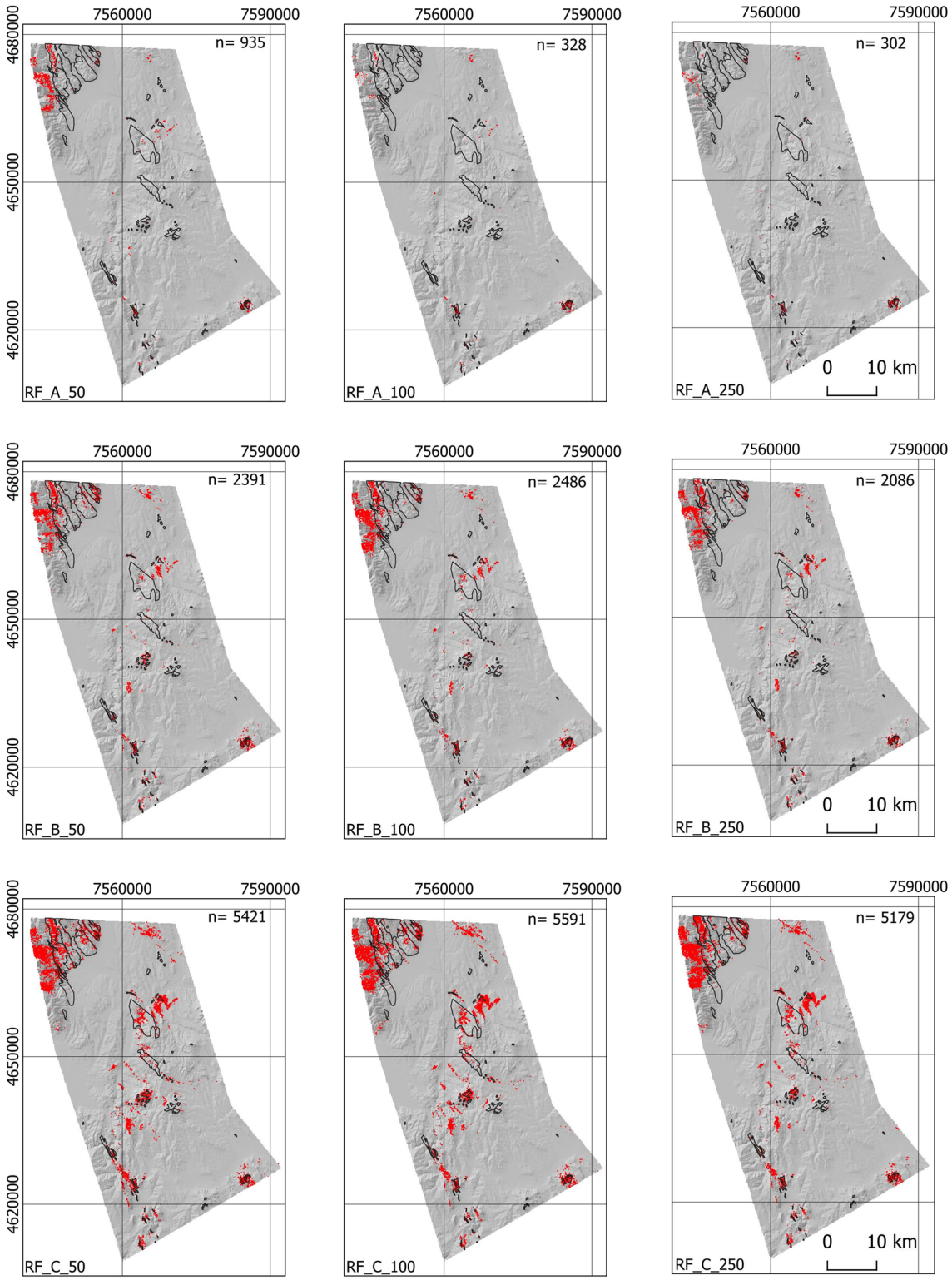

Fig. 7 Selected predictive maps for the RF models; n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

This outcome is consistent with the statistical parity evaluation metric observed earlier. Most of the predictions are situated in the NW region of the EVZ, accompanied by supplementary predictions in the minor enclave in the SE portion of the EVZ. Further predictions pertaining to the ophiolite classification may be discerned at the center of the EVZ, exhibiting a directional orientation of NE-SW. The RF models that employed the B dataset exhibited better performance in comparison to the RF models that employed the A dataset, despite having approximately 50% less data. This observation was further corroborated by the evaluation metrics displayed previously.

Upon visual inspection, it can be inferred that the optimal input dataset among the three RF model groups is dataset C, as supported by the analysis of the evaluation metrics. The dataset denoted as C exhibits many ophiolite classifications in the NW sector of the EVZ. However, a noteworthy advantage of a fully balanced dataset is that the RF C models can accurately predict the geological positioning of ophiolites. The C dataset RF models provide the most accurate prediction of the overall NE-SW strike of ophiolites in the EVZ. Notably, the center of the EVZ exhibits the highest concentration of ophiolite class predictions across all three model groups. Furthermore, it was observed that the RF models trained on the C dataset exhibited the highest number of predictions for the ophiolite class, with a maximum of 5591 predictions achieved by the RF_C_100 model. It is noteworthy that despite having a significantly lower quantity of data, approximately 66% less than the A dataset and 30% less than the B dataset, the C dataset yielded the most accurate predictions.

The study conducted a visual and statistical analysis of evaluation metrics for three groups of RF models with varying numbers of trees. The results indicate that the RF model utilizing the completely balanced C dataset, despite having the least amount of data for model training compared to other input data groups, yields the best overall outcomes for predictive ophiolite mapping.

Feature importance analysis

The RF model, which was trained on the C dataset and consisted of 100 trees, exhibited better performance in terms of evaluation metrics and visual representation through prediction maps. All models employed the complete set of 13 features to generate the prediction map. Additionally, the RF modeling yielded feature importance outputs, which were determined by the MDNA and TINP, as depicted in Fig. 8.

In the RF_C_100 model, the top five features ranked by MDNA and TINP exhibit similarity. These features include the Bouguer anomaly map, distance to fault map, digital elevation model, total intensity of the Earth's magnetic field reduced to the pole, and BR3 map. Likewise, the remaining 14 models exhibited identical outcomes, with the top five most informative parameters consistently ranking in the same order. Additionally, all models shared a commonality in the two parameters that were ranked as the least informative, specifically the Landsat 7 ETM + channel 4 and BR1 map. Regarding the RF_C_100 model, it is significant to mention that the BAM feature is by far the most informative feature for ophiolite prediction, as indicated by both the MDNA and TINP (Fig. 7).

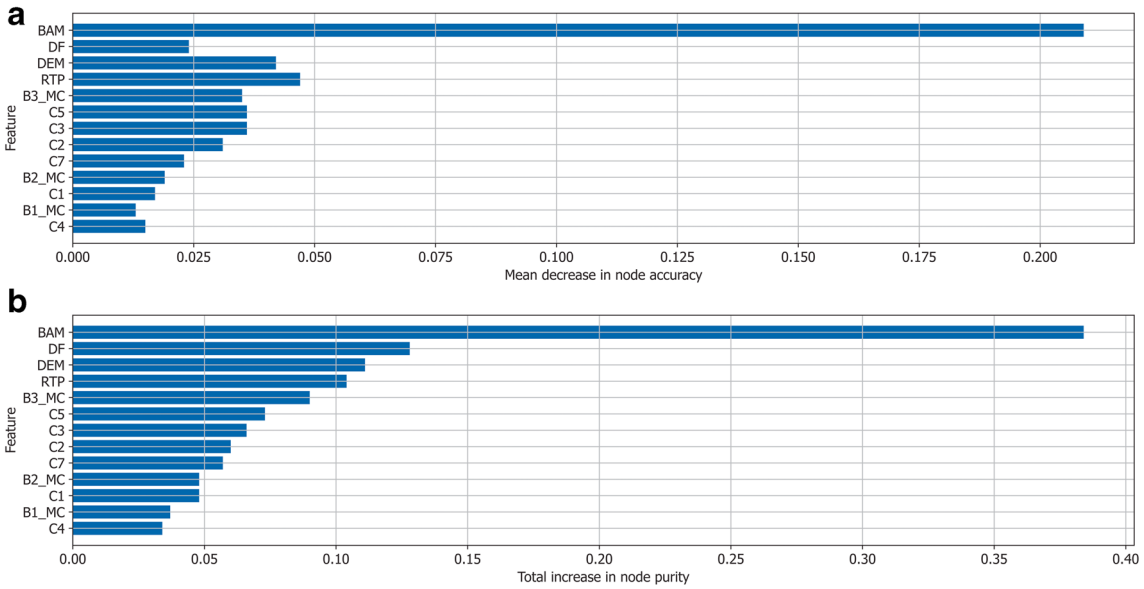

Fig. 8 Feature importance from RF_C_100 model; a) Mean decrease in node accuracy; b) total increase in node purity

Table 3 Comparison between RF_C_100 and RF_C_100_5MI

Model	RF_C_100		RF_C_100_5MI	
/	**0**	**1**	**0**	**1**
Accuracy	0.927	0.927	0.922	0.922
Precision	0.946	0.274	0.944	0.182
F1 Score	0.962	0.174	0.959	0.118
AUC	0.852	0.854	0.829	0.831
TP Rate	0.979	0.127	0.975	0.087
FP Rate	0.873	0.021	0.913	0.025
TN Rate	0.127	0.979	0.087	0.975
FN Rate	0.021	0.873	0.025	0.913
Statistical Parity	0.972	0.028	0.971	0.029
Computational time[s]	210		93	

0- non ophiolite class; **1**—ophiolite class; **TP**- True positive; **FP**- False positive; **TN**- True negative; **FN**- False negative.

To assess the prospective reduction in computational cost linked to RF modelling, the modelling procedure was replicated using the same input dataset (C) and tree count (100). However, unlike RF_C_100, which employed all 13 input features, only the five most informative features, as depicted in Fig. 7, were utilized. The model that was created (RF_C_100_5MI) was compared with the RF_C_100. Table 3 illustrates that the RF_C_100_5MI model shows lower overall performance, as evidenced by the F1 score for the ophiolite class. The discrepancy amounts to approximately 5%. Furthermore, it can be noted that there exists a disparity of approximately 2% in the AUC metric for the ophiolite class, with the RF_C_100 model exhibiting a slightly higher AUC value. The precision metric exhibits a significant contrast, with RF_C_100 demonstrating a 9% higher precision. In general, the evaluation metrics of the RF_C_100_5MI model are most compatible with the evaluation metrics exhibited by the models of the B dataset, except for the statistical parity parameter which is most compatible with the C dataset models. A noteworthy comparison can be drawn between the duration of training and testing for two models. Expectedly, the model that employs all 13 features, which is deemed the overall better model, requires approximately 55% more computational time than the model that employs the same dataset but only the five most informative features.

To fully quantify the performance of the RF_C_100_5MI model, a prediction map was generated using the same methodology as the preceding models (Fig. 9). Table 3 reveals that the RF_C_100_5MI model exhibits a statistical parity evaluation metric that is comparable to that of the RF_C_100 model. Furthermore, while the evaluation metrics of the RF_C_100_5MI model are more comparable with the B input dataset models, the prediction map and the number of ophiolite instances predicted are equivalent to those

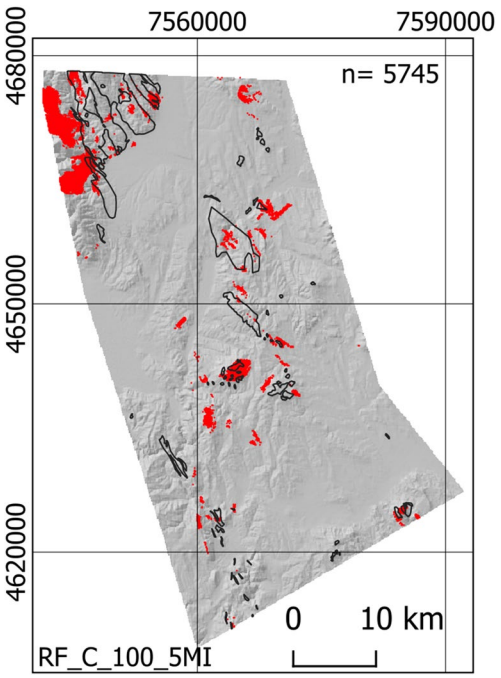

Fig. 9 RF model with only five most informative features (RF_C_100_5MI); n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

of the C dataset models. In general, the RF_C_100_5MI model exhibits comparable prediction maps to those of the C dataset models. Notably, the NW region of the EVZ demonstrates many ophiolite class predictions, along with a minor pocket situated in the SE region of the EVZ. The ophiolites in the EVZ were accurately predicted to have a NE-SW strike. Moreover, the projected count of ophiolite occurrences is comparable to the models in the C dataset, with a predicted value of 5745 ophiolite instances.

In general, the RF_C_100_5MI model exhibits satisfactory evaluation metric statistics and provides an overall adequate geological (visual) prediction of ophiolites in the EVZ. Despite having 8 fewer features, this model offers the added benefit of a 55% reduction in computation time. In this instance, it may be beneficial to make a trade-off by selecting a model that is statistically inferior but yields satisfactory visual outcomes, while requiring 55% less computational time. This holds significant value as it displays that not an overabundance of features is needed for RF modelling, and effective modelling and prediction can be done with five most informative features and a balanced dataset.

K nearest neighbors

The KNN modeling was conducted utilizing solely the top five most informative features that were employed in the RF modeling. The RF_C_100_5MI model demonstrated

reasonable evaluation metric statistics and satisfactory visual predictions of ophiolites, while also offering the advantage of 55% less computational time. Figure 10 presents the evaluation metrics for the five KNN models that were constructed using K values ranging from 1 to 5, with increments of 1.

The F1 score for the ophiolite class indicates that the disparity between the highest-performing model and the lowest-performing model is approximately 1.7%. In contrast, the largest discrepancy in the AUC metric is approximately 8.8%. The difference between the AUC parameter of the optimal KNN model and the RF models is significant, with the KNN models exhibiting markedly inferior AUC values. It is noteworthy to mention that the KNN model exhibits a higher TP rate and statistical parity parameters. The KNN models exhibit a higher likelihood of predicting the ophiolite class in comparison to the RF models, with the former being approximately three times more likely to do so. Specifically, the maximum probability of ophiolite class prediction for the RF models is approximately 2.9%, while that for the KNN models is approximately 6.6%.

The prediction maps for the three KNN models, specifically the models utilizing 1, 2, and 5 nearest neighbors, are depicted in Fig. 11. One notable modification that can be implemented on the RF_C_100_5MI model pertains to the considerable surplus in the number of predicted ophiolite class instances, which surpasses ten thousand, analogous to the statistical parity metric. In addition to variations in the predicted count of ophiolite instances, the KNN models are comparable with the RF models. Most ophiolite class instances predictions are situated in the NW portion of the EVZ. Additionally, the ophiolites' general orientation in the central region of the EVZ is NE-SW, a pattern that all the KNN models accurately predicted. Furthermore, it is possible to observe an area consisting of ophiolite instances in the NE region of the EVZ, which was present in the RF models but not as predominant as in the KNN models.

Based on the KNN modeling, evaluation metric, and analysis, it can be inferred that the KNN model with five nearest neighbors is the most optimal KNN model. Regarding the evaluation metrics, the ophiolite class exhibits the highest F1 score and AUC. However, the AUC value is notably lower than that of the corresponding RF model (RF_C_100_5MI). In general, the KNN models exhibited satisfactory visual and geological foundations for ophiolite prediction, utilizing only five most informative features. However, the predictions are comparatively inferior to those of the RF models.

A comparative analysis can be conducted between RF_C_100_5MI and KNN_C_5, both of which are considered the top-performing models, with RF_C_100 serving as the reference model. A false positive (FP) map can be utilized to draw a comparison, wherein all instances that a particular model has made a FP prediction for are represented (Fig. 12). The significance of the FP map lies in its ability to identify potential areas where ophiolites may be absent from mapping or situated beneath the subsurface and were not able to be mapped with traditional surface geological mapping techniques. The RF_C_100_5MI model exhibited a lower incidence of FP predictions as compared to the KNN model. This

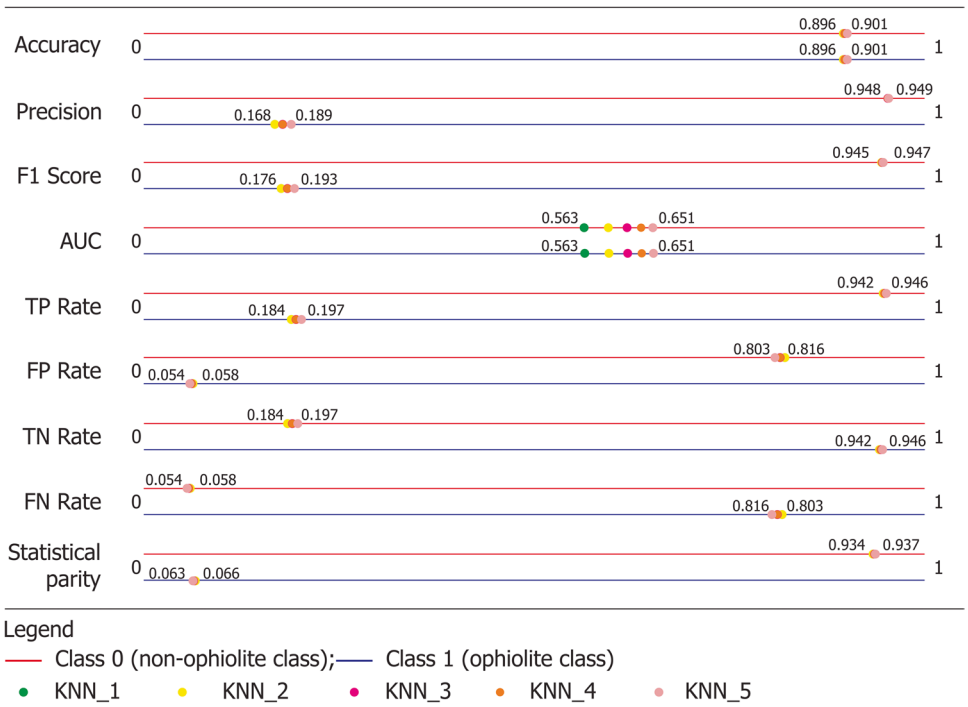

Fig. 10 Selected class evaluation metrics for the KNN modelling

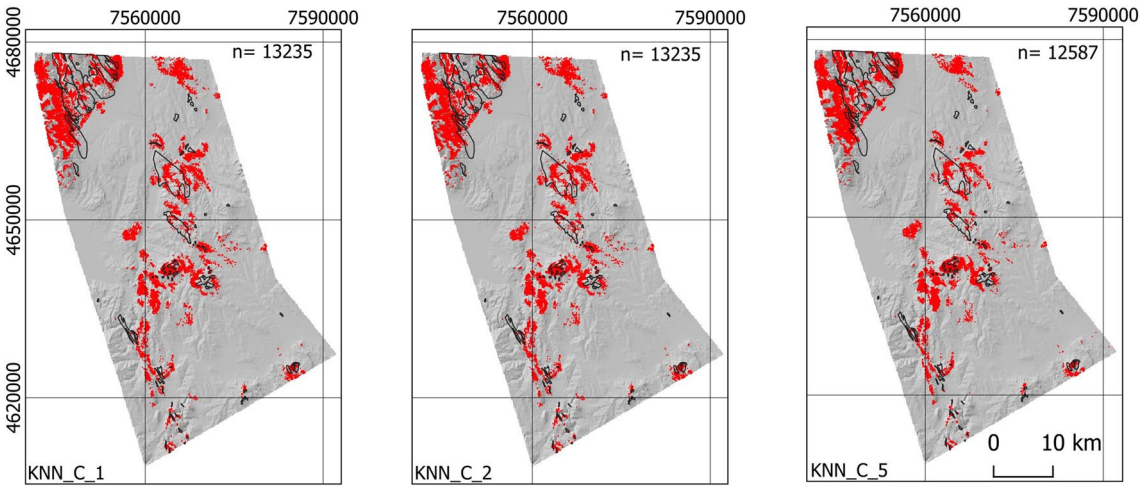

Fig. 11 Selected predictive maps for the KNN models; n- number of ophiolite class instances; black outline- field mapped ophiolites; red markers- predicted ophiolites

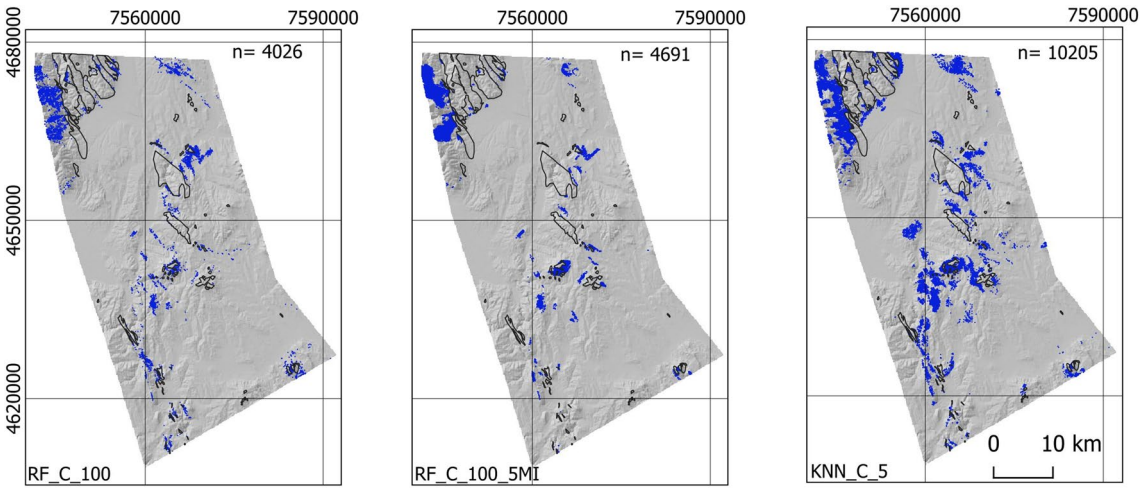

Fig. 12 False positive maps for the RF and KNN models; n- number of false positive instances; black outline- field mapped ophiolites; blue markers- false positive instances

outcome was anticipated, given the KNN model's higher statistical parity metric and the greater number of predicted ophiolite instances overall. The RF_C_100 model exhibited a reduction of 600 FP instances in comparison to the RF model that employed 5 features. The three models exhibited most of the FP predictions in the proximity of pre-existing or mapped ophiolites. Most of them are in the NW part of the EVZ and in the central region of the EVZ, close to the identified ophiolites. When compared, it was observed that the KNN model exhibited a higher number of FP predictions in comparison to both the RF models. It is noteworthy that a specific subset of these FP predictions was situated along the NE-SW strike of the ophiolites through the EVZ.

Discussion

The evaluation of two RF models revealed that the model with a reduced number of features demonstrated lesser performance based on the evaluation metrics. However, it exhibited a notable reduction of approximately 55% in computational time in comparison to the other model. In this instance, the absolute disparity was a modest 137 s. However, when confronted with more extensive datasets, such as a broader region of interest or data of higher resolution, the decrease in computational time may prove to be considerably more noteworthy. The RF_C_100_5MI was selected as the optimal model. Regarding the visual

comparison of the two RF models, it can be observed that the RF_C_100_5MI model produced a prediction map that is satisfactory and closely resembles the one generated by the RF model that employed all 13 features. Furthermore, the KNN model exhibited poorer evaluation metrics statistics, particularly about the AUC parameter, which was not in line with those of the RF model, given the same set of features and input dataset. The KNN model exhibited visually satisfactory prediction maps. Regarding the most informative features, the results revealed intriguing patterns that require further investigation to fully comprehend the geological and geophysical implications. Feature importance is a valuable piece of information that necessitates additional research to establish a comprehensive link between its numerical output and the real nature of the phenomena.

When comparing the RF and KNN models using evaluation metric statistics, it is necessary to take into account various aspects. Although an effort was made to train and test the KNN model using only the five most informative features, as it operates more efficiently on lower dimensional data (i.e., with fewer features), it was ultimately exceeded by the RF model. However, in contrast to the KNN model, the RF model exhibits mode robustness towards outliers, meaning it is less sensitive to them, and it also more effectively captures non-linear relationships within the data. In summary, the RF model exhibits better capability in capturing complex relationships within the data when compared to the KNN model; thus, the KNN model is more appropriate for simpler datasets. This suggests that future research efforts could be focused on more complex models rather than the RF model, as they may be better equipped to handle the utilized data. However, it is important to note that more complex models have a greater number of hyperparameters that require tuning, resulting in a more complicated method of model tuning. While this can be efficiently accomplished through the use of various grid search techniques, it will require greater computational time.

The FP prediction map holds great significance as it enables an investigation into potential ophiolite sites that were not identified through surface mapping methodologies or are situated beneath the surface. In the context of data analysis in areas with dense vegetation, such as the EVZ region, it is possible that the FP prediction map may provide valuable insights for future geological surveys, specifically in identifying potential ophiolite locations.

From a statistical standpoint, it can be observed that all the models generated exhibit a minimal number of predictions related to the total number of ophiolite instances. Despite being well-presented, the most proficient model failed to identify approximately 70% of ophiolite instances within the EVZ. Conversely, incorporating statistical rigor into models that utilize data collected from challenging conditions and incorporating legacy geological data may not be the best course of action. The C dataset models demonstrated a reliable geological prediction map that could potentially serve as a basis for subsequent surveys or more comprehensive ophiolite exploration endeavors. The C dataset models accurately predicted the general occurrence of the ophiolite belt strike in the EVZ, despite their limited ability to predict most individual ophiolite instances. Despite the absence of the most statistically precise models, the visual analysis of the prediction maps demonstrated that the selected models are adequate for the given scale of mapping and the size of the study area.

As demonstrated in a number of studies (Cracknell and Reading 2014; Ge et al. 2020; Shayeganpour et al. 2021; Bachri et al. 2022), the RF method outperforms the KNN model, which was confirmed in this piece of research. There are multiple potential causes for this result; however, the most likely is that the KNN model may not be optimal when dealing with a large feature space, compounded by the likely nonlinear relationship between the features and the target variable. Furthermore, the RF algorithm emerged as a highly favorable initial option, as demonstrated in this study, which emphasized its straightforwardness in implementation (requiring only a few hyperparameters to be adjusted), resistance to overfitting, and overall computational simplicity.

Additionally, it is important to briefly discuss the features employed and their accessibility for researchers seeking to implement the same approach in their own survey region. The feature importance analysis revealed that the five previously mentioned features are sufficient for achieving a sufficiently accurate prediction of ophiolites. The features encompass data that may not be readily accessible to all researchers, such as the Bouguer anomaly map (which can be costly if a gravity survey was not conducted), the total intensity of the Earth's magnetic field (similar to the Bouguer anomaly map), the BR3 map (which necessitates at least two multispectral imagery bands), the distance to fault map (which requires extensive regional geologic mapping), and the digital elevation model (which relies on satellite data). Although ML methods necessitate a substantial amount of high-quality data, additional research can be conducted to provide best practices recommendations regarding the anticipated outcomes when utilizing all available data from various sources (such as this study) and offer insights into the expected outcomes when working with limited data and identify scenarios where attempting ML classification is advisable or not, as ML is a data-driven process.

Conclusion

Based on available geological data, satellite images, and geophysical measurements, this study focused on mapping the extremely complex geological and tectonic units under

conditions of dense vegetation. This approach was evaluated to determine if it provided any useful information. The survey was conducted in the central region of North Macedonia and encompasses a larger portion of the EVZ. The input data (features) for the RF algorithm consisted of bands of Landsat 7 ETM + satellite images, a Total Intensity EMF Anomaly Map reduced to the pole, distance to fault, digital elevation model, and a Bouguer Anomaly Map, unlike other similar studies. The Bouguer Anomaly Map was chosen because it directly indicates the appearance of a body with a higher density, as was the case with gabbro and dolerite as integral components of ophiolite. Due to the 100 m working resolution, local geophysical variations were disregarded.

The RF algorithm was employed to compare input datasets containing varying degrees of ophiolite and non-ophiolite classes. The results of the comparison revealed that dataset C, which was characterized by a balanced 50–50 percent of ophiolites and non-ophiolites in the training dataset, yielded the most optimal outcomes.

According to the results of the feature importance analysis, the Bouguer gravity anomaly map, Total intensity EMF reduced to the pole, distance to fault map, digital elevation model, and BR3 map are the most significant features for predicting ophiolites. The statistical metrics of the model that incorporated all 13 features were better. However, the RF model that employed all features and the model that utilized only the five most informative features exhibited similar ophiolite distribution through the EVZ, as observed visually. Based on statistical analysis, the model that employed only the top five informative features is kin statistically for slightly imbalanced training datasets, such as dataset B. However, upon visual inspection, the models trained on dataset B and the RF model utilizing only the top five informative features and the balanced dataset C exhibit significant dissimilarities. One benefit of utilizing a reduced feature set consisting of only five of the most informative features is a reduction in computational time. Specifically, this approach results in a decrease of approximately 55% in computational time. Furthermore, it is crucial to emphasize the importance of employing a balanced training dataset when dealing with a target class imbalanced testing area. This is because models trained on an imbalanced dataset (dataset A) tend to exhibit significant overestimation of the non-ophiolite class. Implementing a balanced dataset that contains a reduced number of samples (approximately 66% fewer in this instance) yields more precise models. This indicates that in certain ML classification scenarios, a smaller amount of data can be advantageous.

Comparing the RF and KNN algorithms produced results consistent with the existing literature, the RF algorithm produced superior results based on the evaluation metrics of both algorithms. The KNN model, utilizing dataset C as the input dataset, exhibited an FP rate that was nearly two and a half times that of the RF model, which also utilized the same input features and dataset. Furthermore, it was observed that the KNN model exhibited comparatively inferior AUC values, with the maximum value recorded at 0.651 (KNN_C_5). On the other hand, the RF model yielded AUC values of 0.831 (RF_C_100_5MI). Conversely, visual analysis indicated that the KNN algorithm accurately approximated the spatial distribution of the ophiolites within the test area of the EVZ.

In densely vegetated regions with complex geological and tectonic settings, a RF algorithm with appropriate features and a balanced dataset yields satisfactory results for predictive mapping of complex units, such as ophiolites, and may indicate the existence of these units even when they were not field mapped, especially when geophysical data are used as features.

Authors contribution This research project represents a collective effort in which each author played a crucial role in shaping its conception, methodology, and design. The initial idea was ignited by Dragana Đurić and Uroš Đurić, who laid the foundation for the study by developing the very first models and utilizing machine learning algorithms. Subsequently, Mileva Samardžić-Petrović took charge of the statistical analyses and delved into the initial exploration of machine learning algorithms. Building upon this groundwork, Filip Arnaut made significant contributions and novelty by enhancing the machine learning techniques, ultimately introducing the novel application of KNN, which led to the development of entirely new models and enriched the analysis. Moreover, the research was complemented by the invaluable expertise of Igor Peshevski, who conducted an in-depth literature review focusing on geological and tectonic aspects, contributing to a more comprehensive understanding of the context. The collaborative nature of this endeavor is exemplified in the writing process, wherein Filip Arnaut took the lead in crafting the manuscript while benefiting from substantial contributions and assistance from Dragana Đurić, Uroš Đurić, Mileva Samardžić-Petrović, and Igor Peshevski. Throughout this stage, the manuscript underwent internal review process, with each author providing valuable comments and feedback on earlier versions. The culmination of this collaborative effort is the final manuscript, which has been thoroughly reviewed, refined, and approved by all authors.

Funding The study was funded by the Ministry of Education, Science and Technological Development, Republic of Serbia under the "Agreement on the realization and financing of scientific research work in 2023": grant no. 451–03-47/2023–01/200126 (Faculty of Mining and Geology) and grant no. 451–03-68/2020–14/200092 (University of Belgrade, Faculty of Civil Engineering).

Data availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflicts of interest/Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Akhavi MS, Webster TL, Raymond DA (2001) RADARSAT-1 Imagery and GIS Modeling for Mineral Exploration in Nova Scotia, Canada. Geocarto Int 16:57–64. https://doi.org/10.1080/10106040108542183

Albert G, Ammar S (2021) Application of random forest classification and remotely sensed data in geological mapping on the Jebel Meloussi area (Tunisia). Arab J Geosci 14. https://doi.org/10.1007/s12517-021-08509-x

Aliyu A, Adamu LM, Abdulmalik NF et al (2021) Application of remote sensing in lithological discrimination of precambrian basement rocks of Zungeru area, part of sheet 163 (Zungeru NW), North Central Nigeria. FJS 5:390–398. https://doi.org/10.33003/fjs-2021-0503-729

Al-Rawashdeh S, Saleh B, Hamzah M (2006) The use of Remote Sensing Technology in geological Investigation and mineral Detection in El Azraq-Jordan. Cybergeo. https://doi.org/10.4000/cybergeo.2856

Bachri I, Hakdaoui M, Raji M et al (2022) Identification of Lithology Using Sentinel-2A Through an Ensemble of Machine Learning Algorithms. Int J Appl Geospat Res 13:1–17. https://doi.org/10.4018/ijagr.297524

Behnia P, Harris JR, Rainbird RH et al (2012) Remote predictive mapping of bedrock geology using image classification of Landsat and SPOT data, western Minto Inlier, Victoria Island, Northwest Territories, Canada. Int J Remote Sens 33:6876–6903. https://doi.org/10.1080/01431161.2012.693219

Bilibajkić P, Mladenović M, Mujagić S, Rimac I (1979) Explanation for the gravity map of SFR Yugoslavia - Bouguer anomalies - 1:500 000. Federal Geological Institute, Belgrade

Boev B, Cvetković V, Prelević D, Šarić K, Boev I (2018) East Vardar Ophiolites Revisited A Brief Synthesis Of Geology And Geochemical Data. Contrib Sec Nat Math Biotech Sci, Macedonian Academy of Sciences and Arts 39:51–68. https://doi.org/10.20903/csnmbs.masa.2018.39.1.119

Bolt GH, Bruggenwert MGM (1976) Composition of the Soil. Soil Chemistry: A Basic Elements 1–12. https://doi.org/10.1016/s0166-2481(08)70630-5

Breiman L (2001) Random Forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Brockmann CE, Fernandez A, Ballon R, Claure I (Servicio Geologico de Bolivia, La Paz, Bolivia) (1977) Analysis of geological structures based on landsat-1 images. In: Remote Sensing Applications for Mineral Exploration. Dowen, Hutchinson and Ross, Strondsberg, PA, USA, pp 292–317

Carranza EJM, Laborte AG (2015a) Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). Comput Geosci 74:60–70. https://doi.org/10.1016/j.cageo.2014.10.004

Carranza EJM, Laborte AG (2015b) Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. Ore Geol Rev 71:777–787. https://doi.org/10.1016/j.oregeorev.2014.08.010

Cassinis G (1930) Sur L'adoption D'une Formule Internationale Pour la Pesanteur Normale. Bull Géod 26:40–49. https://doi.org/10.1007/bf03030025

Catani F, Lagomarsino D, Segoni S, Tofani V (2013) Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Nat Hazards Earth Syst Sci 13:2815–2831. https://doi.org/10.5194/nhess-13-2815-2013

Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theory 13:21–27. https://doi.org/10.1109/tit.1967.1053964

Cracknell MJ, Reading AM (2014) Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Comput Geosci 63:22–33. https://doi.org/10.1016/j.cageo.2013.10.008

Cracknell JM (2014) Machine Learning for Geological Mapping: Algorithms And Applications. Dissertation, School of Physical Sciences (Earth Sciences), University of Tasmania

Cutler DR, Edwards TC, Beard KH et al (2007) Random forests for classification in ecology. Ecol 88:2783–2792. https://doi.org/10.1890/07-0539.1

Cvetkov V, Đurić D, Lesić V, Starčević M, Petković M, Petrović S (2016) Koenigsberger ratio and Total Magnetic Field Anomaly reduction to the pole for the area of Macedonia. Geologica Macedonica 4:429–534

Dimitrijević MD (1978) Geological mapping. Publishing and informational student center, Belgrade (In Serbian)

Dimitrijević MD (1997) Geology of Yugoslavia. Geological Insti-tute – GEMINI, Belgrade, Spec. Publ., Monograph, 1—197

Dumurdzanov N, Hristov S, Pavlovski B, Ivanova V (1981) Explanatory notes for sheets Vitolište and Kajmakčalan. General geological map (1:100,000) of the Socialist Federal Republic of Yugoslavia, Federal Geological Survey, Belgrade, p 61 (in Macedonian)

Đurić D (2023) Overview on the Results of Potential Field Methods for the Spatial Position of East Vardar Ophiolites. In: Chitea F (ed) Insights of Geosciences for Natural Hazards and Cultural Heritage. Zenodo. https://doi.org/10.5281/zenodo.8103476

Farrand WH (1997) Identification and mapping of ferric oxide and oxyhydroxide minerals in imaging spectrometer data of Summitville, Colorado, U.S.A., and the surrounding San Juan Mountains. Int J Remote Sens 18:1543–1552. https://doi.org/10.1080/01431697218269

Fix E, Hodges JL (1951) Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties. International Statistical Review / Revue Internationale de Statistique 57:238. https://doi.org/10.2307/1403797

Foody GM, Mathur A (2004) A relative evaluation of multiclass image classification by support vector machines. IEEE Trans Geosci 42:1335–1343. https://doi.org/10.1109/tgrs.2004.827257

Ge W, Cheng Q, Tang Y et al (2018) Lithological Classification Using Sentinel-2A Data in the Shibanjing Ophiolite Complex in Inner Mongolia. China Remte Sens 10:638. https://doi.org/10.3390/rs10040638

Ge G, Shi Z, Zhu Y et al (2020) Land use/cover classification in an arid desert-oasis mosaic landscape of China using remote sensed imagery: Performance assessment of four machine learning algorithms. Glob Ecol Conserv 22. https://doi.org/10.1016/j.gecco.2020.e00971

Ge YZ, Zhang ZJ, Cheng QM, Wu GP (2022) Geological mapping of basalt using stream sediment geochemical data: Case study of covered areas in Jining, Inner Mongolia. China J Geochem Explor 232:106888. https://doi.org/10.1016/j.gexplo.2021.106888

Ham J, Chen Y, Crawford MM, Ghosh J (2005) Investigation of the random forest framework for classification of hyperspectral data. IEEE Trans Geosci 43:492–501. https://doi.org/10.1109/tgrs.2004.842481

Harris JR, Grunsky EC (2015) Predictive lithological mapping of Canada's North using Random Forest classification applied to geophysical and geochemical data. Comput Geosci 80:9–25. https://doi.org/10.1016/j.cageo.2015.03.013

Harris JR, Rogge D, Hitchcock R et al (2005) Mapping lithology in Canada's Arctic: application of hyperspectral data using the minimum noise fraction transformation and matched filtering. Can J Earth Sci 42:2173–2193. https://doi.org/10.1139/e05-064

Harris JR, Ford KL, Charbonneau BW (2009) Application of gamma-ray spectrometer data for lithological mapping in a cordilleran

environment, Sekwi Region, NWT. Can J Remote 35:12–30. https://doi.org/10.5589/m09-022
Harris JR, He JX, Rainbird R, Behnia P (2014) A Comparison of Different Remotely Sensed Data for Classifying Bedrock Types in Canada's Arctic: Application of the Robust Classification Method and Random Forests. Geosci Can 41:557. https://doi.org/10.12789/geocanj.2014.41.062
Harris JR, Schetselaar EM, de Kemp E, St-Onge MR (2008) Case study 2. LANDSAT, magnetic and topographic data for regional lithological mapping, southeast Baffin Island. https://doi.org/10.4095/226015
Hossin M, Sulaiman MN (2015). A Review on Evaluation Metrics for Data Classification Evaluations IJDKP. https://doi.org/10.5121/ijdkp.2015.5201
Hristov S, Karajovanovik M, Stračkov M (1965) Basic geological map of SFRJ, sheet Kavadarci, M 1:100,000 (map & interpreter), Federal Geological Survey, Beograd, p 62 (in Macedonian)
Hristov S, Karajovanovik M, Stračkov M (1973) Basic geologic map of Former Yugoslavia 1:100 000, explanatory booklet for sheet Kavadarci (in Macedonian)
Huang C, Davis LS, Townshend JRG (2002) An assessment of support vector machines for land cover classification. Int J Remote Sens 23:725–749. https://doi.org/10.1080/01431160110040323
Ivanovski T, Rakicević T (1966) Basic geologic map of Former Yugoslavia 1: 100 000, sheet of Gevgelija. Federal Geological Survey, Belgrade (In Macedonian)
James G, Witten D, Hastie T, Tibshirani R (2013) An Introduction to Statistical Learning: With Applications in R. Springer, New York
Joshi MV (2002) On evaluating performance of classifiers for rare classes. IEEE Int Conf on Data Mining, Maebashi City, Japan 641:644
Kanevski M, Pozdnukhov A, Timonin V (2009) Machine Learning for Spatial Environmental Data: Theory. CRC Press, Boca Raton, USA, Applications and Software
Karajovanovik M, Hadži-Mitrova S, (1975): Basic geologic map of Former Yugoslavia 1:100 000, explanatory booklet for the sheet of Titov Veles. Federal Geological Survey, Belgrade (In Macedonian)
Karajovanovic M, Hristov S (1976) Explanatory note of the General Geological Map of Kumanovo 1:100,000 map sheet: Skopje, Federal Geological Survey Jugoslavia, p 58 (in Macedonian, with English summary)
Kuhn S, Cracknell MJ, Reading AM (2018) Lithologic mapping using Random Forests applied to geophysical and remote-sensing data: A demonstration study from the Eastern Goldfields of Australia. Geophys 83:B183–B193. https://doi.org/10.1190/geo2017-0590.1
Kuncheva LI (2004) Combining Pattern Classifiers: Methods and Algorithms. John Wiley & Sons Inc., Hoboken, New Jersey
Leverington DW (2010) Discrimination of sedimentary lithologies using Hyperion and Landsat Thematic Mapper data: a case study at Melville Island, Canadian High Arctic. Int J Remote Sens 31:233–260. https://doi.org/10.1080/01431160902882637
Leverington DW, Moon WM (2012) Landsat-TM-Based Discrimination of Lithological Units Associated with the Purtuniq Ophiolite, Quebec, Canada. Remote Sens 4:1208–1231. https://doi.org/10.3390/rs4051208
Longhi I, Sgavetti M, Chiari R, Mazzoli C (2001) Spectral analysis and classification of metamorphic rocks from laboratory reflectance spectra in the 0.4-2.5 μ m interval: A tool for hyperspectral data interpretation. Int J Remote Sens 22:3763–3782. https://doi.org/10.1080/01431160010006980
Lorenz H (2004) Integration of Corona and Landsat Thematic Mapper data for bedrock geological studies in the high Arctic. Int J Remote Sens 25:5143–5162. https://doi.org/10.1080/01431160410001705097
Moores EM (1982) Origin and emplacement of ophiolites. Rev Geophys 20:735–760. https://doi.org/10.1029/RG020i004p00735
Neville RA, Lévesque J, Staenz K et al (2003) Spectral unmixing of hyperspectral imagery for mineral exploration: comparison of results from SFSI and AVIRIS. Can J Remote Sens 29:99–110. https://doi.org/10.5589/m02-085
Novak ID, Soulakellis N (2000) Identifying geomorphic features using LANDSAT-5/TM data processing techniques on Lesvos, Greece. Geomorphology 34:101–109. https://doi.org/10.1016/s0169-555x(00)00003-9
Pendžerkovski J, Rakićević T, Ivanovski T, Gjuzelkovski D (1963) Geological map and guide book for sheet Kožuf (K 34–105), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 47
Petrović D (2015) Spatial position of the East Vardar ophiolite: geophysical-geological model and implications on its geodynamic evolution. University of Belgrade (In Serbian with English abstract), Faculty of Mining and Geology
Petrović D, Cvetkov V, Vasiljević I, Cvetković V (2015) A new geophysical model of the Serbian part of the East Vardar ophiolite: Implications for its geodynamic evolution. J Geodyn 90:1–13. https://doi.org/10.1016/j.jog.2015.07.003
Puissant A, Rougier S, Stumpf A (2014) Object-oriented mapping of urban trees using Random Forest classifiers. Int J Appl Earth Obs Geoinf 26:235–245. https://doi.org/10.1016/j.jag.2013.07.002
Rakikevik T, Stojanov R, Arsovski M (l965) Geological map and guide book for Sheet Prilep (K 34–92), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 65
Rakikevik T, Dumurdžanov N, Petkovski P (l969) Geological map and guide book for Sheet Štip (K 34–81), Basic Geological Map of SFRJ 1:100,000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 70
Rakikevik T, Penđerovski J, Kovačevik M (l973) Geological map and guide book for sheet Strumica (K 34–94), Basic Geological Map of SFRJ 1:100000. Federal Geological Survey (in Macedonian with English abstract). Belgrade, p 69
Resimić-Šarić K, Cvetković V, Balogh K, Koroneos A (2006) Main characteristics of ophiolitic complexes within the eastern branch of the Vardar Zone Composite Terrane in Serbia. In International Symposium on the Mesozoic ophiolite belts ofthe northern part of the Balkan Peninsula, Belgrade (Serbia) and Banja Luka (Bosnia and Herzegovina), May 31st–June 6th, 2006, Abstracts. Faculty of Mining and Geology, Universityof Belgrade, Belgrade, pp 112–115
Robertson A, Karamata S, Šarić K (2009) Overview of ophiolites and related units in the Late Palaeozoic-Early Cenozoic magmatic and tectonic development of Tethys in the northern part of the Balkan region. Lithos 108:1–36. https://doi.org/10.1016/j.lithos.2008.09.007
Rodriguez-Galiano VF, 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. Int J Geogr Inf Sci 28:1336–1354. https://doi.org/10.1080/13658816.2014.885527
Schefer S, Fügenschuh B, Schmid SM, Egli D, Ustaszewski K (2007) Tectonic evolution of the suture zone between Dinarides and Carpatho-Balkan: Field evidence from the Kopaonik Region, Southern Serbia. In: Geophysical Research Abstracts, Vol. 9, 03891. European Geosciences Union
Schetselaar EM, Ryan JJ (2008) A Remote Predictive Mapping Case Study of the Boothia Mainland Area, Nunavut, Canada. In: Harris JR (ed) Remote Predictive Mapping: An Aid for Northern Mapping. Geological Survey of Canada, Open File 5643, pp 261–281. https://doi.org/10.4095/226028

Shayeganpour S, Tangestani MH, Gorsevski PV (2021) Machine learning and multi-sensor data fusion for mapping lithology: A case study of Kowli-kosh area, SW Iran. Adv Sapce Res 68:3992–4015. https://doi.org/10.1016/j.asr.2021.08.003

Sposito G (1989) The Chemistry of Soils. Oxford University Press

Steinbach M, Tan PN (2009) kNN: k- Nearest Neighbors. In: The top ten algorithms in data mining, 1st edn. Chapman and Hall/CRC

Waske B, Braun M (2009) Classifier ensembles for land cover mapping using multitemporal SAR imagery. ISPRS J Photogramm Remote Sens 64:450–457. https://doi.org/10.1016/j.isprsjprs.2009.01.003

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 13:839–856. https://doi.org/10.1007/s10346-015-0614-1

Zelić M, Agostini S, Marroni M, Pandolfi L, Tonarini S (2010) Geological and geochemical features of the Kopaonik intrusive complex (Vardar zone, Serbia). Ofioliti 35:33–47

Zuo R, Carranza EJM (2023) Machine Learning-Based Mapping for Mineral Exploration. Math Geosci 55:891–895. https://doi.org/10.1007/s11004-023-10097-3