

Machine learning-driven mineral prospectivity mapping: A predictive performance analysis in Janja, Iran

Mapowanie perspektywiczne minerałów oparte na uczeniu maszynowym: predykcyjna analiza efektywności obliczeniowej w regionie Janja, Iran

Mohammad Ebdali¹, Ardeshir Hezarkhani¹, Adel Shirazy¹, Amin Beiranvand Pour²

¹ *Amirkabir University of Technology, Iran*

² *University Malaysia Terengganu (UMT), Malaysia*

ABSTRACT: Geochemical analysis is an effective technique for detecting mineral deposits by examining element concentrations. Various statistical techniques have been developed to differentiate abnormal values from background values. A more accurate analysis can be obtained by employing multivariate statistical methods. The use of these methods enables the simultaneous analysis of changes in multiple variables. This research utilized correlation coefficients, cluster analysis, and factor analysis to demonstrate the genetic connections among various elements. The factor analysis method was additionally applied to generate multivariable maps and comprehensive multivariable results. Moreover, the stepwise factor analysis (SFA) method, an enhanced version of traditional factor analysis, was utilized to produce geochemical distribution maps. This technique entails initially recognizing and removing non-representative elements, followed by identifying the most important and impactful representative factors. This study demonstrates the efficacy of the SFA method when applied to geochemical data. This approach removes superfluous elements and increases the variance attributed to the predictive mineralization factor, thereby improving the geochemical halos. Additionally, this research evaluated multivariate analysis approaches alongside machine learning techniques. To achieve this, a multilayer perceptron neural network (MLP) was used to evaluate the levels of gold, silver, copper, lead, and zinc in the study area. The output variable represented the grade of a particular element individually, whereas the input variables encompassed the grades of the remaining four elements. To optimize the model, different quantities of hidden layers and a range of activation functions were applied. Ultimately, an ideal model was developed for each element. The model achieved accuracies of 95%, 88%, 73%, 80%, and 72% for the gold, silver, copper, lead, and zinc, respectively. The results show the significant computational efficiency of this method in assessing element grades. Finally, the element distribution maps generated by both methods indicate that the MLP approach identified the anomalous areas with higher accuracy.

Key words: cluster analysis, correlation coefficients, multivariate statistical methods, stepwise factor analysis, multilayer perceptron.

STRESZCZENIE: Analiza geochemiczna jest skuteczną techniką wykrywania złóż mineralnych poprzez badanie stężeń pierwiastków. Opracowano różne techniki statystyczne w celu odróżnienia nieprawidłowych wartości od wartości tła. Dokładniejszą analizę można uzyskać stosując wielowymiarowe metody statystyczne. Zastosowanie tych metod umożliwia jednoczesną analizę zmian wielu zmiennych. W niniejszym badaniu wykorzystano współczynniki korelacji, analizę skupień i analizę czynnikową w celu wykazania powiązań genetycznych między różnymi pierwiastkami. Metoda analizy czynnikowej została dodatkowo zastosowana do wygenerowania map wielu zmiennych i kompleksowych wyników wielu zmiennych. Co więcej, zastosowano metodę stopniowej analizy czynnikowej (SFA), ulepszoną wersję tradycyjnej analizy czynnikowej, w celu stworzenia map rozkładu geochemicznego. Technika ta polega na wstępnym rozpoznaniu i usunięciu pierwiastków niereprezentatywnych, a następnie zidentyfikowaniu najważniejszych i najbardziej wpływowych czynników reprezentatywnych. Przeprowadzone badania wykazały skuteczność metody SFA w analizie danych geochemicznych. Podejście to pozwala na eliminację zbędnych elementów oraz zwiększenie wariancji przypisanej predykcijnemu czynnikowi mineralizacji, co prowadzi do lepszego zdefiniowania aureoli geochemicznych. Dodatkowo, w badaniu tym oceniono wielowymiarowe podejścia analityczne wraz z technikami uczenia maszynowego. Ponadto, w badaniu oceniono metody analizy wielowymiarowej w połączeniu z technikami uczenia maszynowego. W tym celu wykorzystano sieć neuronową – perceptron wielowarstwowy (MLP) do oceny zawartości złota, srebra, miedzi, ołowiu i cynku w badanym obszarze. Zmienną wyjściową była zawartość konkretnego pierwiastka, natomiast zmiennymi wejściowymi – zawartości pozostałych czterech pierwiastków. W celu optymalizacji modelu zastosowano różne liczby warstw ukrytych oraz szereg funkcji aktywacji. Ostatecznie opracowano model optymalny dla każdego pierwiastka. Modele osiągnęły dokładności wynoszące odpowiednio 95%, 88%, 73%, 80% i 72% dla złota, srebra, miedzi, ołowiu i cynku.

Corresponding author: A. Hezarkhani, e-mail: ardehez@aut.ac.ir

Article contributed to the Editor: 07.01.2025. Approved for publication: 18.03.2025.

Wyniki te wskazują na wysoką efektywność obliczeniową tej metody w ocenie zawartości pierwiastków. Ponadto mapy rozmieszczenia pierwiastków wygenerowane obiema metodami wykazały, że podejście oparte na MLP identyfikowało obszary anomalii z większą dokładnością.

Słowa kluczowe: analiza skupień, współczynniki korelacji, wielowymiarowe metody statystyczne, krokowa analiza czynnikowa, perceptron wielowarstwowy.

Introduction

The geochemical study of waterways sediments is a method applied during the early stages of identifying and locating promising regions. One of the practical challenges in geochemical research is assessing how representative a sample is for predicting the type of mineralization. To identify the best indicators and pathfinder elements for exploration, it is crucial to pinpoint and assess promising areas for a particular mineral. A proposed solution to this issue is the use of multivariate statistical analysis (Halfpenny and Mazzucchelli, 1999; Chandrajith et al., 2001; Grunsky et al., 2009). Factor analysis is a multivariate analytical method frequently used in the geochemical investigation of waterway sediments to identify prospective regions for mineralization (Borovec, 1996; Reimann et al., 2002; Kumru and Bakaç, 2003; Van Helvoort et al., 2005; Sun et al., 2009). This approach can address specific limitations of factor analysis in geochemical exploration, improve geochemical halos, and produce a more effective geochemical map (Yousefi et al., 2012).

Statistical techniques can be employed to examine the relationship between data in geochemical studies. Before performing multivariate statistical analysis, univariate analyses are required. Conducting these analyses allows for the examination of the data and the identification of certain characteristics, such as abnormal samples in the study area. The aim of assessing heavy mineral sample results in multivariate statistical studies is to determine the type of mineralization present in the region (Govett, 2013). Considering the extent of the region, the most appropriate sampling technique is the collection of waterways sediments. However, heavy minerals investigations may also prove to be quite beneficial. In this method—unlike geochemical sampling, which seeks to evaluate the overall amount of an element within the rock and analyze its variability and distribution—the examination of the mineralogical phase emphasizes the presence of elements as distinct minerals. Hence, analyzing heavy mineral samples is utilized to pinpoint anomalous regions, identify mineral varieties and possible types of mineralization, understand mineralogical generative relationships, and assess potential paragenesis. The only limitation of this method is that the outcomes are partially quantitative and partially qualitative. Typically, this method

can be highly advantageous when applied alongside the geochemical technique. Taking this into account, and evaluating the advantages and disadvantages of this method with three goals in mind, heavy mineral sampling was conducted in the area. This involved confirming extractive anomalies through the hydrogeochemical approach, determining the dispersion phase of various elements, and addressing any information gaps arising from sampling method constraints or the generation of waterway geochemical findings. Therefore, integrating the findings from these two approaches can offer a more precise representation of the exploration environment.

Evaluating mineral grade is essential for assessing the deposit reserve; therefore, different stages of a mining project, such as feasibility, design, and planning, significantly depend on the accuracy of grade estimation. Geostatistical methods are frequently used to predict grade fluctuations and evaluate reserves in deposits; nevertheless, these techniques encounter several issues, such as uncertainty regarding the validity of assumptions, the determination of variographic parameters, being time-consuming, and the inability to provide accurate estimates under conditions of weak spatial structure of regional variability. As a result, this has led to the investigation of artificial intelligence-driven estimators for grade assessment in recent times. Research in this field involves the use of multilayer perceptron neural networks, support vector regression, and fuzzy neural systems for estimating copper grades (Valizadeh and Sharghi, 2014); optimization methods for neural networks focused on copper grade estimation (Tahmasbi and Hezarkhani, 2011); the application of perceptron-genetic neural networks for analyzing limestone composition (Chatterjee et al., 2008); artificial neural networks for assessing gold grades (Mostafaei et al., 2023); evaluation of grade in a gold placer deposit horizon using a simple ray neural network (Samanta and Bandopadhyay, 2009); predicting geochemical behavior of copper through artificial neural networks (Shirazy et al., 2020); estimating lead and zinc grades while assessing deposit reserves with artificial neural networks (Arinze et al., 2019); analysis of lignite impurities using a fuzzy neural system (Tutmez, 2009); determining platinum grade in a placer deposit horizon with support vector regression (Chatterjee and Bandopadhyay, 2011); geochemical studies along with tonnage and grade via probabilistic neural networks (Singer, 2006); classification of

organic content in sediments (Weller et al., 2005); digitization and categorization of volcanic rock geochemical data using artificial neural networks (Lacassie et al., 2004, 2006), and employing neural networks for locating deposits (Singer and Kousta, 1996).

This study aims to explore the computational effectiveness of the intelligent multilayer perceptron neural network approach for modeling element grades in the Janja polymetallic gold deposit and to compare its results with multivariate statistical methods.

Study area

The Janja region is situated in Saberi town in the province of Sistan and Baluchistan (Figure 1). Saberi town is located at a geographic coordinates of $61^{\circ} 29'$ longitude and $31^{\circ} 1'$ latitude, positioned in the southeast of Iran and to the northwest of Zahedan city.

The Sistan fault zone in eastern Iran is located in the north-south-trending Cretaceous-Tertiary orogenic belt between the Central Iranian and Afghan continental blocks (Agard et al.,

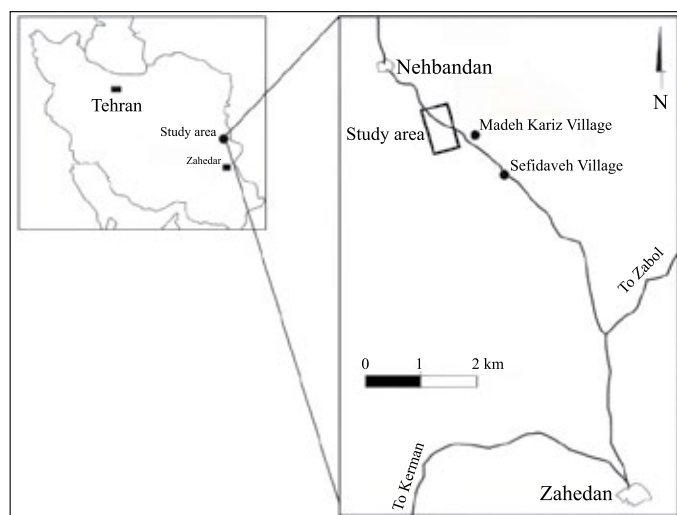


Figure 1. The location of the study area

Rysunek 1. Lokalizacja obszaru badań

2011). Rock outcrops in this area include sedimentary rocks such as sandstone and shale of the Shemshak Formation, Jurassic limestones, and marly shales, in addition to Quaternary sediments. Igneous outcrops in the area also include dacite and layered tuffs (Aghanabati, 2004) (Figure 2).

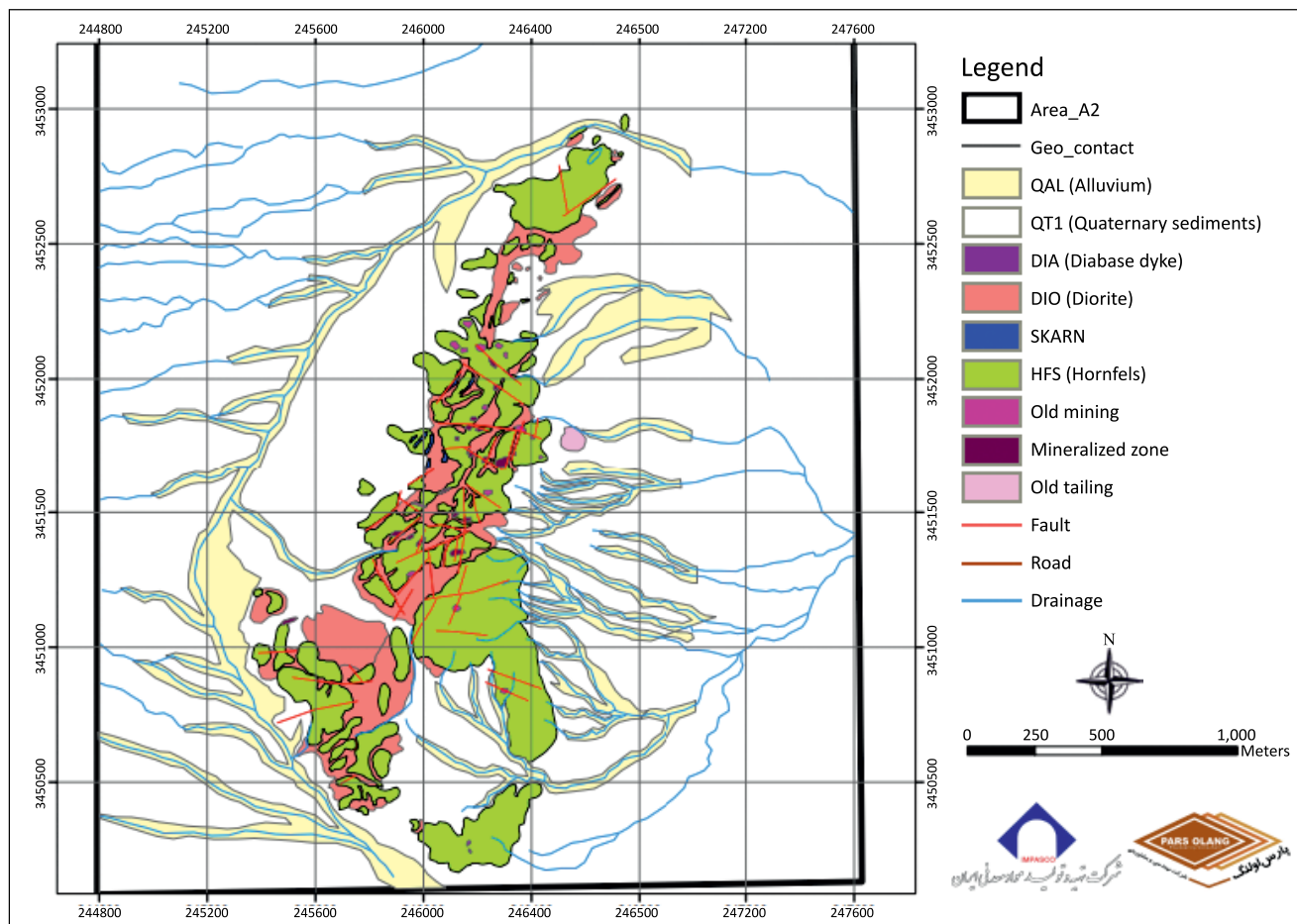


Figure 2. Geological map of the study area

Rysunek 2. Mapa geologiczna obszaru badań

Sefidabeh Formation unit

The oldest geological unit found in this region is the Sefidabeh Formation, dating from the Late Cretaceous to Early Paleocene epochs. In the study area, this unit appears as metamorphosed equivalents in the study region (hornfels and skarn). A Miocene-aged mass of diorite to quartz diorite has intruded into this sedimentary layer, causing contact metamorphism and resulting in the formation of hornfels and skarn from the Sefidabeh layer. The dominant alteration in this unit is propylitic, accompanied by the development of epidote, chlorite, and calcite. In some outcrops, localized sericite alteration is also present.

Hornfels unit

Due to the intrusion of the diorite-quartz mass in the area, the Sefidabeh sedimentary formation has undergone constant metamorphism, resulting in the formation of the hornfels unit. In surface observations, this unit is frequently seen on the diorite-quartz mass. The color of this unit varies from pea yellow to light green, depending on the quantity of epidote and chlorite present. Microscopic analyses indicate that this unit mainly consists of epidote, chlorite, quartz, and calcite.

Skarn Unit

This unit formed as a result of the intrusion of the diorite-quartz mass into the Sefidabeh sedimentary formation. It typically develops with limited extent at the boundary of the diorite mass or in more reactive layers of the Sefidabeh formation, such as sections of impure limestone. The thickness of this unit ranges from a few tens of centimeters to several meters. In surface exposures, the formation appears dark green due to the abundant presence of epidote and chlorite. The primary minerals composing the skarn unit comprise epidote, chlorite, calcite, and garnet.

Diabase dyke unit

This unit has limited extent in the study area. Only two minor exposures of these dykes have been observed in the Hornfels unit located in the southern part of the area. In surface outcrops, the unit appears dark gray to black, and many of its exposures are weathered and fractured. The unit exhibits a northeast-southwest orientation, and its thickness ranges from 1 to 3 meters. Microscopic analyses show that the unit comprises faceted pyroxene crystals and plagioclase within a more finely crystalline groundmass of the same composition. The dominant alteration in these dykes is weak chloritization.

Quaternary sedimentary units

These are the most widespread units in the area and consist of two components. One component comprises coarse sedi-

ments, which are most prevalent, while the other consists of alluvial sediments found in the channels of rivers and streams in the area.

Materials and methods

Geochemical sampling

The design of the sampling network is a crucial phase in geochemical studies. Its effectiveness depends on the type of exploratory deposit and the scale of the study. To achieve this goal, it is essential to collect and integrate all relevant information, including bedrock classification (igneous, sedimentary, and metamorphic), intersections between igneous masses and sedimentary layers, varieties of metamorphic facies, and concealed intrusive bodies. In developing the sampling network, information from the 1:100,000 Khunik geological map, along with topographic maps, tectonic structures, and aerial and satellite images, was collected and integrated. As a result, 153 river sediment samples were collected from a 144 km² area, yielding an estimated sampling density of approximately 1.1 samples per square kilometer (Figure 3). The samples were analyzed using the ICP-AES method for 44 elements. It is important to note that the Fire Assay preparation technique was employed for gold analysis. Furthermore, the heavy mineral sampling network was designed to collect one sample for every 2.4 km². As a result, 59 heavy mineral samples were planned and collected, taking into account the necessary quantity of samples required to generate heavy mineral samples in both outcrop and plain areas. During field operations, a total of 7 liters of unsieved samples were collected. Collecting heavy mineral samples from locations where coarse particles accompany finer ones yields improved results. Sampling involved excavating holes with diameters of 30–40 cm and depths of 30–50 cm, collecting more than 5 kg of soil, which was sieved through a 20 mesh to obtain the final sample. The collected samples were prepared following volume measurement, mud washing, and drying, using bromoform liquid in the heavy mineral lab. Subsequently, with the use of a magnet, the recovered minerals were classified into three groups: magnetic (AA), weakly magnetic (AV), and non-magnetic minerals. The results were then reported according to the percentage of minerals identified using a binocular microscope.

Factor analysis

The observable random vector $x = (x_1, x_2, \dots, x_p)^T$, which has p components, has a mean represented by $\mu = (\mu_1, \mu_2, \dots, \mu_p)^T$, and a covariance matrix denoted by $\Sigma = (\sigma_{ij})_{p,p}$. The factor model suggests that x is linearly dependent on several m -dimensional vectors of unobserved variables $f = f_1, f_2, \dots, f_m$,

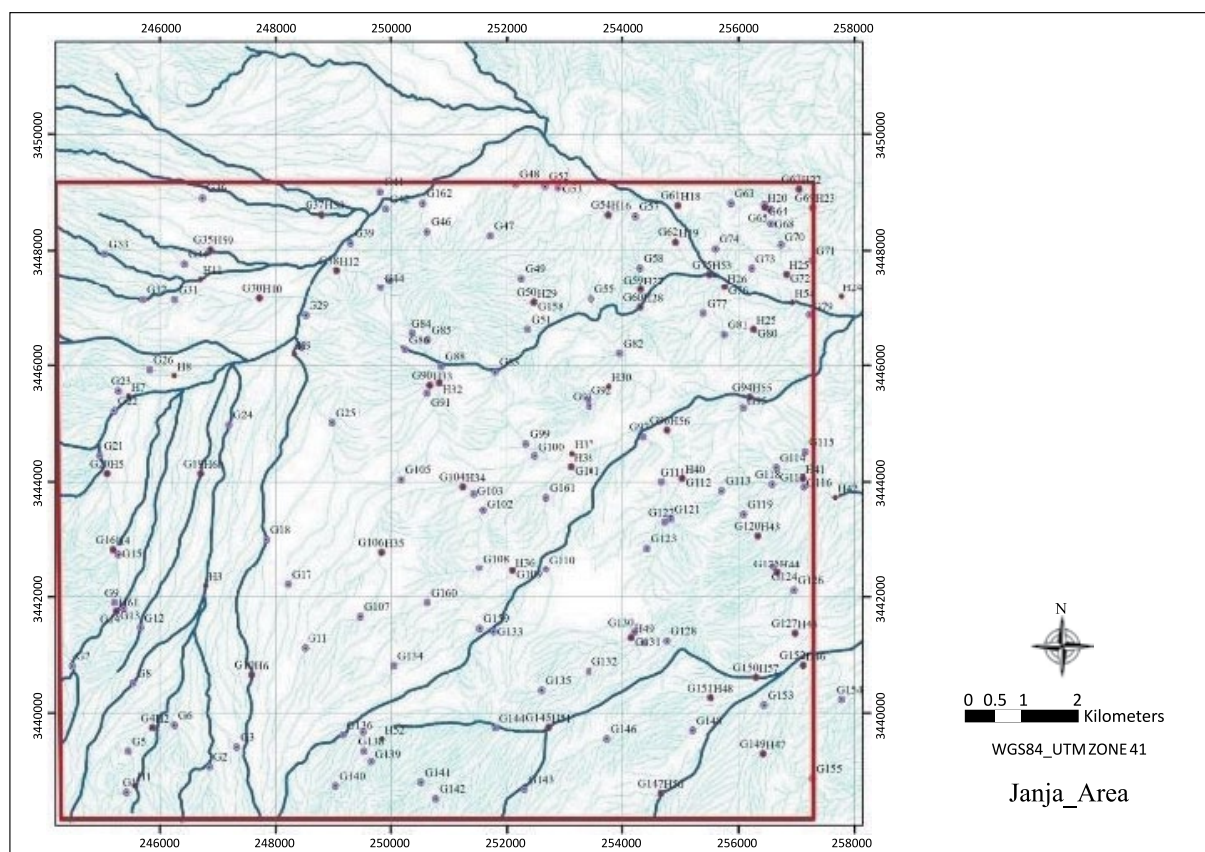


Figure 3. Geological sampling network map of the study area (Ebdali and Hezarkhani, 2024)

Rysunek 3. Mapa geologicznej sieci pobierania próbek na badanym obszarze (Ebdali i Hezarkhani, 2024)

known as common factors, and a p -dimensional vector of unobserved variables $\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$, known as specific factors. The matrix representation of FA is given by $x - \mu = Af + \varepsilon$, where $A = (\sigma_{ij})_{p \times m}$ ($i = 1, 2, \dots, p; j = 1, 2, \dots, m$), is known as factor loading.

In FA, the goal is to account for the covariance or correlation between variables. The covariance matrix is represented as $\Sigma = AA^T + D\hat{A}$, the variance matrix is $\text{var}(x_i) = \sum_{j=1}^m a_{ij}^2 \sigma_i^2$, and the common variance is $h_i^2 = \sum_{j=1}^m a_{ij}^2$.

Multilayer Perceptron Neural Network

Artificial neural networks are inspired by the complex structures of the human brain, in which millions of neurons communicate to solve problems and store information. The task of a neural network is to learn. This process begins with training, or gaining experience, which is carried out using a series of desired input and output data. In this way, a set of correct inputs and outputs is presented to the network, and the neural network uses these inputs to construct a complex mathematical model that generates the correct response when provided with new inputs (Krose and Smagt, 1996).

The multilayer perceptron neural network is a multi-layered feedforward neural network. In this network, spatial relationships between input and output variables are detected by a group of processing units called neurons, organized into input, hidden, and output layers. This network detects the spatial variability of mineral grade using mapping functions ϕ , created by a set of association weights between the inputs and the output (Mahmoudabadi et al., 2009). The output o (element grade) can be considered a function of the association weights $w: o = \phi(x)$ (Dutta et al., 2010). Optimization algorithms can be used to determine the optimal structure of a multilayer perceptron neural network in terms of the number of hidden layers and the number of neurons in those layers.

Discussion

Data processing

Exploration objectives can be achieved through the analysis of geochemical data. In geochemical exploration, numerous samples and multiple factors must be considered. As a result, the application of statistics and probability in data analysis has become essential. Data processing is considered acceptable when the laboratory error (assessment of repeated samples)

is below 10%. By incorporating additional methods such as heavy mineral studies, anomaly control, and comparison of geochemical anomalies with data processing results, the accuracy and reliability of the analysis can be improved. The data management procedure includes organizing raw laboratory data, substituting censored data, conducting statistical analyses, detecting outliers, and presenting results using graphs and tables.

Analysis error determination

To effectively use the measurement results, it is essential to understand the degree of confidence in the measurements. Quality control tests are performed to identify any errors in the preparation and analysis steps. The reliability of measurement data generally depends on the magnitude of both random and systematic errors. It is important to note that the sampling error is primarily systematic, while random error arises from natural variability and is present to some extent in all types of measurements (Parsons and Clegg, 2009). The relative error of analyses is calculated using Equation (1):

$$e = 2/n \left[\sum |x_i - y_i| / (x_i + y_i) \right] \times 100 \quad (1)$$

where:

e – relative error value,

n – number of repeated samples,

x_i, y_i – values measured in the main sample and its corresponding repeated sample.

The sample analysis error values were calculated for different elements (Table 1). The results show that cadmium and sulfur have errors exceeding 20%, while all other elements have errors within acceptable limits.

Table 1. Mean calculated error between two pairs of corresponding repeated samples using the ICP-AES technique to evaluate lab precision

Tabela 1. Średni obliczony błąd między dwiema parami odpowiednich powtarzających się próbek przy użyciu techniki ICP-AES do oceny dokładności badań laboratoryjnych

Elements	Average Lab Error	Elements	Average Lab Error
Ag	13	Mo	13
Al	3	Na	2
As	11	Ni	6
Ba	3	P	5
Be	3	Pb	15
Ca	4	S	28
Cd	21	Sb	6
Ce	3	Sc	3
Co	5	Sr	2
Cr	11	Th	3
Cu	7	Ti	1
Fe	3	U	6
K	3	V	3
La	3	Y	4
Li	4	Yb	4
Mg	5	Zn	6
Mn	4	Zr	4

Censored data estimation

Censored data refers to values that lie beyond the sensitivity thresholds of the device, either exceeding the upper threshold or falling below the lower threshold. In the context of relative measurements, such as differentiating between background and anomaly, the presence of censored data can lead to inaccurate evaluations. To address this issue, estimated values should

Table 2. Number of censored data and their replacement values in the dataset

Tabela 2. Liczba danych cenzurowanych i ich wartości zastępcze w zbiorze danych

Element	Unit	Detection Limit	Total Sample	No. of Censored	Percentage [%]	Replace Value
Au	ppb	5	153	135	88.2	3.75
Ag	ppm	0.1	153	0	0	0.075
Al	ppm	100	153	0	0	75
As	ppm	0.5	153	1	0.6	0.375
Ca	ppm	100	153	0	0	75
Cd	ppm	0.1	153	0	0	0.075
Ce	ppm	1	153	0	0	0.75
Co	ppm	1	153	0	0	0.75
Cr	ppm	1	153	0	0	0.75
Cu	ppm	1	153	0	0	0.75
Fe	ppm	100	153	0	0	75
La	ppm	1	153	0	0	0.75
Li	ppm	1	153	0	0	0.75

cont. Table 2/cd. Tabela 2

Element	Unit	Detection Limit	Total Sample	No. of Censored	Percentage [%]	Replace Value
Mg	ppm	100	153	20	13	133
Mn	ppm	5	153	0	0	3.75
Mo	ppm	0.5	153	0	0	0.375
Ni	ppm	1	153	0	0	0.75
P	ppm	10	153	0	0	7.5
Pb	ppm	153	153	0	0	0.75
S	ppm	50	153	0	0	37.5
Sb	ppm	0.5	153	0	0	0.375
Sc	ppm	0.5	153	0	0	0.375
Th	ppm	0.5	153	0	0	0.375
V	ppm	1	153	0	0	0.75
Yb	ppm	0.2	153	0	0	0.15
Zn	ppm	1	153	0	0	0.75

be used to substitute censored data. Various techniques can be employed to estimate censored data values, one of which is the simple substitution method (Sanford et al., 1993). This approach entails replacing values exceeding the sensitivity threshold with 4/3 of that threshold, and those below it with 3/4 of the threshold. If the proportion of censored data is less than 10% of the total dataset, this method is usually considered acceptable. The outcomes of substitution for the censored are presented in Table 2. The table indicates that only 2 elements contain missing data. For gold, 88.2% of the data is censored and cannot be substituted, with only the samples containing the censored data being excluded from this element, allowing the remaining available data to be used. In addition, in the case of magnesium, 20 samples were censored, representing 13% of the total.

Outlier values elimination

Outlier are values that lie significantly outside the data range and are considered as either very low or very high values. These values are important, as they might indicate anomalies associated with mineralized areas. However, they may also arise from substantial sampling or laboratory errors.

Most statistical distributions in exploratory projects are non-normal and typically exhibit right-skewed distributions (Reimann and Filzmoser, 2000). Distributions of this kind emphasize the importance of the values located on the right side of the distribution. These exceptionally high values may represent anomalies (on a regional scale) or rich ore deposits (on a local scale). The box plot diagram was utilized in this study to identify outlier values. Boxplots serve as a method to illustrate the data distribution within statistical populations, highlighting quartiles and the interquartile range. These graphs show the minimum, 25th percentile, median, 75th percentile,

and maximum values, identifying outlier or extreme samples using the interquartile range, where outlier samples satisfy equations (2) and (3).

$$\text{Outlier} = 1.5 (75\text{th} - 25\text{th}) \quad (2)$$

$$\text{Extreme} = 3 (75\text{th} - 25\text{th}) \quad (3)$$

Samples that exceed these thresholds are classified as outliers or extreme values. The elongation of the tails in boxplots may suggest the presence of samples exhibiting unusual values. From the boxplots, it is possible to identify the detection limit level, interquartile range, median position, spread of outliers and extreme values, distinctions between normal and anomalous distributions, as well as potential for unusual values.

It is important to emphasize that during Pearson's correlation analysis and cluster analysis, outliers were adjusted solely by replacing them with the highest outlier value to compute the standard function F . Furthermore, in differentiating anomalies from background values, mean values and standard deviations were calculated while retaining outlier values, and no substitutions were made in the computations. Once the censored values were replaced, a boxplot was generated for different elements. The boxplot illustration presents gold and silver as examples (Figure 4).

Data normalization

It is common for geochemical data to exhibit a log-normal distribution, while a normal distribution is rare. Therefore, prior to performing calculations, it may be necessary to convert the abnormal data into normally distributed data through the use of transformations. To achieve this, logarithmic transformation, three-parameter transformation, generalized exponential transformation, or other available techniques can be employed (Pawlowsky-Glahn et al., 2015). Nonetheless, logarithmic transformation is often used due to its simplicity of implementation.

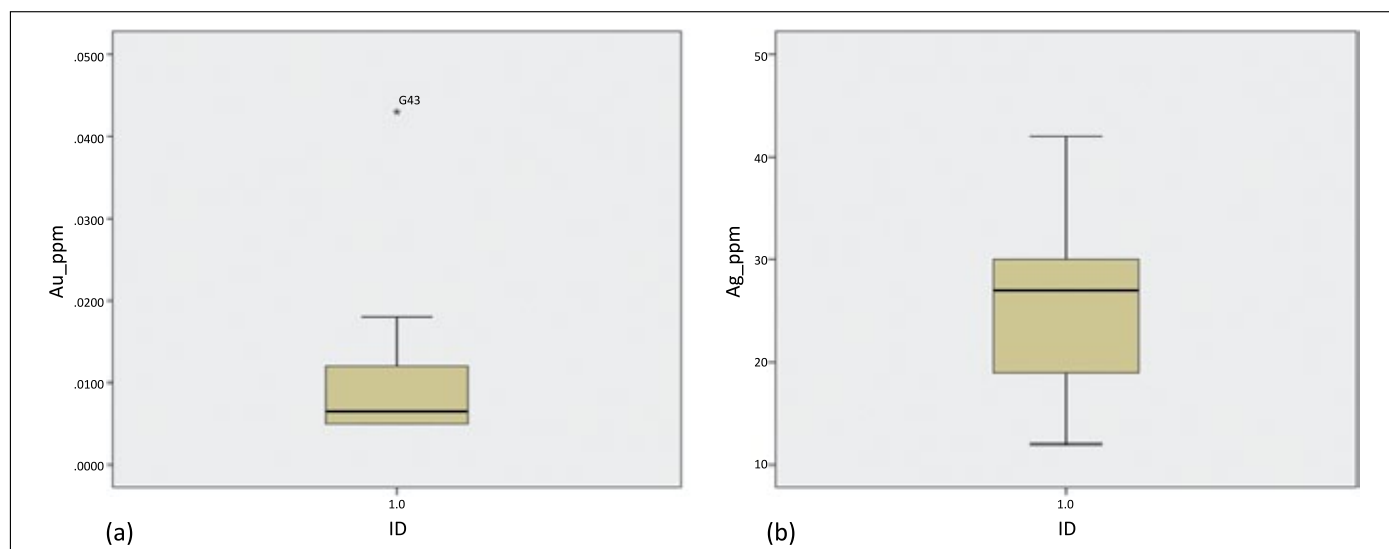


Figure 4. Boxplot of gold (a) and silver (b)

Rysunek 4. Wykres pudełkowy dla złota (a) i srebra (b)

When the data follow a normal distribution, the average of the sample group provides a more accurate estimate than the average of the entire population. Despite the advantages of converting the data distribution to normal, such transformation should not always be applied. The decision to use a transformation depends on a specific context. Transformations may be appropriate if a desired outcome can be achieved using the transformed values, particularly when there are no limitations in accessing the or when reverse transformation can be performed easily. If the estimation derived from the initial data is sufficiently precise, it is preferable to avoid transformation when possible. The F function is a statistical measure used to determine whether a normal distribution exists among different elements. This function is calculated using equation (4).

$$F = 2 \left| \text{Sk} \right| + \left| 3 - K \right| \quad (4)$$

where:

Sk – Skewness value,

K – amount of stretching in the distribution of data.

Hence, a distribution (raw or logarithmic) with an F value closer to zero is considered more typical. Equations (5) and (6) are used to calculate the mean and standard deviation by applying inverse transformations in a normal log transformation to find the values of X and S .

$$S = \bar{X} \times \sqrt{(e^{\beta^2} - 1)} \quad (5)$$

$$\bar{X} = e^{\alpha + (\beta^2/2)} \quad (6)$$

where:

α – mean of logarithmic data,

β – standard deviation of the logarithmic data.

The geochemical data were examined after replacing outlier values and applying logarithmic transformation. By examining the shape of the distribution function, the maximum value of each element, and the skewness of its distribution, it was possible to assess whether a given element has a favorable detection potential. In this region, zinc, lead, arsenic, copper, and cadmium exhibit significant skewness and peak values relative to global background levels, indicating a relative potential for mineralization.

Outlier values may suggest the occurrence of a special geological phenomenon. Unusual concentrations of elements such as zinc, lead, arsenic, silver, cadmium, and antimony may indicate metal mineralization or localized concentration of these elements. Variations in lithological units, such as the presence of carbonate rocks in the region, may result in the occurrence of elements such as potassium, cerium, calcium, and vanadium.

Establishing the normality of each element can be done by examining its minimum, maximum, average, median, standard deviation, and skewness. Based on this analysis, silver and copper exhibit nearly normal conditions, while lead and zinc show standard logarithmic conditions.

Multivariate statistical calculations

Multivariate analysis refers to the concurrent assessment of multiple variables. Based on this definition, many multivariate methods are fundamentally built upon and have evolved from univariate and bivariate analyses (Denis, 2021; Maiti, 2022; Mardia et al., 2024). Each distinct group of elements exhibits a degree of sensitivity that is relatively similar under different environmental conditions. Identifying the shared genetic connections among different elements can enhance

our understanding of changes in geochemical environments. These connections can serve as a straightforward reference for interpreting deposit types, while conversely, the accumulation of some elements might suggest the presence of unimportant and misleading anomalies. In general, grasping the genetic relationships among elements is crucial for the accurate analysis of geochemical information.

Overall, multivariate statistical analyses offer two main advantages. The composite halos obtained through multivariate statistical methods show a greater correlation with the deposit's structural features, geology, and origin, resulting in a more clearly defined relationship between the elements. Moreover, by employing composite halos, random errors can be reduced, and the quantity of data and maps minimized, leading to more effective outcomes. In this research, genetic connections were illustrated through correlation coefficients among different components, along with cluster analysis and factor analysis. The factor analysis technique is also employed for generating multivariable maps and achieving overall multivariable outcomes.

Correlation coefficients calculation

Opting for two-variable studies as the main approach facilitates the detection and examination of genetic and secondary relationships among variables. The examination of two variables includes analyzing the scatter plot and calculating the correlation coefficient between them (Ghannadpour and Hezarkhani, 2015; Ghannadpour et al., 2015). The correlation coefficient and its variations are frequently employed in exploratory stages, particularly in geochemical investigations. When determining correlation coefficients, it is essential to assume that the data follow a normal distribution, similar to other statistical measures. If the data distribution is not normal, Spearman's non-parametric correlation coefficient is used to calculate the correlation coefficients, irrespective of the data distribution function. The correlation coefficient matrix was computed by examining the chemical composition of 153 river sediment samples for various elements. Based on this information, copper is closely associated with lead, zinc, arsenic, and cadmium. Consequently, rather than anomalies in these elements, copper mineralization can be observed alongside other essential elements such as lead and zinc, as well as hydrothermal alteration. The correlation between copper and lead is 0.89, whereas with zinc it stands at 0.95.

A significant correlation exists between lead and zinc, along with cadmium and arsenic, which may indicate the possibility of vein mineralization in the area. The correlation of cadmium and arsenic with lead is 0.90 and 0.96, respectively, whereas with zinc it is 0.82 and 0.95. Further correlations include the relationships between iron and chromium, cobalt, manga-

nese, and vanadium, which have little ore value and probably originate from the lithology and rocks in the surrounding area. It is important to note that these results reveal polymetallic mineralization of lead, zinc, and copper in the exploration region, resulting in the formation of a significant system as indicated by the data.

Cluster analysis

Cluster analysis is a multivariate technique intended to categorize variables or samples by their similarities within the group and distinctions between groups. In this instance, the clustering algorithm and the correlation coefficient served as the main methods for assessing similarity (Barnett et al., 2014). The results of the centroid cluster analysis for the geochemical data are classified into 3 categories: a) The first group includes lead, zinc, copper, cadmium, and arsenic. In this group, base metals and arsenic elements are primarily identified as tracers, possessing the highest mineralization value and aligning with the polymetallic mineralization discovered in the exploration zone. b) The second category includes iron and cobalt associated with vanadium, indicating mafic to intermediate intrusive bodies in the area, alongside the distribution of the porphyry diorite mass. c) The third group includes manganese and chromium, along with cerium and ytterbium, which are probably related to the intrusive bodies in the region and lack mineralization significance.

Factor analysis

Geochemical reference layers are developed and established through geochemical multivariate analyses to identify the optimal reagent or mixture of reagents for the desired mineralization. Multivariate analyses can evaluate the significance of different combinations of geochemical variables (elements) more effectively (Garrett and Grunsky, 2001; Carranza, 2004, 2010). Considering that mineralization is an uncommon occurrence, and that this is merely one of the numerous factors influencing the alteration of substances in waterway sediments (Bonham-Carter et al., 1989; Carranza and Hale, 1997), deducing the optimal combination of geochemical reagents for a particular mineralization type is likewise a challenge.

Factor analysis is frequently utilized in the analysis of geochemical data. It seeks to explain variability in a set of multivariate geochemical data. This process entails reducing the dimensions of the data and variables to reveal hidden connections among elements by identifying a specific number of factors (Tripathi, 1979). Factor analysis is a statistical method that uses the complete data matrix, such as the correlation or variance-covariance matrix of variables, to produce a linear factor. Each component in the linear relationship is given a coefficient that reflects its importance within the intended factor.

Although it is recognized that geochemical data seldom follow a normal distribution, factor analysis, like many statistical techniques, is ideally performed on data that are either normally distributed or approximate normality. To improve the outcomes of factor analysis and identify the most effective multi-element combinations that indicate mineralization, a refined stepwise factorial analysis (SFA) technique was developed to determine the ideal multi-element combination(s) representing a particular type of mineralization. This method consists of two main phases known as "clean factor extraction" and "evaluation of multi-element effects including importance and calculation of reliable factor scores". Each of these primary phases contains multiple internal stages.

Clean factors identification

At this stage, factor analysis is initially conducted on the main data set that encompasses all the chosen elements. This phase essentially constitutes the first stage of factor analysis. In the results of the initial stage, elements that do not contribute to any factor based on the chosen threshold are excluded, as the presence or lack of even a single element in the input data set can significantly influence the result. These components can be labelled as disruptive geochemical elements and factors that hinder analysis. Therefore, it is essential to remove these elements from the dataset and perform a second-stage factor analysis. This updated analysis may uncover various factors depending on the combinations of elements, their coefficients, and the sample's score values. If, during the second stage, there are components that still fail to meet the defined threshold for involvement in any factor, they must be removed from the dataset. Subsequently, factor analysis must be performed again. If every component in the second stage can be assigned to a factor, then the initial phase of analysis is complete, and the extracted factors will be free from any interfering elements.

Usually, traditional factor analysis results in several factors with different combinations of elements, whereas the SFA technique provides clearer and more reliable outcomes by identifying and discarding non-representative elements. In relation to this issue, and considering the data and the type of mineralization sought, one or more significant key factors can be identified. Nonetheless, in comparison with traditional factor analysis, the number of factors is reduced, the range of changes covered is expanded, and, importantly, the predictive accuracy is improved.

In factor analysis, the primary emphasis is on removing geochemical noise that continuously hinders the statistical assessment of the data. Consequently, clean factors were established by eliminating geochemical noise, resulting in the identification of ultimate mineralization-representative factors that guide the mineralization process. Moreover, since a single

factor influences various variables, methods exist to simplify the interpretation of factors while preserving the same degree of involvement. These methods involve factor rotation, specifically employing the Varimax technique for data rotation in this research (Kaiser, 1958). This method applies the KMO value to evaluate the accuracy of the factor analysis results and to verify the adequacy of the data sample size. High KMO values support factor analysis, whereas low values hinder it. KMO values near 0.9 are considered very suitable for factor analysis, while values close to 0.8 are deemed adequate. Values near 0.7 suggest balanced factor analysis, values close to 0.6 represent average adequacy, and values at or below 0.5 are considered insufficient. It is essential to note that the factor analysis was performed on log-transformed data, and the factor threshold was established at 0.5 within this interval. According to Table 3, the KMO value derived from the SFA is 0.79, which falls within the suitable category as defined in geostatistical literature (Barrie et al., 2023).

Table 3. KMO parameter value in the study area dataset

Tabela 3. Wartość parametru KMO w zbiorze danych badanego obszaru

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.79
Bartlett's Test of Sphericity	Approx. Chi-Square	4579.08
	Df	105
	Sig.	0

Factor analysis was performed exclusively on arsenic, cadmium, cerium, cobalt, chromium, copper, iron, manganese, lanthanum, nickel, lead, sulfur, ytterbium, vanadium, and zinc. Analysis of these 15 elements resulted in 4 factors, as shown in Table (4). This analysis includes only the clean factors, with all geochemical interferences removed.

According to Table 4, the first factor accounts for around 50.3% of the variance and includes cerium, cobalt, chromium, copper, manganese, lanthanum, lead, and zinc. This factor distinguishes two separate categories of elements. The main group is mainly influenced by elements in the silicate minerals network (lanthanum, cerium, cobalt, and nickel), whereas the secondary group is influenced by elements found in ore-forming minerals (lead, zinc, and copper). These elements may be found in both igneous rocks and mineral veins. In general, this factor is particularly important due to the presence of copper, lead, and zinc. The second factor, which represents about 23.65% of the variance, includes arsenic, cadmium, copper, lead, and zinc. The presence of these elements together is significant. This factor is associated with the development of ore in poly-metallic hydrothermal veins or intrusive bodies, which aligns with expectations based on their quality in geochemical data.

Table 4. Result of factor analysis conducted in the study area**Tabela 4.** Wyniki analizy czynnikowej przeprowadzonej w obszarze badań

Factor	Component			
	1	2	3	4
As	0.574	0.700	-0.017	0.095
Cd	0.514	0.839	-0.014	-0.019
Ce	0.766	-0.332	0.421	0.190
Co	0.886	-0.341	-0.199	-0.026
Cr	0.837	-0.35	0.147	0.091
Cu	0.733	0.616	0.047	-0.088
Fe	0.846	-0.346	-0.182	-0.264
La	0.804	-0.355	-0.012	0.238
Mn	0.934	-0.274	0.158	-0.104
Ni	0.411	-0.128	-0.486	0.736
Pb	0.602	0.782	-0.017	0.022
S	-0.166	0.167	0.679	0.281
Yb	0.720	-0.338	0.509	-0.122
V	0.693	-0.334	-0.403	-0.260
Zn	0.741	0.650	-0.047	-0.068

The third factor includes sulfur and ytterbium. This factor is unlikely to be related to mineralization, given the common values in this area. The fourth factor includes nickel, which accounts for around 6% of the variance and does not appear to be particularly significant for exploration based on its values in the exploratory area.

Multilayer perceptron estimator

Data augmentation

As previously noted, the database of the study area contained only 153 geochemical samples. This posed a significant problem due to the limited number of samples, potentially resulting in imbalances in error calculations. The high imbalance in computations can hinder the classification rules, leading to inadequate model estimation and accuracy assessment. Moreover, a limited number of samples in the training data causes models to struggle to capture the inherent complexity of the input-output relationships in high-dimensional data.

To address these problems and enhance the training of machine learning models, data augmentation techniques were employed to expand the data space by generating artificial samples based on the fundamental structure of the actual data (Chatterjee et al., 2022). Various data augmentation techniques have been proposed, including flipping and rotation (Simard et al., 2003), clipping and scaling (Dieleman et al., 2015), and altering the intensities of the RGB channels (Krizhevsky et al., 2012). The key principle of data augmentation is that the applied transformations should not change the semantic meaning of the labels (Hinton et al., 2012). A suitable data augmentation technique is presented that maintains both the diversity and geological integrity of the augmented data. The basic concept behind this augmentation technique is to incorporate random noise into the data while ensuring the information remains geologically meaningful. Neural networks such as the multilayer perceptron emphasize the spatial distribution and correlations of the data rather than the values at specific points (typically the cell for deposit or non-deposit). Therefore, adding noise to the data can not only preserve the majority of its spatial features but also address the issue of limited training samples. In this research, two data augmentation techniques were used: a) Adding noise: In this technique, random noise drawn from a normal distribution is added to the values (including Au, Ag, Cu, Pb, and Zn). This noise is scaled to the standard deviation of the values, giving the data a more natural and realistic appearance. b) Shifting coordinates: In this method, the coordinates (X, Y, and Z) are randomly shifted within a designated interval (default 0.1), helping to generate data variability and simulate different conditions. Ultimately, the original data were merged with the newly generated data from these two techniques, increasing the total data count from 153 to 4105.

Modelling

According to geochemical sampling of the study area, the grades of gold, silver, copper, lead, and zinc were available. Initially, the mean and standard deviation were used to distinguish the anomalous population from the background (Table 5).

Using the derived intervals, the elemental grade values were categorized into 5 classes, ranging from class 1 (of least

Table 5. Classification values in the Rose method**Tabela 5.** Wartości klasyfikacji w metodzie Rosego

Classified	Ag [ppm]	Au [ppm]	As [ppm]	Cu [ppm]	Pb [ppm]	Zn [ppm]
MIN_AVE	0.12_0.25	0.005_0.009	6_11.6	19_26.2	9_22.1	46_69.8
AVE_AVE + STDEV	0.25_0.32	0.009_0.013	11.6_17.2	26.2_29.3	22.1_36.8	69.8_82.4
AVE + STDEV_AVE + 2STDEV	0.32_0.39	0.013_0.018	17.2_22.9	29.3_32.3	36.8_51.4	82.4_95.1
AVE + 2STDEV_AVE + 3STDEV	0.39_0.46	0.018_0.022	22.9_28.6	32.3_35.4	51.4_66.1	95.1_107.7
AVE + 3STDEV_MAX	>0.46	>0.022	28.6_45.7	35.4_36	66.1_128	>107.7

exploration importance) to class 5 (of greatest exploration importance). Subsequently, to assess the grade using the intelligent multilayer perceptron method, the model output was the grade of one of the elements (target variable), and the model input consisted of the grades of the other four elements. Next, the hyperparameters influencing the performance of the multilayer perceptron estimator for assessing the grade in the test set were determined for each element (Table 6).

It is important to highlight that Adam is a widely used adaptive learning rate optimization algorithm in machine learning, offering improvements over traditional gradient descent methods. It calculates individual learning rates for each parameter based on both momentum (smoothing updates) and RMSProp (normalizing updates based on squared gradients). This solver operates based on an initial setup. The procedure is as follows: initialize parameters θ , the first moment vector $m_0 = 0$, and the second moment vector $v_0 = 0$. Set the learning rate α , the exponential decay rates β_1 and β_2 , and a small constant ϵ . This solver uses the update rule for each iteration t :

1. Calculate gradient: $g_t = \nabla_{\theta} f_t(\theta)$;
2. Update first moment estimate: $m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$;
3. Update second moment estimate: $v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$;
4. Compute bias-corrected first moment estimate:
 $\hat{m}_t = m_t / (1 - \beta_1^t)$;
5. Compute bias-corrected second moment estimate:
 $\hat{v}_t = v_t / (1 - \beta_2^t)$;
6. Update parameters: $\theta_t = \theta_{t-1} - \alpha / \sqrt{\hat{v}_t + \epsilon} \hat{m}_t$.

Adam's advantages include computational efficiency, suitability for large datasets, robustness to noisy gradients, and relative insensitivity to hyperparameter tuning. Potential limitations include generalization issues in some cases and increased memory requirements. It is generally considered a good default choice for many deep learning tasks and for rapid prototyping.

Estimators validation

Following the estimation of the test data by the estimator, the root mean square error (RMSE) and mean absolute error (MAE) criteria were computed using the observed and estimated values in the test set, as outlined in equations (7) and (8), respectively.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - a_i)^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |t_i - a_i| \quad (8)$$

where:

t_i – actual value of the parameter,

a_i – value assessed by the estimator.

Additionally, accuracy, precision, recall, and f1-score criteria were calculated to validate and compare the efficiency of each model (Table 7). The values of these criteria demonstrate the strong evaluative performance of this approach.

Table 6. Optimal hyperparameters for each model

Tabela 6. Optymalne hiperparametry dla każdego modelu

Input variables	Target	Network geometry	Activation function	Hidden layer size	Max iteration	Solver	Learning rate initial
Ag, Cu, Pb, Zn	Au	4/5/10/1	tanh	2	200	adam	0.025
Au, Cu, Pb, Zn	Ag	4/5/10/1	tanh	2	200	adam	0.010
Au, Ag, Pb, Zn	Cu	4/5/10/1	logistic	2	200	adam	0.025
Au, Ag, Cu, Zn	Pb	4/10/1	tanh	1	200	adam	0.010
Au, Ag, Cu, Pb	Zn	4/5/10/1	tanh	2	100	adam	0.025

Table 7. Validation criteria calculated for the multilayer perceptron neural network

Tabela 7. Kryteria walidacji obliczone dla sieci neuronowej – perceptron wielowarstwowy

Input variables	Target	RMSE	MAE	Accuracy	Precision	Recall	F ₁ score
Ag, Cu, Pb, Zn	Au	0.47	0.08	95.78	95.74	95.74	95.74
Au, Cu, Pb, Zn	Ag	0.35	0.12	88.05	87.77	87.77	87.77
Au, Ag, Pb, Zn	Cu	0.68	0.33	73.17	72.80	72.80	72.80
Au, Ag, Cu, Zn	Pb	0.61	0.26	79.08	77.60	77.60	77.60
Au, Ag, Cu, Bp	Zn	0.61	0.30	72.93	72.83	72.83	72.83

Results

Anomalies evaluation using multivariate analysis

Geological findings in a particular area lead to the development of anomaly maps, which are crucial for identifying potential areas. Typically, defining a geochemically anomalous area involves considering parameters such as the number of anomalous samples and their distribution of each element, the trend of the anomaly expansion, the extent of the promising zone, the exact location of the anomalous samples, the intersection of geochemical anomalies with aerial geophysical anomalies and tectonic features, the rock facies in the anomalous environment, and, ultimately, the comparability of the emission value of the target element with its emission value across diverse primary and secondary environments. After conducting a statistical analysis of the data, anomaly maps for gold, silver, copper, lead, and zinc were generated and presented. It is important to note that the maps were developed and displayed using ArcGIS software. The collected samples were statistically analyzed and subsequently classified and color-coded within the software. The classification intervals employed in the neural network method (Table 5) were also used to create the element distribution maps.

The concentration of gold ranges from values below the detection limit of the analyzer to a maximum of 43 ppm. This element exhibits no first- or second-degree anomalies, while third-degree anomalies are more concentrated in the northeast, southeast, and central parts of the area (Figure 5). The largest anomaly of this element is located in the northern section of the region. The predominant lithology in these areas comprises diorite masses that have intruded into the sedimentary rocks of the Sefidabeh formation and associated metamorphosed rocks.

The silver concentration ranges from a minimum of 0.12 ppm to a maximum of 0.42 ppm. According to the distribution map, a significant silver anomaly is present in the northern section of the area (Figure 6). The geology of these areas consists of a diorite body and the altered Sefidabeh formation. This anomaly corresponds with those observed for gold, arsenic, and copper.

Copper concentrations range from a minimum of 19 ppm to a maximum of 103 ppm. According to the copper distribution map, the main anomalies are scattered across the northern, northeastern, and southeastern regions of the area (Figure 7). In terms of lithology, the outcrop area includes units of the Sefidabeh formation and a diorite mass.

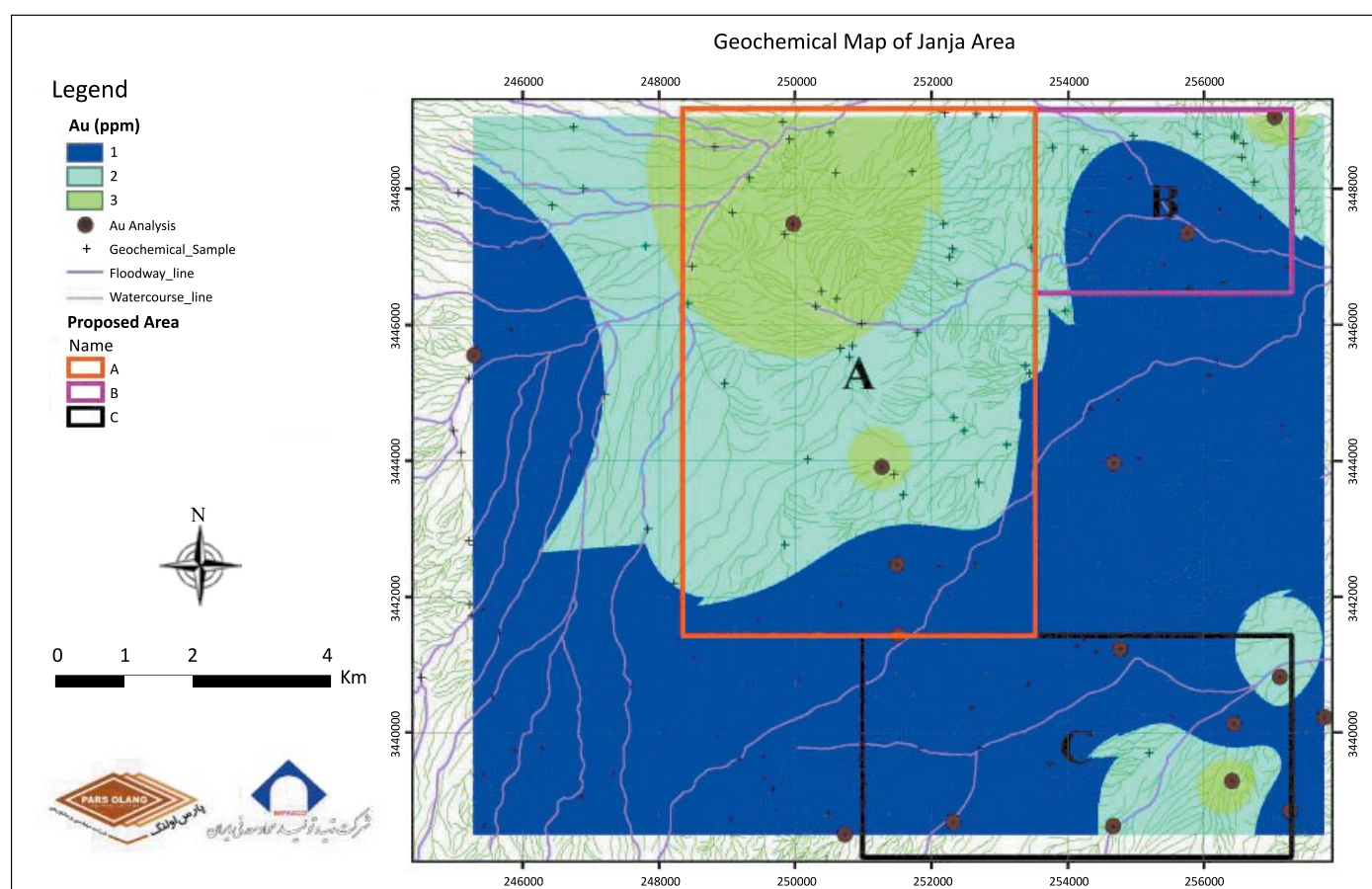


Figure 5. Gold distribution map by Rose method

Rysunek 5. Mapa rozkładu złota według metody Rosego

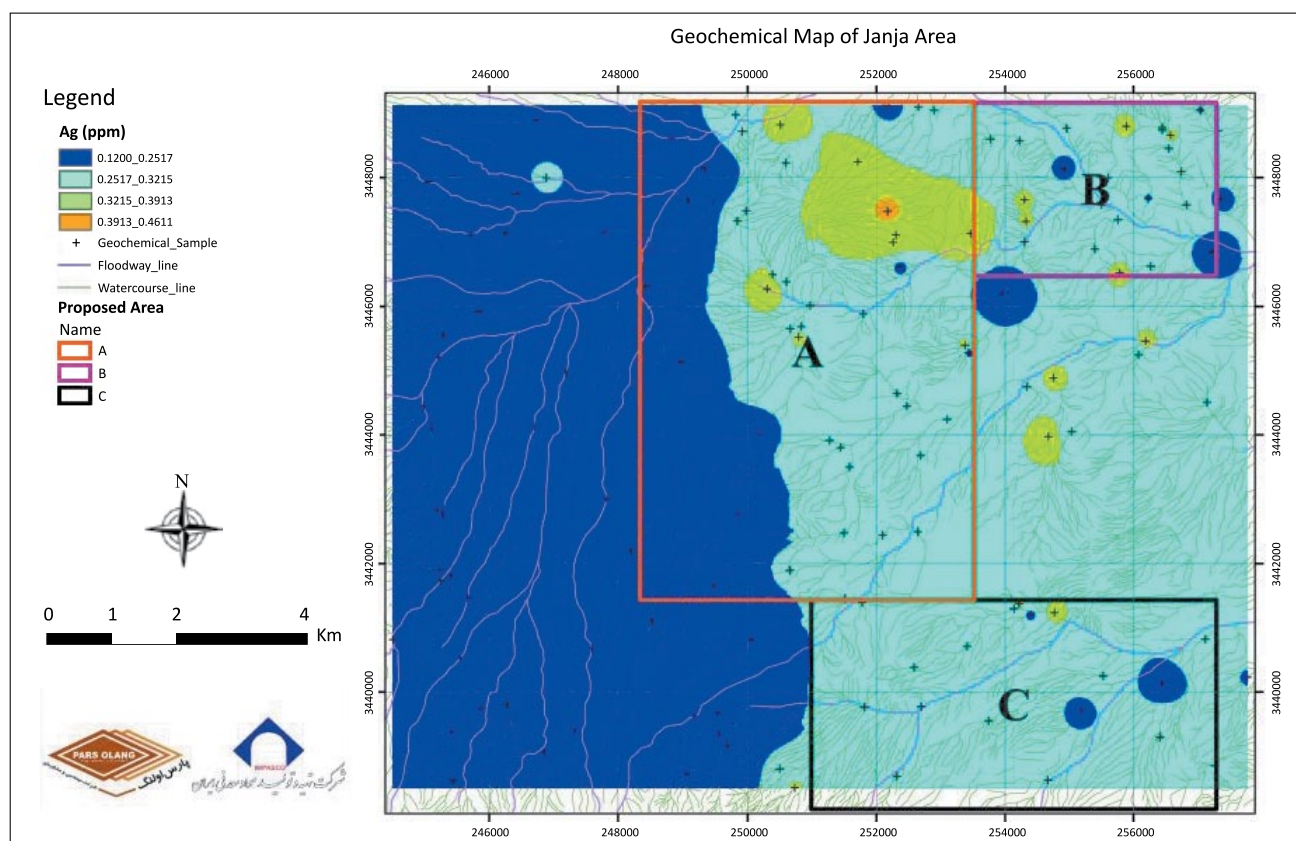


Figure 6. Silver distribution map by Rose method

Rysunek 6. Mapa rozkładu srebra według metody Rosego

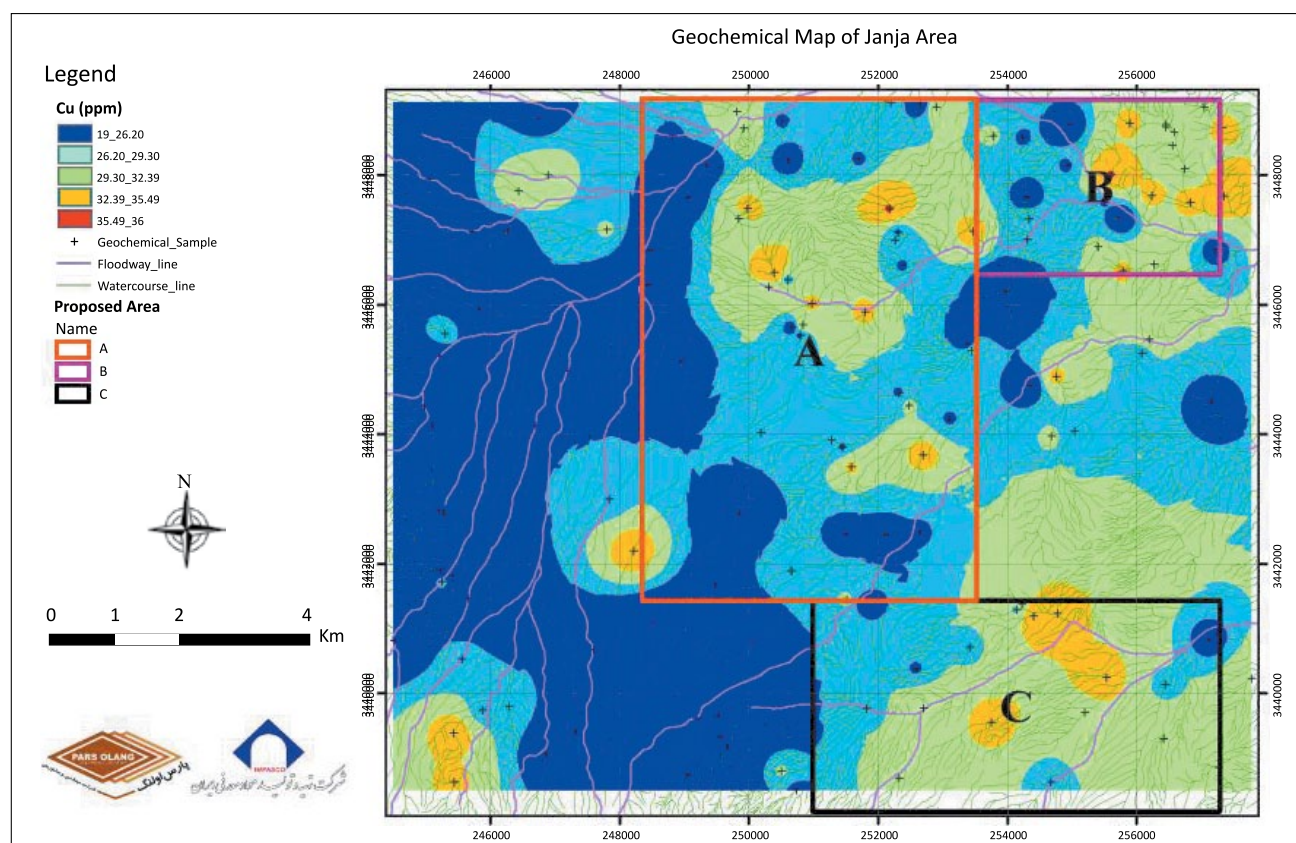


Figure 7. Copper distribution map by Rose method

Rysunek 7. Mapa rozkładu miedzi według metody Rosego

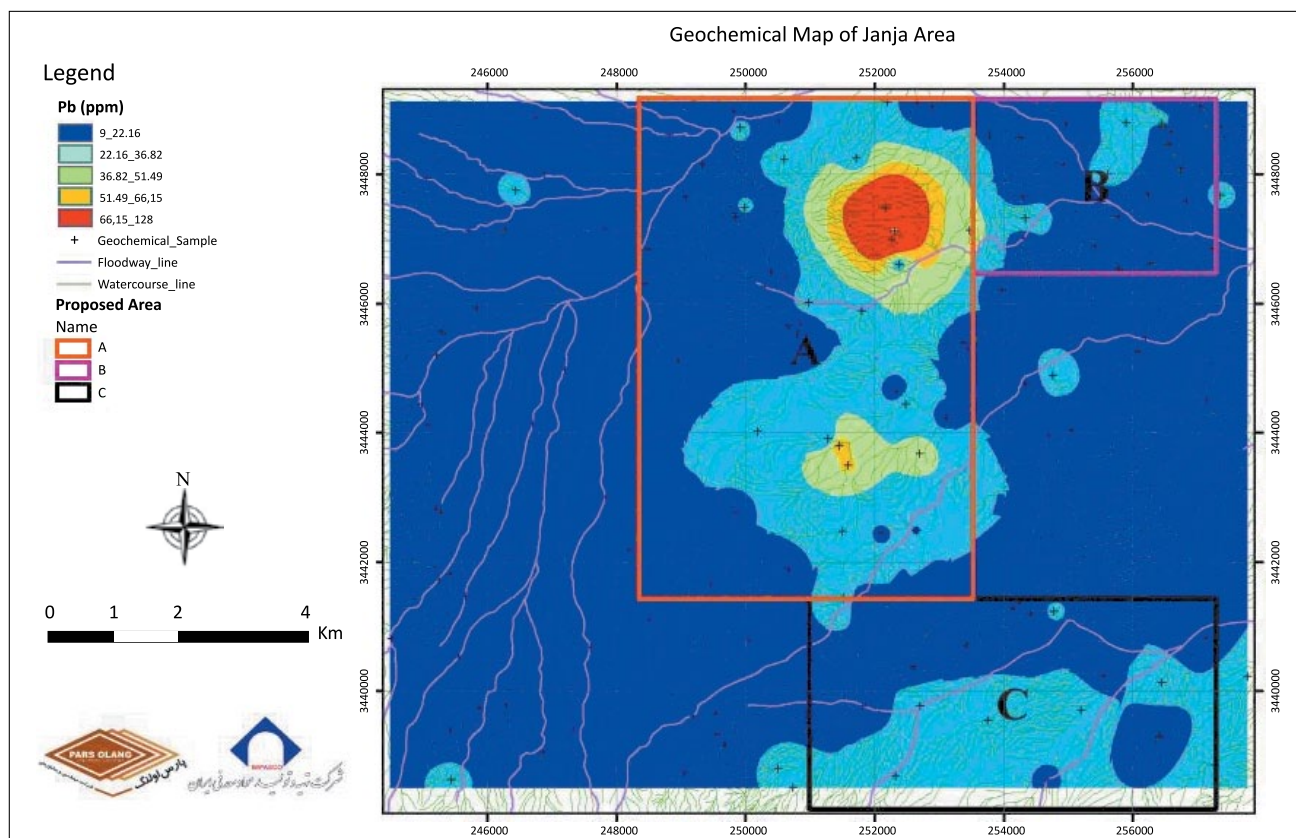


Figure 8. Lead distribution map by Rose method

Rysunek 8. Mapa rozkładu ołowiu według metody Rosego

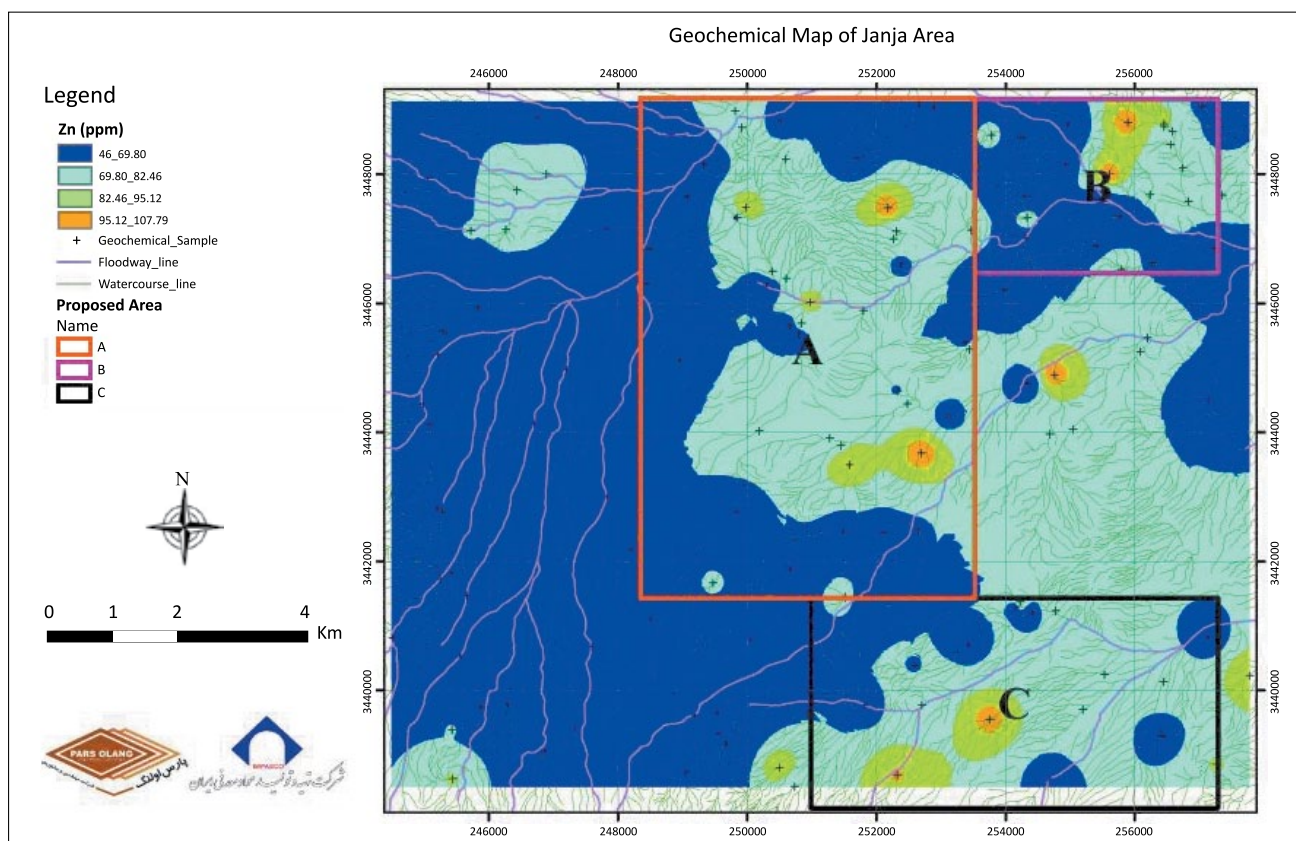


Figure 9. Zinc distribution map by Rose method

Rysunek 9. Mapa rozkładu cynku według metody Rosego

Lead concentrations in the region vary from 9 ppm to 544 ppm. The most notable anomalies of this element are located in the northern and central parts of the area. The primary lead anomaly in the northern area is associated with the lithology of the metamorphosed Sefidabeh formation and the diorite mass (Figure 8).

Zinc levels range from 46 ppm to 538 ppm. According to the zinc distribution map (Figure 9), anomalies are dispersed throughout the northern, northeastern, southeastern, and central regions of the area. Zinc anomalies in the northern and central areas align with those observed for lead.

The presence of gold, arsenic, and silver anomalies in the study area indicates hydrothermal vein type mineralization. Moreover, the occurrence of copper, lead, and zinc anomalies suggests the potential for polymetallic mineralization within the study area.

Heavy minerals studies

The examinations of heavy minerals is essential for identifying both primary deposits and placer deposits during exploration (Dill, 1998). Studies of heavy minerals in Central Europe and Canada (Eyles and Kocsis, 1989), as well as alluvial sediments in New Zealand (Youngson and Craw, 1996), have led to the identification of gold mineralization. Furthermore, the study of heavy minerals can aid in determining the characteristics of the source area (sedimentary-volcanic rock types) and the extent of alteration in the geological formations of the area (Westerhof, 1986). Therefore, identifying the origin and source can significantly influence the concentration of heavy minerals as tracers for assessing mineral potential.

Heavy mineral studies have demonstrated that sampling alluvium can assist in locating prospective areas of mineral deposits. In the study area, 59 samples were selected and analyzed from suitable sites based on the distribution of rock types and extensive alluvial networks. Results from these studies indicate 10 records of mineral deposits (associated with mineralization). The identified minerals include native gold, native copper, galena, cinnabar, malachite, cerussite, pyromorphite, mimetite, vanadinite, and wulfenite. Additionally, chromite appeared as rounded grains in most heavy mineral samples, which was considered unusual, as there are no ultramafic outcrops or chromite mines in the vicinity. Based on this observation, analyses of heavy minerals were conducted on 2 sand samples from the region, both of which contained chrome. This study found that chromite minerals from distant areas, such as the 1:100,000 Bandan sheet containing ophiolitic formations and chromite mines, were transported to the region by wind-blown sand. This phenomenon is likely explained by the 120-day winds of Sistan. Various other minerals such as magnetite, hematite, sapphire, andalusite, pyrite, oxidized

pyrite, limonite, and others were also frequently recorded. Due to the semi-quantitative characteristics of the results, conventional statistical inferences cannot be directly applied to heavy mineral studies. Given the importance of ore deposits in the study area, attempts were made to represent minerals exhibiting similar behaviors or distinctive traits (such as the presence of a specific element) on a single map. The Janja region is known for significant minerals belonging to the gold, copper, and lead mineral families. The heavy mineral distribution map (Figure 10) illustrates the anomalous regions recognized for various minerals based on the heavy mineral analyses.

The results of heavy mineral investigations indicate that the study area is rich in mineral resources. Celestine and barite, as non-magnetic minerals, hold significant value. Additionally, other non-magnetic minerals such as apatite, rutile, zircon, leucosene, sphene, and pyrite are present in minor and scattered quantities, along with calcium carbonate measured in grams per ton. Minerals found in the medium magnetic section, also measured in grams per ton, include hematite, ilmenite, pyroxene, amphibole, and occasionally pyrite. The only mineral present in the hypermagnetic fraction is magnetite; however, its concentrations render it economically unviable.

Ore-forming minerals, with the exception of a few such as magnetite, ilmenite, chromite, and hematite, fall into the category of non-magnetic minerals. Gold occurs as part of the crystal structure in non-magnetic minerals and can be detected through laboratory methods (Carman, 2005). In general, heavy minerals studies regarding ore significance in the area can be classified into three categories (Figure 11): 1) Main minerals associated with gold mineralization, including native gold and cinnabar. 2) Indicator minerals of copper mineralization, such as native copper and malachite. 3) Indicator minerals associated with polymetallic mineralization, such as lead, cerussite, mimetite, pyromorphite, vanadinite, and wulfenite.

Lithogeochemical explorations

Geochemical exploration is a technique used to locate mineral deposits, which contributes to reducing exploration costs and identifying promising regions for further investigation (Prendergast, 2007). Lithogeochemical studies, as an exploration method, involve examining the distribution and concentration of various elements in rocks, detecting anomalies, assessing geochemical halos, and developing zoning models. These approaches aid in revealing concealed mineral deposits (Venkataraman et al., 2000).

Bedrock samples are commonly evaluated to detect primary dispersion halos associated with hidden mineral deposits (Eilu et al., 2001). Lithogeochemical sampling was conducted in both the northern and southern drainage basins of the study area.

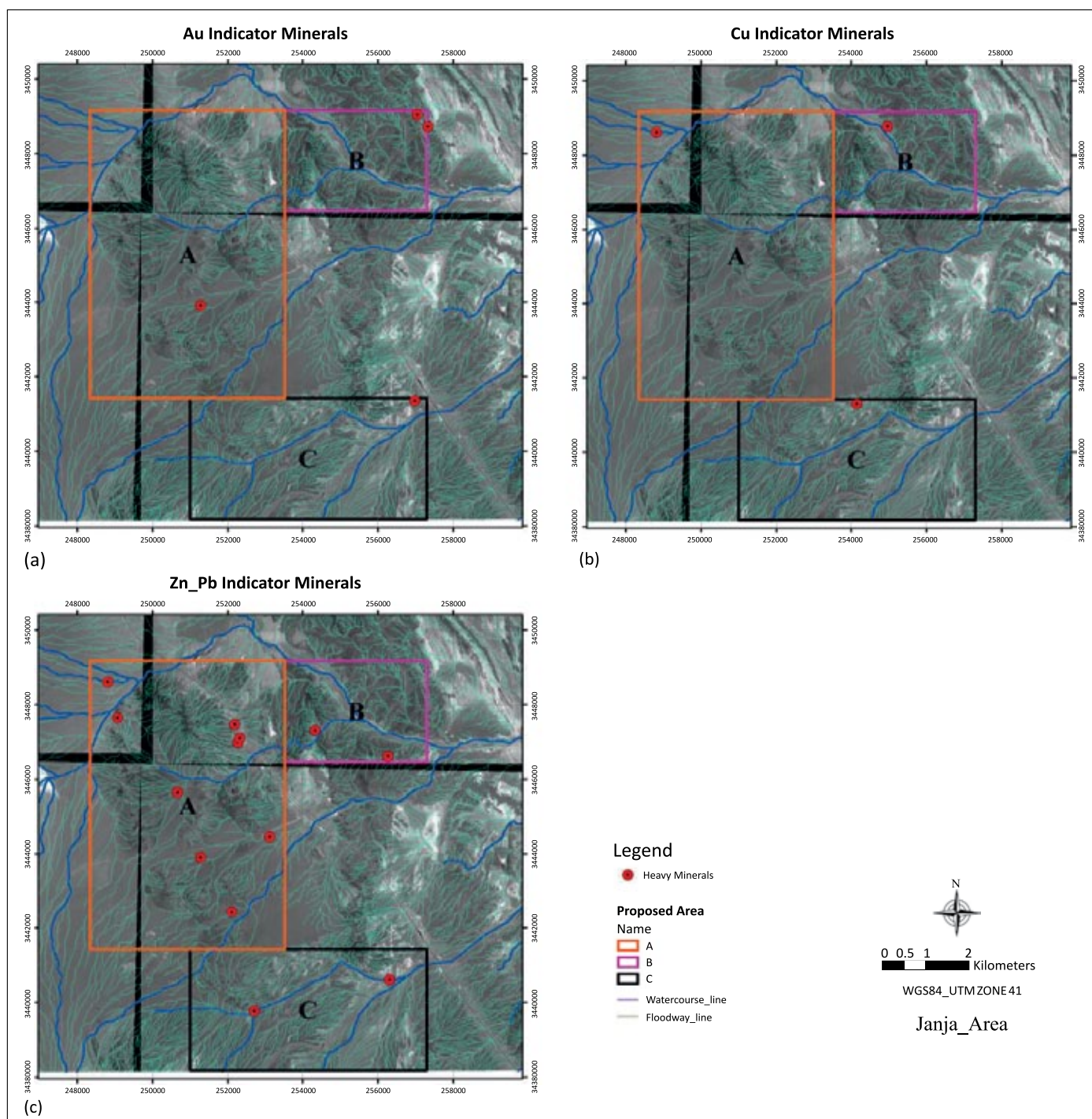


Figure 10. a) Scatter map of gold indicator minerals (gold + cinnabar); b) Scatter map of copper indicator minerals (native copper + malachite); c) Scatter map of lead and zinc indicator minerals (galena, cerussite, mimetite and pyromorphite) in the study area

Rysunek 10. a) Mapa występowania minerałów wskaźnikowych złota (złoto + cynober); b) mapa występowania minerałów wskaźnikowych miedzi (miedź rodzima + malachit); c) mapa występowania minerałów wskaźnikowych ołowiu i cynku (galena, cerusyt, mimetytu i piromorfit) na badanym obszarze

To verify the position of anomalous samples identified through geochemical and heavy mineral investigations, hammer prospecting was conducted in the study area. These investigations were conducted in the northern and southern regions, where the anomalies were concentrated, leading to the identification of polymetallic veins containing gold, silver, copper, lead, and zinc mineralization in both sections. In

the northern region, 13 mineralized samples were collected, and the concentration of polymetallic elements was assessed. Consequently, elemental maps illustrating the distribution of gold, silver, copper, and lead in the northern region are presented (Figure 12). During hammer investigations in the southern part of the specified area, various veins of polymetallic mineralization were identified.

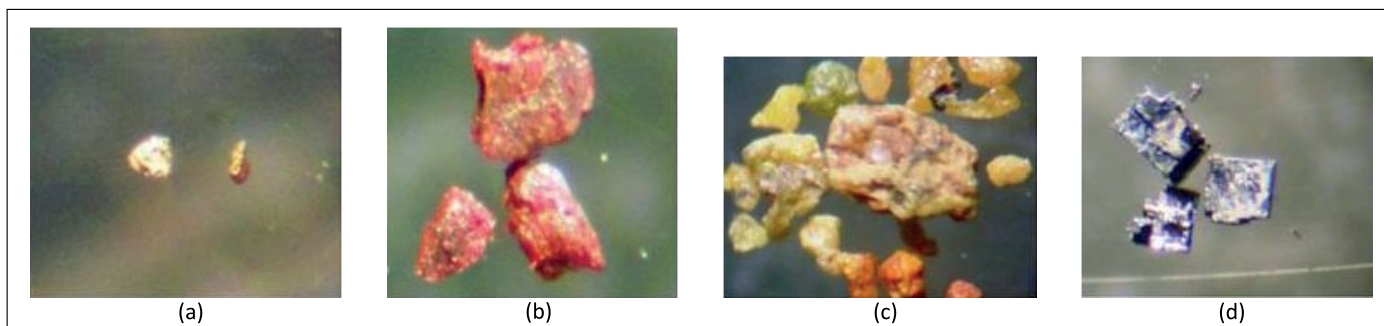


Figure 11. a) Presence of gold; b) Gold along with cinnabar; c) Mimetite, pyromorphite and vanadinite; d) Galena in heavy mineral samples (The gold grains size is about 25 to 100 μ)

Rysunek 11. a) Obecność złota; b) złoto wraz z cynobrem; c) mimetyt, piromorfit i wanadynit; d) Galena w próbkach minerałów ciężkich (wielkość ziaren złota wynosi około 25–100 μ)

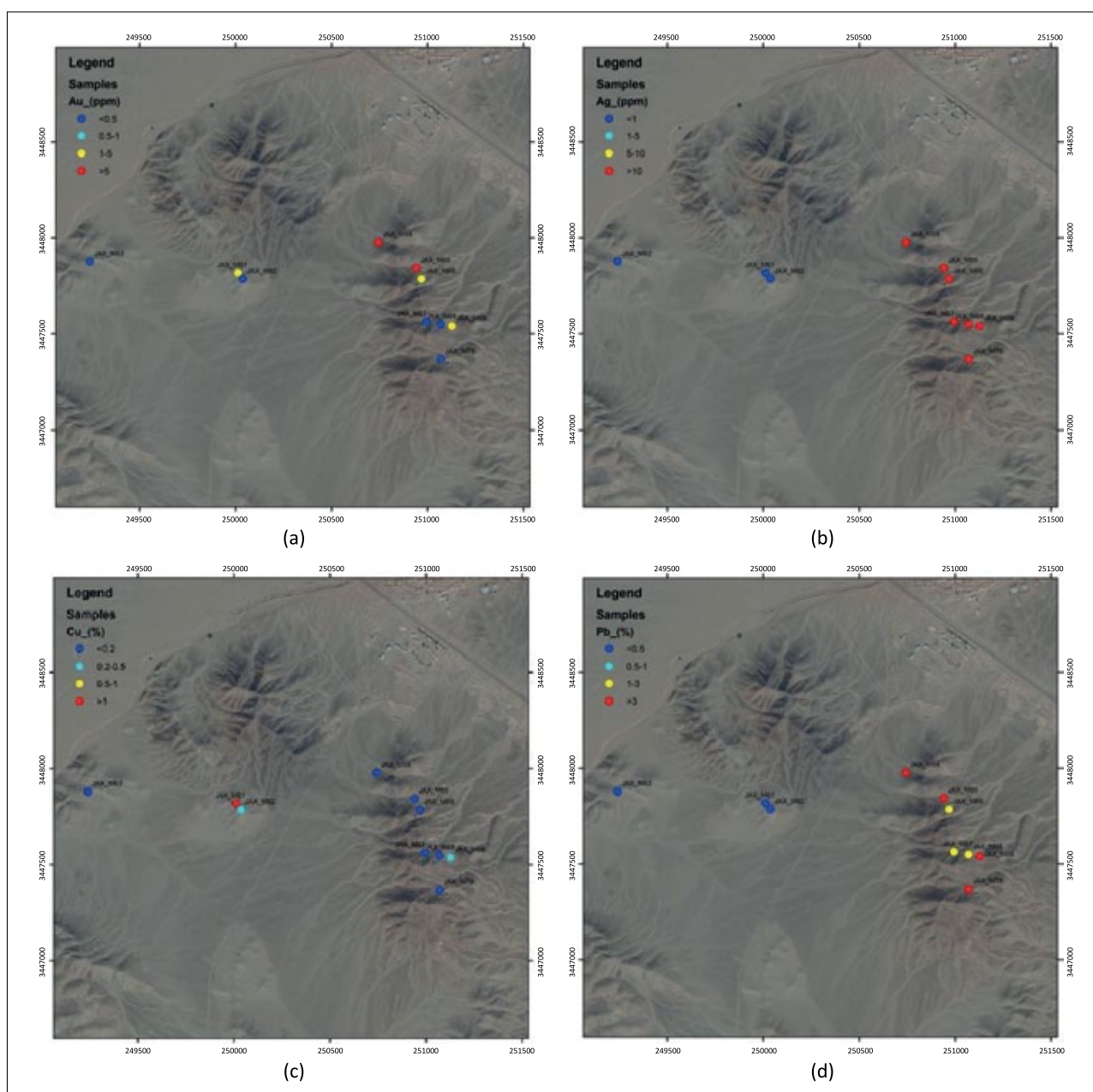


Figure 12. Grade maps:: a) gold; b) silver; c) copper; d) lead for mineralized samples obtained from the northern region

Rysunek 12. Mapy koncentracji: a) złota; b) srebra; c) miedzi; d) ołowiu dla zmineralizowanych próbek uzyskanych z regionu północnego

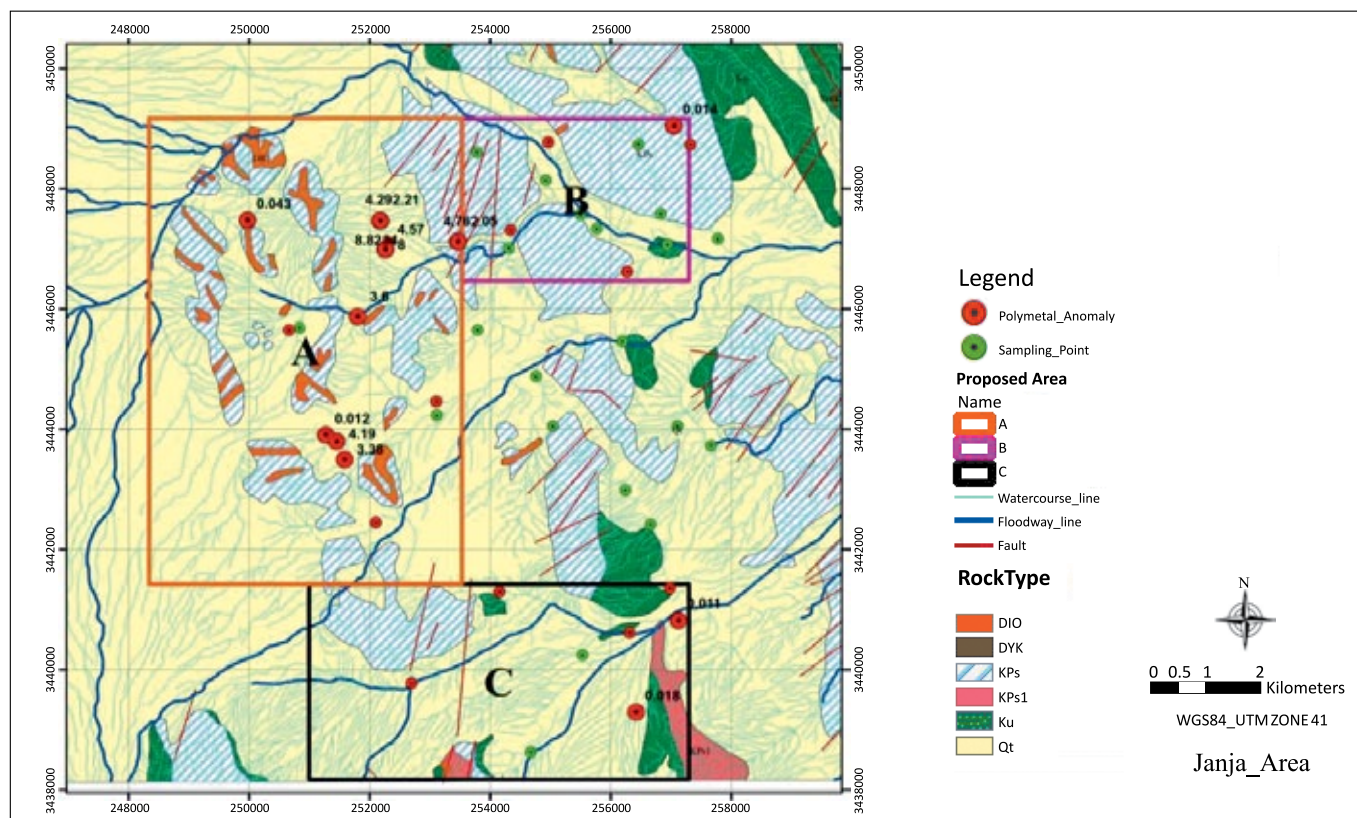


Figure 13. Proposed zones for the continuation of exploration operations

Rysunek 13. Proponowane strefy dla kontynuacji działań poszukiwawczych

Field operations and anomaly control

Following the analysis of geochemical samples using the Rose method, a set of encouraging results was obtained. These anomalous areas may be associated with mineralization events and geochemical halos observed at the surface, or they may be related to mineral contamination or spurious anomalies within the area. It is important to note that the anomalous watershed regions were re-examined during field operations. Samples were collected from mineralized and altered zones and for further tests, including chemical analysis and preparation of thin or polished sections. By analyzing the results of these samples, the presence of economically valuable mineralization in the anomalous area can be confirmed. Based on geological evidence and data obtained from the investigated area, three zones encompassing approximately 71 km² have been identified in the study area as potentially rich in polymetallic deposits of gold, silver, copper, and lead (Figure 13).

Anomalies evaluation using multilayer perceptron network

As previously noted, the multilayer perceptron neural network approach was also employed to identify prospective mineralized zones. The outcomes of this method were presented as distribution maps for gold, silver, copper, lead, and zinc (Figure 14).

Conclusions

Analysis and assessment the geochemical studies and anomaly monitoring indicate that the study area has significant potential for polymetallic mineralization of gold, silver, copper, lead, and zinc. Evidence supporting this mineralization includes the following:

- High concentrations of arsenic, lead, and zinc in the waterway sediment samples from the area;
- High correlation (often exceeding 90%) between arsenic, cadmium, lead, zinc, and copper in river sediment samples;
- Relationship between lead, zinc, copper, and arsenic anomalies in the area;
- The presence of lead, zinc, and copper as an ore-forming group shows a variability of 50.3% in factor analysis;

Results from preliminary heavy mineral analysis and anomaly control in indicate the presence of polymetallic deposits of gold, silver, copper, lead, and zinc in the region. Indicators of this mineralization include:

- High concentrations of arsenic, lead, and zinc found in waterway sediments;
- Compatibility of lead, zinc, copper, and arsenic anomalies in the area;
- Identification of ore minerals such as native gold, native copper, galena, cinnabar, malachite, cerussite, pyromorphite,

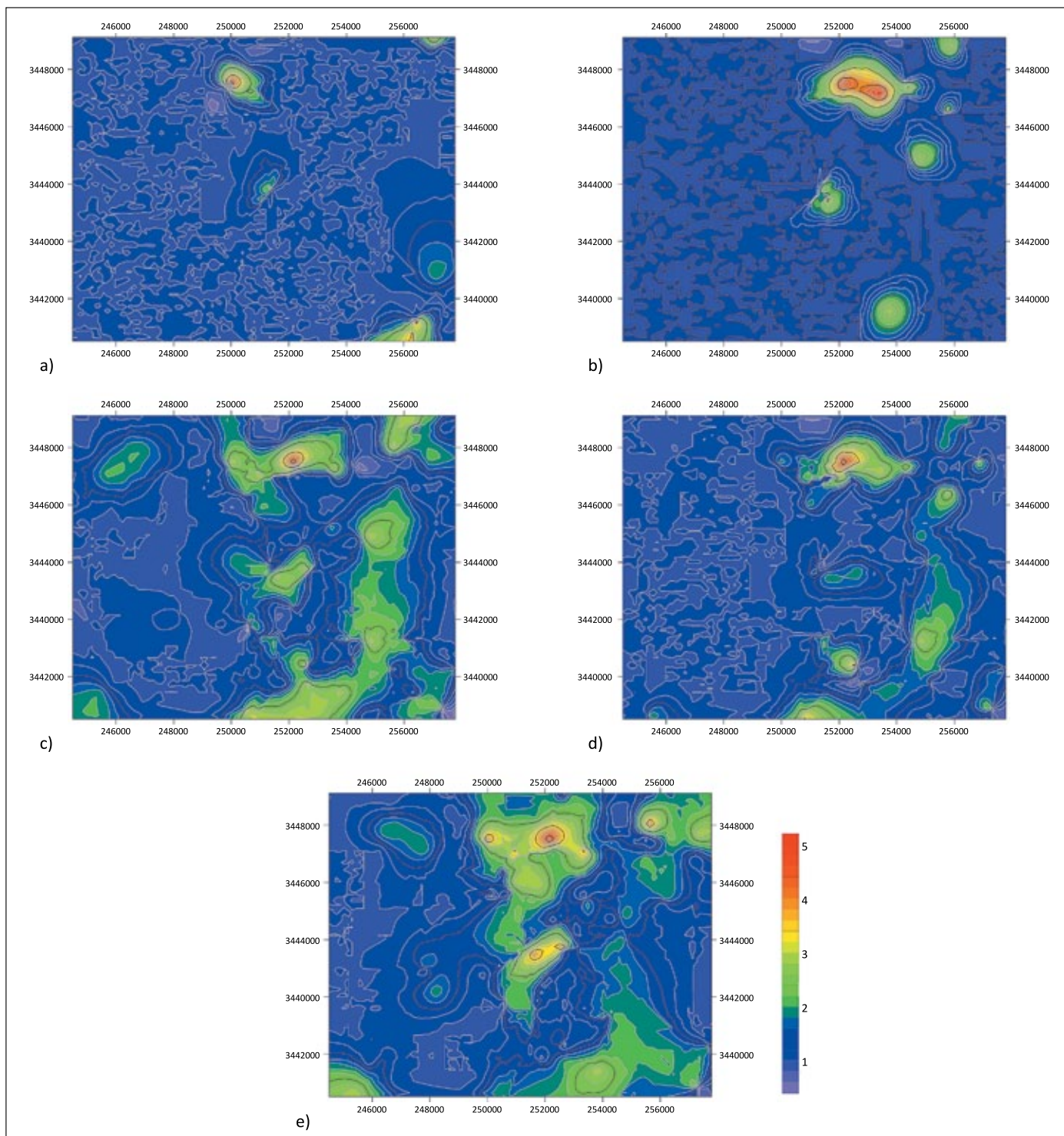


Figure 14. Maps generated and threshold values determined using Multilayer Perceptron for: a) gold; b) silver; c) copper; d) lead; e) zinc
Rysunek 14. Wygenerowane mapy i wartości progowe określone przy użyciu perceptronu wielowarstwowego dla: a) złota; b) srebra; c) miedzi; d) ołowiu; e) cynku

mimetite, vanadinite, and wulfenite in heavy mineral samples.

Heavy mineral anomalies were identified, and field validation of these anomalies was conducted through hammer prospecting. In this study, over 40 highly concentrated mining sites and polymetallic veins containing gold, silver, copper, lead, and zinc were identified in anomalous areas. These poly-

metallic veins commonly occur as siliceous veins ranging in thickness from 1 to 10 meters and contain minerals such as galena, sphalerite, chalcopyrite, and malachite. The maximum concentration of gold, copper, and silver detected in the mineralized samples from these anomalous areas is 15.9, 10775, and 270.3 ppm, respectively. The maximum concentrations of lead and zinc both exceeded 3%. It is believed that mineral

deposits in the anomalous areas were formed by ascending fluids and hydrothermal processes along 100 to 110 degree faults. Based on geological evidence, as well as geochemical and heavy mineral analyses, three zones covering approximately 71 km² within the study area have been identified as potentially rich in polymetallic deposits of gold, silver, copper, and lead.

Artificial neural networks are parallel systems used to identify complex patterns in data. Owing to their parallel structure, neural networks are capable of performing complex computations without relying on mathematical models or assuming the linearity of many variables. One of the primary factors influencing the performance of a neural network is the number of neurons, which significantly impacts the model's output. Therefore, input variables must be selected in a manner that maximizes the number of factors influencing the output. It is important to note that the decision on the number of neurons in the hidden layer depends on experience and testing the network with different numbers of neurons. In addition to the number of neurons, other factors such as the activation function, network structure, initial learning rate, and other factors influence the performance of the neural network model. These factors should also be considered when designing the network to achieve the best network configuration.

This research examined the effectiveness of the multilayer perceptron neural network's intelligent estimator in predicting the grade of gold, silver, copper, lead, and zinc, alongside multivariate statistical techniques. In this approach, the grade of one element served as target variable, while the grades of the remaining four elements were used as input variables. The results confirmed the suitable performance of the model in estimating element grades. The element grade distribution maps produced by both the multivariate statistical analysis methods and the multilayer perceptron neural network indicate that the neural network output maps have identified the promising areas with greater accuracy. This can lead to reduced exploration costs, including those related to sampling, drilling, and geophysical operations.

References

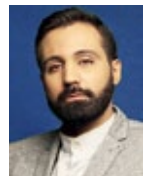
- Agard P., Omrani J., Jolivet L., Whitechurch H., Vrielynck B., Spakman W., Monié P., Meyer B., Wortel R., 2011. Zagros orogeny: A subduction-dominated process. *Geological Magazine*, 148(5–6): 692–725. DOI: 10.1017/S001675681100046X.
- Aghanabati A., 2004. Geology of Iran, Tehran, Iran. *Geological Survey*.
- Arinze I.J., Emedo C.O., Ugbor C.C., 2019. A scalar-geometric approach for the probable estimation of the reserve of some Pb–Zn deposits in Ameri, southeastern Nigeria. *Journal of Sustainable Mining*, 18(4): 208–225. DOI: 10.1016/j.jsm.2019.07.004.
- Barnett R.M., Manchuk J.G., Deutsch C.V., 2014. Projection pursuit multivariate transform. *Mathematical Geosciences*, 46: 337–359. DOI: 10.1007/s11004-013-9497-7.
- Barrie A., Agodzo S., Frazer-Williams R., Awuah E., Bessah E., 2023. A multivariate statistical approach and water quality index for water quality assessment for the Rokel river in Sierra Leone. *Heliyon*, 9(6): e16196. DOI: 10.1016/j.heliyon.2023.e16196.
- Bonham-Carter G.F., Agterberg F.P., Wright D.F., 1989. Integration of geological datasets for gold exploration in Nova Scotia. *Digital Geologic and Geographic Information Systems*, 10: 15–23.
- Borovec Z., 1996. Evaluation of the concentrations of trace elements in stream sediments by factor and cluster analysis and the sequential extraction procedure. *Science of the Total Environment*, 177(1–3): 237–250. DOI: 10.1016/0048-9697(95)04901-0.
- Carman G.D., 2005. Geology, mineralization, and hydrothermal evolution of the Ladolam gold deposit, Lihir Island, Papua New Guinea. [In:] Simmons S.F., Graham I. (ed.), Volcanic, geothermal, and ore-forming fluids: rulers and witnesses of processes within the earth. *Society of Economic Geologists*, 10, 247–284. DOI: 10.5382/SP.10.14.
- Carranza E.J.M., 2004. Usefulness of stream order to detect stream sediment geochemical anomalies. *Geochemistry: Exploration, Environment, Analysis*, 4(4): 341–352. DOI: 10.1144/1467-7873/03-040.
- Carranza E.J.M., 2010. Improved wildcat modelling of mineral prospectivity. *Resource Geology*, 60(2): 129–149. DOI: 10.1111/j.1751-3928.2010.00121.x.
- Carranza E.J.M., Hale M., 1997. A catchment basin approach to the analysis of reconnaissance geochemical-geological data from Albay Province, Philippines. *Journal of Geochemical Exploration*, 60(2): 157–171. DOI: 10.1016/S0375-6742(97)00032-0.
- Chandrajith R., Dissanayake C.B., Tobschall H.J., 2001. Application of multi-element relationships in stream sediments to mineral exploration: a case study of Walawe Ganga Basin, Sri Lanka. *Applied Geochemistry*, 16(3): 339–350. DOI: 10.1016/S0883-2927(00)00038-X.
- Chatterjee S., Bandopadhyay S., 2011. Goodnews Bay Platinum resource estimation using least squares support vector regression with selection of input space dimension and hyperparameters. *Natural Resources Research*, 20: 117–129. DOI: 10.1007/s11053-011-9140-6.
- Chatterjee S., Bandopadhyay S., Rai P., 2008. Genetic algorithm-based neural network learning parameter selection for ore grade evaluation of limestone deposit. *Mining Technology*, 117(4): 178–190. DOI: 10.1179/037178409X405732.
- Chatterjee S., Mastalerz M., Drobniak A., Karacan C.Ö., 2022. Machine learning and data augmentation approach for identification of rare earth element potential in Indiana Coals, USA. *International Journal of Coal Geology*, 259(1–10): 104054. DOI: 10.1016/j.coal.2022.104054.
- Denis D.J., 2021. Applied univariate, bivariate, and multivariate statistics: Understanding statistics for social and natural scientists, with applications in SPSS and R. *John Wiley & Sons*.
- Dieleman S., Willett K.W., Dambre J., 2015. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, 450(2): 1441–1459. DOI: 10.1093/mnras/stv632.
- Dill H.G., 1998. A review of heavy minerals in clastic sediments with case studies from the alluvial-fan through the nearshore-marine environments. *Earth-Science Reviews*, 45(1–2): 103–132. DOI: 10.1016/S0012-8252(98)00030-0.
- Dutta S., Bandopadhyay S., Ganguli R., Misra D., 2010. Machine learning algorithms and their application to ore reserve estimation of sparse and imprecise data. *Journal of Intelligent Learning Systems and Applications*, 2(2): 86–96. DOI: 10.4236/jilsa.2010.22012.

- Ebdali M., Hezarkhani A., 2024. A comparative study of decision tree and support vector machine methods for gold prospectivity mapping. *Mineralia Slovaca*, 56(2): 165–180. DOI: 10.56623/ms.2024.56.2.4.
- Eilu P., Mikucki E.J., Dugdale A.L., 2001. Alteration zoning and primary geochemical dispersion at the Bronzewing lode-gold deposit, Western Australia. *Mineralium Deposita*, 36(1): 13–31. DOI: 10.1007/s001260050283.
- Eyles N., Kocsis S.P., 1989. Sedimentological controls on gold in a late Pleistocene glacial placer deposit, Cariboo Mining District, British Columbia, Canada. *Sedimentary Geology*, 65(1–2): 45–68. DOI: 10.1016/0037-0738(89)90005-5.
- Garrett R.G., Grunsky E.C., 2001. Weighted sums – knowledge based empirical indices for use in exploration geochemistry. *Geochemistry: Exploration, Environment, Analysis*, 1(2): 135–141. DOI: 10.1144/geochem.1.2.135.
- Ghannadpour S.S., Hezarkhani A., 2015. Introducing a software to calculate the grade and ore deposit probability related problems. *Global Journal of Computer Sciences*, 5(1): 01–06.
- Ghannadpour S.S., Hezarkhani A., Sabetmobarhan A., 2015. Some statistical analyses of Cu and Mo variates and geological interpretations for Parkam porphyry copper system, Kerman, Iran. *Arabian Journal of Geosciences*, 8(1): 345–355. DOI: 10.1007/s12517-013-1096-x.
- Govett G.J.S., 2013. Rock geochemistry in mineral exploration. *Elsevier*.
- Grunsky E.C., Drew L.J., Sutphin D.M., 2009. Process recognition in multi-element soil and stream-sediment geochemical data. *Applied Geochemistry*, 24(8): 1602–1616. DOI: 10.1016/j.apgeochem.2009.04.024.
- Halfpenny R., Mazzucchelli R.H., 1999. Regional multi-element drainage geochemistry in the Himalayan Mountains, northern Pakistan. *Journal of Geochemical Exploration*, 67(1–3): 223–233. DOI: 10.1016/S0375-6742(99)00069-2.
- Hinton G., Deng L., Yu D., Dahl G.E., Mohamed A.-r., Jaitly N., Senior A., Vanhoucke V., Nguyen P., Sainath T.N., 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6): 82–97. DOI: 10.1109/MSP.2012.2205597.
- Kaiser H.F., 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23(3): 187–200. DOI: 10.1007/BF02289233.
- Krizhevsky A., Sutskever I., Hinton G.E., 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25(2). DOI: 10.1145/3065386.
- Krose B., Smagt P., 1996. An introduction to neural networks. *The University of Amsterdam*.
- Kumru M., Bakaç M., 2003. R-mode factor analysis applied to the distribution of elements in soils from the Aydın basin, Turkey. *Journal of Geochemical Exploration*, 77(2–3): 81–91. DOI: 10.1016/S0375-6742(02)00271-6.
- Lacassie J.P., Del Solar J.R., Roser B., Hervé F., 2006. Visualization of volcanic rock geochemical data and classification with artificial neural networks. *Mathematical Geology*, 38: 697–710. DOI: 10.1007/s11004-006-9042-z.
- Lacassie J.P., Roser B., Del Solar J.R., Hervé F., 2004. Discovering geochemical patterns using self-organizing neural networks: a new perspective for sedimentary provenance analysis. *Sedimentary Geology*, 165(1–2): 175–191.
- Mahmoudabadi H., Izadi M., Menhaj M.B., 2009. A hybrid method for grade estimation using genetic algorithm and neural networks. *Computational Geosciences*, 13: 91–101. DOI: 10.1007/s10596-008-9107-9.
- Maiti J., 2022. Multivariate statistical modeling in engineering and management. *CRC Press*. DOI: 10.1201/9781003303060.
- Mardia K.V., Kent J.T., Taylor C.C., 2024. Multivariate analysis. *John Wiley & Sons*. DOI: 10.1093/jrssa/qnae103.
- Mostafaei K., Maleki S., Jodeiri Shokri B., Yousefi M., 2023. Predicting gold grade by using support vector machine and neural network to generate an evidence layer for 3D prospectivity analysis. *International Journal of Mining and Geo-Engineering*, 57(4): 435–444. DOI: 10.22059/ijmge.2023.362951.595087.
- Parsons S., Clegg W., 2009. Random and systematic errors. [In:] William C., and others (ed.) *Crystal Structure Analysis: Principles and Practice*, 2nd eds. *International Union of Crystallography Texts on Crystallography*. DOI: 10.1093/acprof:oso/9780199219469.003.0016.
- Pawlowsky-Glahn V., Egozcue J.J., Tolosana-Delgado R., 2015. Modeling and Analysis of Compositional Data. *John Wiley & Sons*. DOI: 10.1002/9781119003144.
- Prendergast K., 2007. Application of litho-geochemistry to gold exploration in the St Ives goldfield, Western Australia. *Geochemistry: Exploration, Environment, Analysis*, 7(2): 99–108. DOI: 10.1144/1467-7873/07-134.
- Reimann C., Filzmoser P., 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statistical treatment of geochemical and environmental data. *Environmental Geology*, 39: 1001–1014. DOI: 10.1007/s002549900081.
- Reimann C., Filzmoser P., Garrett R.G., 2002. Factor analysis applied to regional geochemical data: problems and possibilities. *Applied Geochemistry*, 17(3): 185–206. DOI: 10.1016/S0883-2927(01)00066-X.
- Samanta B., Bandopadhyay S., 2009. Construction of a radial basis function network using an evolutionary algorithm for grade estimation in a placer gold deposit. *Computers & Geosciences*, 35(8): 1592–1602. DOI: 10.1016/j.cageo.2009.01.006.
- Sanford R.F., Pierson C.T., Crovelli R.A., 1993. An objective replacement method for censored geochemical data. *Mathematical Geology*, 25: 59–80. DOI: 10.1007/BF00890676.
- Shirazy A., Ziaii M., Hezarkhani A., 2020. Geochemical Behavior Investigation Based on K-means and Artificial Neural Network Prediction for Copper, in Kivi region, Ardabil province, IRAN. *Journal of Mining Engineering*, 14(45): 96–112. DOI: 10.22034/ijme.2020.37388.
- Simard P.Y., Steinkraus D., Platt J.C., 2003. Best practices for convolutional neural networks applied to visual document analysis. *7th International Conference on Document Analysis and Recognition (ICDAR 2003)*, Edinburgh, Scotland, UK. DOI: 10.1109/ICDAR.2003.1227801.
- Singer D.A., 2006. Typing mineral deposits using their associated rocks, grades and tonnages using a probabilistic neural network. *Mathematical Geology*, 38: 465–474. DOI: 10.1007/s11004-005-9023-7.
- Singer D.A., Kouda R., 1996. Application of a feedforward neural network in the search for Kuroko deposits in the Hokuroku District, Japan. *Mathematical Geology*, 28: 1017–1023. DOI: 10.1007/BF02068587.
- Sun X., Deng J., Gong Q., Wang Q., Yang L., Zhao Z., 2009. Kohonen neural network and factor analysis based approach to geochemical data pattern recognition. *Journal of Geochemical Exploration*, 103(1): 6–16. DOI: 10.1016/j.gexplo.2009.04.002.
- Tahmasbi P., Hezarkhani A., 2011. Presentation a Method for Optimization of Neural Network for Ore Grade Estimation Based on the Porphyry Copper System of Sonajil-Ahar. *Scientific*

- Quarterly Journal of Geosciences*, 21(81): 31–36. DOI: 10.22071/gsj.2011.54200.
- Tripathi V.S., 1979. Factor analysis in geochemical exploration. *Journal of Geochemical Exploration*, 11(3): 263–275. DOI: 10.1016/0375-6742(79)90004-9.
- Tutmez B., 2009. Use of hybrid intelligent computing in mineral resources evaluation. *Applied Soft Computing*, 9(3): 1023–1028. DOI: 10.1016/j.asoc.2009.02.001.
- Valizadeh N., Sharghi Y., 2014. Performance Comparison of Estimators Based on Artificial Intelligence for Ore Grade Estimation in Masjed Daghi Copper Deposit. *Journal of Analytical and Numerical Methods in Mining Engineering*, 4(8): 29–37.
- Van Helvoort P.-J., Filzmoser P., van Gaans P.F., 2005. Sequential factor analysis as a new approach to multivariate analysis of heterogeneous geochemical datasets: an application to a bulk chemical characterization of fluvial deposits (Rhine–Meuse delta, The Netherlands). *Applied Geochemistry*, 20(12): 2233–2251. DOI: 10.1016/j.apgeochem.2005.08.009.
- Venkataraman G., Madhavan B.B., Ratha D., Antony J.P., Goyal R., Banglani S., Roy S.S., 2000. Spatial modeling for base-metal mineral exploration through integration of geological data sets. *Natural Resources Research*, 9(1): 27–42. DOI: 10.1023/A:1010157613023.
- Weller A.F., Corcoran J., Harris A.J., Ware J.A., 2005. The semi-automated classification of sedimentary organic matter in palynological preparations. *Computers & Geosciences*, 31(10): 1213–1223. DOI: 10.1016/j.cageo.2005.03.011.
- Westerhof A.B., 1986. Heavy minerals in exploration the present state of an old art. *ITC Journal*, 4: 290–296.
- Youngson J., Craw D., 1996. Recycling and chemical mobility of alluvial gold in Tertiary and Quaternary sediments, Central and East Otago, New Zealand. *New Zealand Journal of Geology and Geophysics*, 39(4): 493–508. DOI: 10.1080/00288306.1996.9514728.
- Yousefi M., Kamkar-Rouhani A., Carranza E.J.M., 2012. Geochemical mineralization probability index (GMPI): a new approach to generate enhanced stream sediment geochemical evidential map for increasing probability of success in mineral potential mapping. *Journal of Geochemical Exploration*, 115: 24–35. DOI: 10.1016/j.apgeochem.2005.08.009.



Mohammad EBDALI, M.Sc.
Ph.D. Student at the Department of Mining Engineering
Amirkabir University of Technology, Iran
No. 350, Hafez Ave, Valiasr Square, Tehran, Iran
E-mail: mohammadebdali@aut.ac.ir



Adel SHIRAZY, Ph.D.
Postdoctoral Researcher at the Department of Mining Engineering
Amirkabir University of Technology, Iran
No. 350, Hafez Ave, Valiasr Square, Tehran, Iran
E-mail: adel.shirazy@gmail.com



Prof. Ardeshtir HEZARKHANI, Ph.D.
Professor at the Department of Mining Engineering
Amirkabir University of Technology, Iran
No. 350, Hafez Ave, Valiasr Square, Tehran, Iran
E-mail: ardehez@aut.ac.ir



Amin Beiranvand POUR, Ph.D.
Associate Professor at the Institute of Oceanography and Environment (INOS), Higher Institution Center of Excellence (HICoE) in Marine Science, University Malaysia Terengganu (UMT)
Kuala Nerus 21030, Terengganu, Malaysia
E-mail: beiranvand.pour@umt.edu.my