Characterization of rare earth elements by coupling multivariate analysis, factor analysis, and geostatistical simulation; case-study of Gazestan deposit, central Iran

F. Soltani[1], P. Moarefvand[1*], F. Alinia[1] and P. Afzal[2]

1. Department of Mining and Metallurgy Engineering, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
2. Department of Petroleum and Mining Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

Abstract

The traditional approaches of modeling and estimation of highly skewed deposits have led to incorrect evaluations, creating challenges and risks in resource management. The low concentration of the rare earth element (REE) deposits, on one hand, and their strategic importance, on the other, enhances the necessity of multivariate modeling of these deposits. The wide variations of the grades and their relation with different rock units increase the complexities of the modeling of REEs. In this work, the Gazestan Magnetite-Apatite deposit was investigated and modeled using the statistical and geostatistical methods. Light and heavy REEs in apatite minerals are concentrated in the form of fine monazite inclusions. Using 908 assayed samples, 64 elements including light and heavy REEs from drill cores were analyzed. By performing the necessary pre-processing and stepwise factor analysis, and taking into account the threshold of 0.6 in six stages, a mineralization factor including phosphorus with the highest correlation was obtained. Then using a concentration-number fractal analysis on the mineralization factor, REEs were investigated in various rock units such as magnetite-apatite units. Next, using the sequential Gaussian simulation, the distribution of light, heavy, and total REEs and the mineralization factor in various realizations were obtained. Finally, based on the realizations, the analysis of uncertainty in the deposit was performed. All multivariate studies confirm the spatial structure analysis, simulation and analysis of rock units, and relationship of phosphorus with mineralization.

1. Introduction

The strategic importance of rare earth elements (REEs) in the future of mineral economics on one hand, and the relatively low concentrations of this type of deposit, on the other, make it necessary to explore, model, and estimate the concentration of these deposits precisely. The traditional approaches of exploration and modeling along with the inaccurate estimation of the elemental contents create high risks for the future of mining of these deposits. The novel modeling based on the modern methods of using multivariate data and also geostatistical simulations are considered as useful tools in the investigation of spatial structure and uncertainty in terms of concentration. Application of these methods can lead to the management of REEs toward identification of uncertainty factors and, therefore, reduce the modeling risks, and it can provide solutions to the uncertainty management.

The term 'rare earth elements (REEs) is used to describe 17 elements including the lanthanide series, yttrium, and also scandium. In the strategic industries such as defence, renewable energy, communications, healthcare, advance optic, and many high tech sectors, the physical and chemical properties of REEs are very critical in the 21st century. China has been introduced as the main

producer and also consumer of REE elements and products in the last decade [1].

The increasing demand, and consequently, the increasing price of REEs at the global scale has encouraged investment in the exploration of relevant deposits in advanced industrial countries [2]. REEs have diverse applications in modern technology, and provide many materials necessary for the industry [3].

According to their chemical characteristics, REEs have been classified into light REEs (LREEs) including Ce, Eu, La, Nd, Pr, Sm, and Pm, and heavy REEs (HREEs) including Tm, Dy, Er, Gd, Ho, Lu, Tb, Yb, Sc, and Y [4-6]. LREEs are not only more common in nature but also less valuable than HREEs. Strong correlations have been reported between REEs of similar atomic sizes [4]. Although REEs are chemically similar, their end uses in technology and industry are very diverse and specific.

The general methods regarding modelling REE deposits utilize estimation as a moving average to get a smoothed REE grades, by which the resolution of estimated blocks where higher grades of REEs are located will be destroyed and so the uncertainty corresponding to the real REE spatial variability, and therefore, the related volumes of the grades will be misrepresented. As mentioned earlier, the failure in modeling geological uncertainties may cause the feasibility study results to deviate from mining business expectations [4, 5-7].

Conventional estimates can capture the global trends but will be locally different from the reality and so will cause an overly-smoothed picture from the geological zones and the related volumes [8].

Goodfellow *et al*. (2012) [9] have shown the side-effects of the conventional estimation methods such as ordinary kriging (OK) on the grades, corresponding volumes, and metal content when modelling a mineral deposit. The stochastic simulation based on geostatistics is capable of avoiding the conventional estimations' limitations and capture the grade variability and quantify uncertainties for all attributes of the mineral deposit [4, 10-12].

The recent studies on REEs show a move towards mathematical and statistical modeling. Sadeghi *et al*. (2013) [13] have conducted multivariate and geostatistical studies, aiming at investigating the distribution of REEs and associated mineralization and the distinction between HREEs and LREEs in Sweden. Studies of Petrosino *et al*. (2013) [14] have been performed on REEs of

stream sediment samples in various geological environments in Sweden and Italy, and the pattern recognition has been performed by the principal component analysis methodology. Using indicator kriging (IK) and OK, Hellman and Dunkam (2014) [15], have modeled the host rock of REEs, estimating the REE content based on the cut-off grades. In a carbonatite deposit containing REEs, Mikhailov *et al*. (2016) [16], have modeled the rock units associated with mineralization, investigating the relationship between increasing the concentration of different elements in rock unit controlers. Using the clustering and Bayesian methods, Zaremotlagh and Hezarkhani (2016) [17] have modeled the various concentrations of REEs and the geochemical patterns of concentrations in some parts of the Choghart iron ore. Quigley *et al*. (2017) [4] have applied a geostatistical simulation on the total REE to generate the deposits' spatial variability and then have coupled the results with stochastic mine planning to offer an optimum mine design.

The present work provides a combination of multivariate statistical methods and simulation, aiming at modeling the light and heavy REEs in the iron-apatite metasomatism of Gazestan. Therefore, a brief literature review is presented in the following. As a multivariate analysis method, the factor analysis was widely used to interpret exploratory data [18]. Yousefi *et al*. (2014) [19] have used staged factor analysis to eliminate the effects of noise elements in the covariance matrix and increase the effects of indicator and trace elements.

The fractal modeling methods are used for exploratory studies, and specifically, for distinction between the various geochemical anomalies and rock units [20-23]. Using a combination of geostatistical simulation and fractal modeling, Soltani *et al*. (2014) [24] have established a relationship between alterations of a porphyry copper system with high grade variations. Asghari *et al*. (2009) [25] have conducted a research work on the comparison between sequential Gaussian simulation geostatistical estimation in the Choghart iron deposit. Using a combination of simulation and estimation, Talebi *et al*. (2015) [26] have modified the estimated results of the Sungun porphyry copper deposit. Hajsadeghi *et al*. (2016) [27] have used the indicator simulation and multivariate statistics to model a massive-sulfide system.

Metasomatic iron ores can be one of the main economic producers of REEs; thus exploration

and modeling of these deposits have been among the exploration priorities in the recent years. Gazsestan, the case study of this research work, is one of the apt and somehow known areas of Iran in terms of REEs, which is located in the central Iran zone, Bafgh, Posht-e-Badam sub-zone, and on the Esfordi 1:100,000 sheet. The region is characterized by magmatic and metasomatic belts, resulting in iron ores, manganese, apatite, magnetite-apatite, REEs, uranium, and thorium in central Iran [28]. Thus the zone is important in terms of economic geology of REEs in Iran.

A two-step geostatistical simulation approach is used to model both the grade uncertainty and volumetric. Based on 64 elemental analyses, firstly, the elements related to the mineralization of REEs were identified and simulated in deposits during several screening stages using the multivariate statistical method of staged factor analysis. On the derived factor, a concentration-number multi-fractal modeling was then applied, and based on the confusion matrix, the correlation of concentration zones was investigated in mineralization controlling rock units. Then the total contents of REEs, light and heavy, were simulated by sequential Gaussian simulation. Finally, according to the simulation results, the probability of exceeding the variables from the threshold resulting from the fractal modeling was investigated. The main benefit of applying the procedure is to complete reproducing REE distribution statistically and geostatistically in the deposit.

2. Methodology
In this work, different statistical and mathematical methods are used, and these methods are introduced theoretically in the following:

2.1. Staged factor analysis
A significant part of variability is described by a limited number of new variables. The new variables, which are the linear combination of correlated variables, do not show any correlation between them [29]. Generally, there are two steps to this method: first, elimination of the elements that are not included in any of the factors (factor loads are less than the threshold), and repeating until all irrelevant elements are eliminated and all factors are free from the noise elements, and then finding the factors associated with the target such as the type or types of target mineralization [19, 30].

In multivariate statistical methods, the relationships between multiple variables are examined simultaneously. Multivariate methods are often used to reduce the dimensionality so that the process and variability of the dataset can be better interpreted using the reduced data. Principal component analysis is an orthogonal linear transformation that transmits the data to the new coordinate system. The number of calculated principal components depends on the number of initial correlated variables. Staged factor analysis is an improved version of the factor analysis [31]. The first component among the principal components, which is shown by y_1, is written as a linear combination of the initial variables x_1 to x_P:

$$y_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1P}x_P$$

The above equation can be represented as follows in the form of a matrix:

$$Y_1 = [a_1]^T [X]$$

and the most variance and variability along this axis. Since the special values are the same as the diffraction of the principal components, for calculation of the first variable, the first component or the y_1 variable is calculated as follows:

$$S^2_{y_1} = [a_1]^T [S][a_1]$$

where $[S]$ is the covariance matrix is the principal variables.

First, a general factor analysis is performed on the primary data, and taking into account the threshold, if there is an element that does not participate in any of the factors, i.e. an element that is below the desired threshold, that element is deleted and the factor analysis continues with the dataset of the remaining elements until all noise elements are eliminated. The resulting factors are called clean factors. Then the factors that involve the desired mineralization elements are identified and selected, and the operations of the first phase are implemented on the elements of those factors, and the scores of the factor obtained in the last phase are used in the exploratory operations [31, 32].

2.2. Concentration-number (C-N) fractal model
This branch of geometry is able to investigate the variability of complex variables. The power distribution of the fractal model represents the dimensionality, which can be a decimal number, and shows the complexity and variability of the variable. The fractal analyses are tools for the distinction between various communities. The logarithmic graphs in the fractal modeling are tools for distinction among the geological and

mineralization communities in geochemical information. In drawing the concentration-number $(C - N)$ logarithmic plots, with a sudden break in the gradient of the graph, the geological and mineralization community undergoes transformation. This method is basically dependent on the reverse relationship between the concentration and the cumulative frequency of each concentration and the higher concentrations. This method is based on the following equation:

$$N (\geq C)\infty\rho^{-\beta}$$

where $N(\geq C)$ is the number of samples with the concentration equal to or above C, ρ is the concentration, and β is the fractal dimension [18].

2.3. Geostatistical simulation

Geostatistical simulation is a technique for producing the data compatible with a regional variable. The main characteristic of the data obtained from simulation is that it can regenerate the statistics (histogram) and spatial variability (variogram) of the conditioning data [33]. The most important characteristic of the geostatistical simulation is that instead of the best estimate, it can produce a set of models (images) that include a range of possible realizations [34]. In contrast with kriging, which is a smoothed map of the regional variable and in which the variance of the estimated values is reduced, the simulation tries to reproduce the variability of the regional variable in all scales. Estimation on sampling locations gives proper results, and the smoothing effect can be higher as the distance with the sampling location increases; however, the distribution is reproduced in the simulation [35].

The method that is mostly adopted today for modeling applications is called sequential Gaussian simulation. What is known as the primary principal in all Gaussian methods is the normalized primary data, i.e. all data should be transformed into the standard Gaussian distribution $N(0,1)$ [36]. Sequential Gaussian simulation is a direct simulation method that is used for continuous variables such as concentration of minerals. Using the same neighborhood, this approach could be a precise simulation method; however, since it is not possible in practice (it is impossible to form and solve a matrix with a large number of points), this method is considered to be among the

approximate classification algorithms. Sequential Gaussian simulation is used to construct models that reproduce the histogram and also spatial continuity based on simple kriging (SK). SK is the only estimation method that yields exactly the estimated variance and mean [36].

$Let Z = (Z_1, Z_2,...Z_N)'$ be a random variable and $Z = (Z_1, Z_2,...Z_M)'$ a known vector with the $(z_1, z_2,...z_M)'$ values, where $0 \leq M < N$. The distribution of the vector Z under the condition of $Z_i = z_i : i = 1, 2,..., M$ can be re-written as:

$$\Pr\begin{Bmatrix} z_{M+1} \leq Z_{M+1} < z_{M+1} + d_{z_{M+1}},...,\\ z_N \leq Z_N < z_N + d_{z_N} \mid z_1,...,z_M \end{Bmatrix}$$

$$= \Pr\{z_M \leq Z_{M+1} < z_{M+1} + d_{Z_{M+1}} \mid z_1,...,z_M\}$$

$$\times \Pr\{z_{M+2} \leq Z_{M+2} < z_{M+2} + d_{Z_{M+2}} \mid z_1,...,z_M, z_{M+1}\}...$$

$$\times \Pr\{z_N \leq Z_N < z_N + d_{Z_N} \mid z_1,...,z_M, z_{M+1},...,z_{N-1}\}$$

Therefore, the components of Z_i could be simulated in a sequence and with a random selection of the conditional distribution of $\Pr\{Z_i < z_i \mid z_1,...,z_{i-1}\} : i = M+1,..., N.$ The value of Z_i obtained at every stage will be the next point of a conditional data.

What is obtained from a simulation is a realization among an infinite number of possible simulations in the environment. If the random field (RF) of simulation is small $(n < 1000)$, then the above equation will be completely solvable up to the last point, and the method will be accurate (the reproduction of the parameters of the model will be performed completely). If the field is big, this distinction is not made completely in the above equation and is performed with a limited number of previously simulated data. Using a limited number of data for simulation is known as local neighborhood screen [37].

Following, a flowchart describes the procedure of this research work step by step. As shown in Figure 1, at first, a multivariate data reduction based on staged factor analysis is performed, and then a multi-fractal approach is utilized, and finally, a conditional simulation is applied on the REE data.
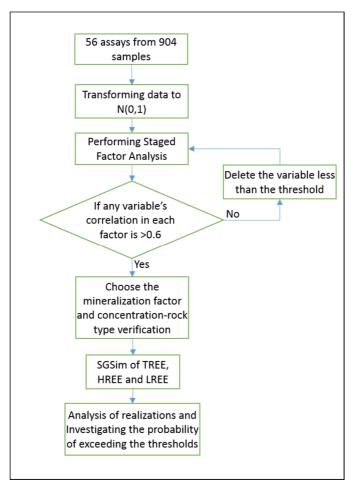
Figure 1. Flowchart of the procedure from data pre-processing to simulation post-processing.

3. Geological settings of Gazestan Deposit

The Gazestan deposit is located in the Yazd Province 78 km from Bafgh and 10 km from the east of the Gazestan village. It is located between the longitudes of $55°55'20''$ and $55°59'00''$ and latitudes of $31°39'00''$ and $31°40'25''$. In terms of geological classifications, the Gazestan exploratory area is placed in the zone of central Iran, sub-zone of Bafgh, Posht-e-Badam, and on the Esfordi $1:100,000$ sheet [38].

So far, several prospecting and exploratory studies have been conducted in central Iran with the aim of identifying REEs, the most important among which are the study in the Bafgh, Posht-e-Badam area, and the study on the prospecting project for metal trace elements by Geological Survey of Iran (GSI) (2001), which lead to the introduction of iron, iron-apatite, and phosphate deposits in the central Iran as the best and most important regions for REE mineralization. The Gazestan low grade iron and apatite deposit is one of the best known deposits in terms of REEs. Geological studies on this deposit included a general exploration in the surface and deep drilling operations with 19 boreholes, from which 427 samples were taken by GSI in 2005, drilling of 15 new exploratory boreholes from which 540 samples were taken by the Madankav Co. in 2015, and modeling the deposits and estimation of the reserves in a separate manner for the iron mineralization, phosphate, and REEs by the Kusha Madan Consulting Engineers Co. in 2015.

The rock units of Gazestan belong to the Rizu series, which contain carbonate rocks, shale, tuff, sandstone, and volcanic rocks. In addition to the sedimentary and volcanic rocks, the plutonic intrusions in the form of stocks and dikes have outcropped in the area with an intermediate to basic rocks. The composition of dikes usually consists of diorite-gabbro and diabase. Green or metasomatic rocks, with intermediate acidic compositions, which are green due to the alteration, are hosted by the iron and phosphate mineralization. The metasomatic processes occur simultaneously or a bit after the mineralization. The intensity of the processes increases when we get closer to the mineralization zone (Figure 2). The rock units of Gazestan are usually composed of acidic and micro-granular volcanic rocks. Thus

the origin of the formation of the acidic rocks of the area is the granitic magma [38].

The Gezestan deposit is in the form of small veins, veinlets, and lenses of apatite or fills the interface between the magnetite grains in a green chlorite-epidote rock unit. The iron ore is initially in the form of magnetite but it turns into hematite due to the martitization. The green altered unit including intrusive volcanics, andesite, microdiorite, tuff, and mafic-ultramafic rocks, and it seems that the vein and lenses of magnetite and apatite have been concentrated in it as immiscible phases. The apatite ore is in the form of fluorine apatite, and REEs in the apatite are in the form of fine inclusions of monazite, xenotime, and allanite, in which a high ratio of LREE/HREE has

been reported with LREE enrichment and Eu depletion. The mineralization area in Gazestan is 2.2 km in length and > 0.7 km in width [38]. According to the studies by GSI, the dip of the mineralization controller rock units is around 50° toward North.

Faulting and tectonic displacement in the area are relatively high and play an important role in the depositing processes. The dominant process of faulting structures occurs along the Northwest-Southeast strike. Investigation of joints shows that the stress in the main direction (East-West) causes metasomatism and alteration of the magnetite mass, and displaces the magnetite mass along the second stress axis (Northeast-Southwest) [28].
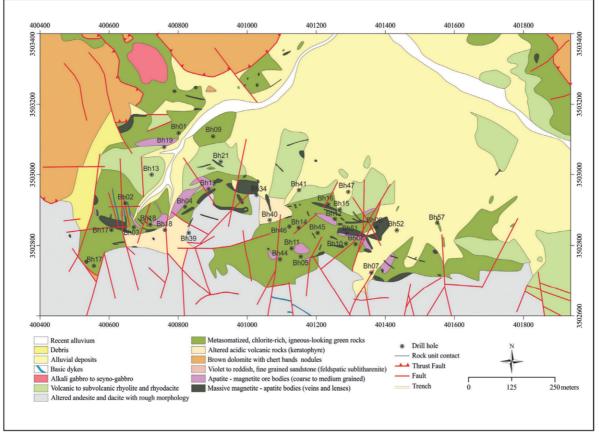

Figure 2. 1:1000 scale geology map of the Gezestan deposit [38].

4. Dataset

In this work, 908 core samples were obtained from 15 exploratory boreholes of 4218 m lengths. The locations of boreholes and samples are shown in Figure 3 in a 3D space. 64 elements, especially the light and heavy REEs and the total REEs were analyzed by the ICP method. The analyzed LREEs included Ce, Eu, La, Nd, Pr, and Sm, and analyzed HREEs included Tm, Dy, Er, Gd, Ho,

Sc, Yb, Tb, and Lu. In order to identify the distribution of REEs, the statistics were separately calculated for each rock unit (Table 1) so that the target rock units could be specified for further modeling.

Table 1 shows the statistical parameters of the 2 m composited data. As it could be seen, the magnetite units not only had a higher average concentration compared to the other units but also

had a high LREE/HREE ratio. The ratio of the light to heavy REEs in magnetite units was 3.8, and in diorite units with very small number of samples $(n = 7)$ it was 3.9. The proportion of the heavy elements was sensibly higher in dacite compared to the light elements. The ratio of the light to heavy elements was 3.1 in dacite, 3 in andesite, 3.2 in tuff, 2.1 in dacite, and 1.2 in gangue. In other words, this ratio significantly increased in each rock unit with increase in the total average concentration.
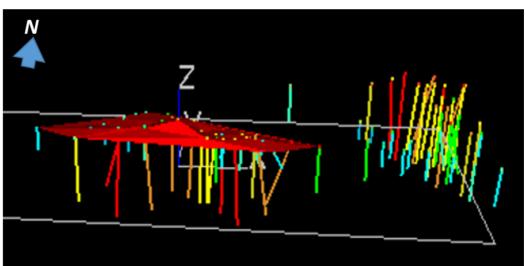

Figure 3. Western drilled boreholes of the Gazestan deposit utilized for further simulation. View to the North.

Table 1. Statistical parameters of light, heavy, and total REEs in each rock unit along with the ratio of light to heavy elements.

Rock type	Total count	Mean of total REE (ppm)	Mean of HREE (ppm)	Mean of LREE (ppm)	LREE/HREE ratio
Andesite	218	827	208	619	3.0
Dacite	64	342	109	233	2.1
Diorite	7	1373	280	1093	3.9
Magnetite	244	1363	283	1080	3.8
Rhyodacite	17	569	135	434	3.2
Tuff	346	941	222	720	3.2
Waste	12	482	215	267	1.2

5. Results

5.1. Staged factor analysis

Using the staged factor analysis, the main components related to REE mineralization were determined and extracted. First, all data was transformed to Guassian distribution N(0,1). In the first stage of the factor analysis, 13 factors were obtained, which explained 81.1% of the total variance of the data. The first factor accounted for 34.24% of the total variability with an eigenvalue value of 19.17. At this stage, taking into account the threshold of 0.6 for the correlation values with 13 factors, the noise elements of As, Ba, Bi, Co, Ga, Ni, Sr, W, and Zn were eliminated at this stage as noise or uncorrelated elements. This process was continued until the 5th stage to obtain the noise-free (clean) factors. Then the factors that contained trace elements and the intended mineralization were selected. Considering the geological and mineralogical conditions, factors 1, 2, and 3 were taken as the objectives because they contained trace elements that included phosphorus and REEs. Thereafter, Ag, Ca, Cd, Cs, Li, Mg, Mn, Nb, and Ti with a correlation below the threshold were eliminated from the factors, and finally, three factors were obtained in the 6th stage. All elements had a correlation coefficient higher than the threshold. The elements are presented in Table 2.

Table 2. Results of factor analysis of the 6th or final stage. As observed, all elements remaining in the three factors have a correlation more than 0.6 with the corresponding factor.

Rotated Component Matrix[a]

	Component 1	2	3
AL	-.451	-.706	.077
CE	.920	.310	-.005
DY	.938	.244	.084
ER	.938	.187	.069
EU	.926	.126	.030
FE	.143	.864	-.031
GD	.937	.314	.018
GERMANIUM	.138	.102	.889
HREE	.922	.206	.221
HF	-.031	.820	-.027
HO	.939	.213	.049
K	-.340	.738	-.066
LREE	.926	.310	.001
LA	.910	.310	.008

Rotated Component Matrix[a]

	Component 1	2	3
LU	-.079	-.066	-.973
ND	.924	.324	-.004
P	.765	.435	.000
PR	.926	.292	-.030
RB	-.336	-.715	.039
SC	-.069	-.526	.652
SM	.938	.305	.050
SN	.197	.717	.129
TB	.944	.266	.049
TH	.726	-.248	.097
TM	-.079	-.066	-.973
TOTALREE	.944	.291	.052
U	.705	-.361	.154
V	.173	.886	.023
YB	.921	.138	.069

5.2. Modeling of $(C - N)$ fractal analysis

Among the factors obtained from the staged factor analysis, the first factor that was directly related to mineralization was selected for fractal modeling and relation with rock units. On the first factor, the $(C - N)$ fractal diagram was plotted by taking the logarithm of the values and taking into account the cumulative order of the factor (Figure 4). The fractal diagram shown in Figure 4 reveals the breaking point between the fitted lines in the diagram, and accordingly, the threshold values for the separation of the rock units. These breaking points correspond to the values of $2.77, 1.42, 0.60$, and 0.163 of the factor's normal distribution, respectively. Considering that at the two final breaking points a multi-fractal behavior is observed, it demonstrates a highly concentrated rock unit. As in the Gazestan deposit, the mineralization is a mixture of magnetite-apatite; thus the highly concentrated part could be compatible with the magnetite-apatite rock units.

Then using the confusion matrix, the correlation of the zone obtained from the number-N fractal method and the geological units was investigated. Based on the thresholds obtained from the fractal modeling and the correlation of the concentration intervals with the various rock units, the highest overall accuracy was obtained for the high-grade magnetite-apatite rock units. The results obtained are presented in Table 3.
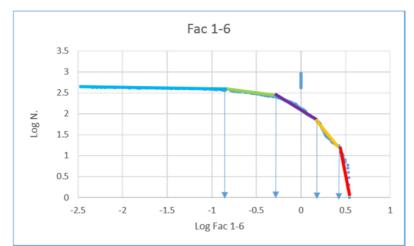

Figure 4. C-N logarithmic diagram of scores of the first factor from the 6th stage of the staged factor analysis and the corresponding break points.

Table 3. Overall accuracy, and errors type 1 and type 2 calculated for correlation of the concentration model of the factor with the geological model.

Geological model			
Outside the zone (magnetite-apatite)	Within the zone (magnetite-apatite)		
False positive (B) = 51	True positive (A) = 26	Within the zone 1.42037-3.47384	**Fractal model**
True negative (D) = 612	False negative (C) = 218	Outside the zone 1.42037-3.47384	
Error type 2 = 0.07	Error type 1 = 0.893		
Overall accuracy = 70%			

5.3. Simulation of mineralization factor and REEs

Since the sequential Gaussian simulation was used for simulation, the data was transformed to Gaussian distribution N(0,1). Figures 5a to 5f show the data for REEs before and after normalization. The score values (Figure 5h) are factors with Gaussian distribution because they are the product of the rotation of the normalized data vectors. Figure 5g shows the linear correlation between the total REE and the mineralization factor, both in a normal space.
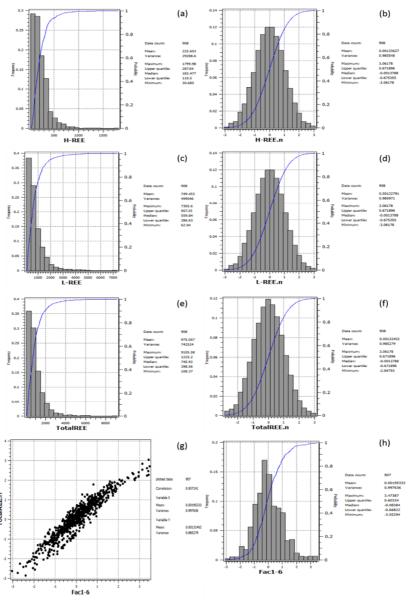

Figure 5. Histograms of a. HREE, c. LREE, and e. total REE on the left and the corresponding normalized values on the right. Scatter plot between the total normalized concentration values and the mineralization factor (g), histogram of the mineralization factor (h).

As it can be seen, there is a high linear correlation between the variables that show a correlation between the mineralizing elements, specifically phosphorus, and REEs.

In order to investigate the spatial variability and probable anisotropy of the deposit, the experimental variogram and fitted model were plotted and investigated. Considering the insufficient number of boreholes that were drilled, specifically perpendicular to the continuity of the deposit (i.e. in the north-south strike), identification of the varigram ranges were uncertain.

However, according to the geological continuity along the controller rock units (a dip of 50° toward north), modeling was performed for the second axis of anisotropy. The directional and non-directional variograms of the total REE are provided in Figure 6. Other variables have approximately the same spatial structures.

From the vertical variogram, the nugget effect value was taken to be 0.3 out of 1 for the variables, which was due to the relatively high variations of variables in very small scale structures. Using the Gaussian data for REEs and the factor scores and parameters obtained from the variography, the concentration simulation was performed 10 times in the SGEMS software (version 2.5b) in the defined block model. The block model is based on the borehole data and geological map, and is defined according to the continuity of the deposit with the dimensions of $20 \times 20 \times 10$ m and with 31421 parent blocks. For each variable, conditional simulation was performed 10 times based on SK. The parameters used in the simulation are given in Table 4. The number of conditioning data plays an important role in the accuracy of the model, and after iterations, 100 conditioning data was used for the final model.
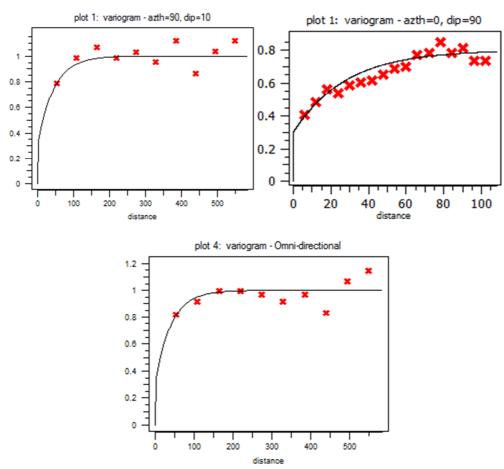

Figure 6. Directional variograms along the major axis (azimuth 90) and vertical veriogram, and omni-directional variogram of total REEs.

Table 4. Parameters used to simulate normal variables.

Variables	Nb of realizations (%)	Kriging type	Seed Nb	Azimuth	Dip	Rake	Nb of conditioning data
Total REE LREE-HREE-Fac	10	SK	14071789	90	10	-40	100

By investigation of the 10 realizations and considering that each realization is simulated from different simulation random paths, different previously simulated nodes, and drawing random values from different estimation variance CDF's, it can be understood that the realizations are completely different from each other but each can reproduce the statistical and geostatistical parameters of the primary data. Considering the data histogram as the reference distribution, each realization was then back-transformed from Gaussian to data distribution. Factor values were simulated with similar simulation parameters such as number of conditioning data, variogram model, and search ellipse. Since factor values are in the Gaussian space (i.e. all vectors are calculated in a normal space through staged factor analysis), no back-transformation was performed.

6. Discussion
6.1. Analysis of realizations
Comparison of the realization statistics with the data indicates the proper reproduction of the minimum, maximum, mean, variance, and also the range of classical statistics for each one of the variables. Figure 7 shows the histograms of one of the realizations for each variable (HREE, LREE, total REE) along with the statistical parameters. One of the investigation methods for the reproduction of data by the simulation results is to study the QQ plots. According to the QQ-plot diagrams, the quantiles of the two distributions are compared and plotted versus each other. The fact that the values are located on the $y = x$ line shows that the quantiles of the two distributions of the data and the values obtained from the simulation are exactly the same and the simulation not only reproduces the mean and variance of the data but also reproduces the distribution itself. In Figure 7, for each histogram, the corresponding QQ-plots are provided on the right.

One of the results of simulation is the E-Type map. This map results from (arithmetical) averaging of all realizations for each block. Due to the averaging, the E-type map provides a smoothed map of changes and cannot represent the changes in the upper and lower tails of the distribution, and therefore, displays a very low variance in comparison with each one of the realizations. In Table 5, the statistical characteristics of each realization in the normal space are compared with the E-type map. The sharp decrease of the variances of realizations from 1 to 0.1 in the E-Type values indicates the effects of averaging (the same as the smoothing effect caused by kriging).

In Figure 8, the E-type values of factor scores, total, light, and heavy REEs are given. The trend of changes of factor scores and REEs indicate a high correlation between the mineralization factor and REEs. Since the factor scores are in the Gaussian space, all other variables are compared with factor scores in the normal mode.

Figure 9 shows the back-transformed simulations into original distribution for the three variables HREE, LREE, and total REE. It should be noted that the mineralization factor (FAC) is meaningful in Gaussian space and so it does not require to back-transform into the original distribution. The range of each variable has been shown in the corresponding legend in the right column of the figure.

Table 5. Basic statistics of total REE realizations and E-type in the Gaussian space.

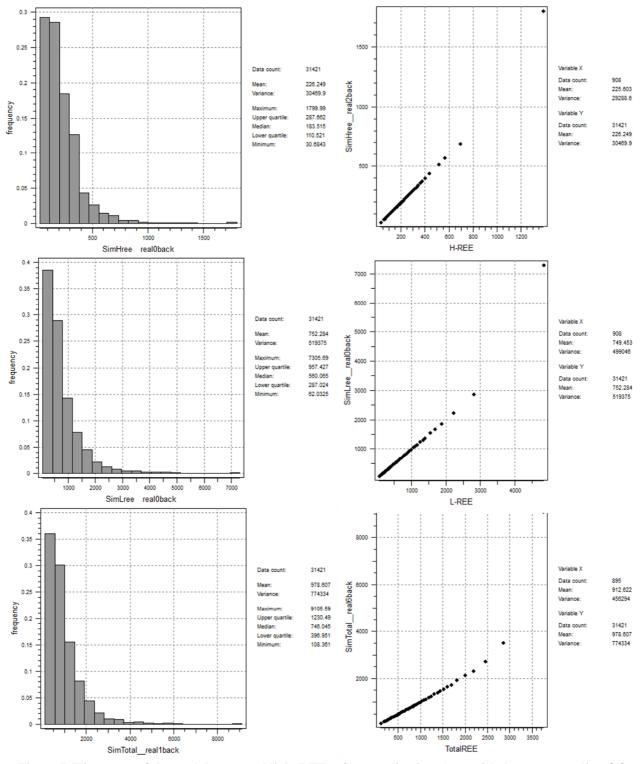|	Range (%)	Min. (%)	Max. (%)	Mean (%)	Var. (%)^2
Sim 1	8.8063	-4.65115	4.15519	-0.04137	1.00497
Sim 2	8.16448	-4.0281	4.13638	0.034255	0.94383
Sim 3	8.41554	-4.33506	4.08048	0.007614	0.94965
Sim 4	7.59244	-3.69934	3.8931	0.037463	0.9617
Sim 5	7.83214	-4.04621	3.78593	-0.088198	0.9853
Sim 6	8.34621	-3.85944	4.48677	-0.093815	1.022
Sim 7	8.50574	-3.7846	4.72114	0.096358	0.9758
Sim 8	7.78277	-3.74586	4.03691	0.021579	1.0175
Sim 9	7.90209	-4.066	3.83609	-0.03183	0.9962
Sim 10	8.53126	-4.06154	4.46972	-0.02246	0.95185
E-type	2.4859	-1.26239	1.22359	-0.00804	0.099

Figure 7. Histogram of the total, heavy, and light REEs of one realization along with the corresponding QQ diagram to be compared with the data distribution used in the simulation.
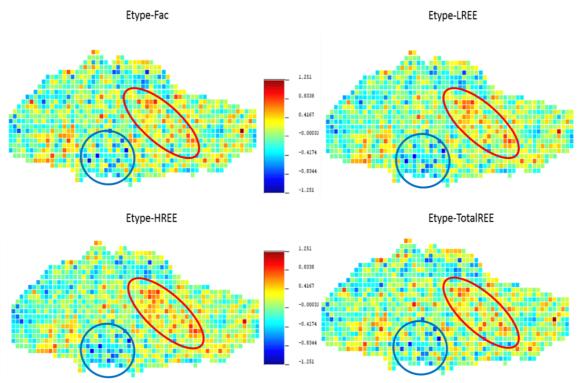
Figure 8. E-type map of factor values, heavy, light, and total REEs in Gaussian space.
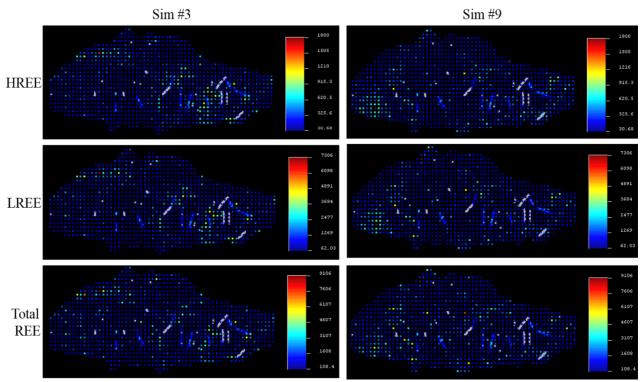

Figure 9. Plans of back-transformed variables into original distribution. Two random realizations (#3 and #9) have been shown for back-transformed HREE, LREE, and total REE.

6.2. Uncertainty analysis

The uncertainty analysis was performed based on two criteria, the "conditional variance" and the "interquartile range". Both CV and IQR are the criteria to quantify analysis based on the statistical methods. CV is calculated the same as the variance, and IQR is Q3 (third quartile)-Q1 (first quartile) of the corresponding distribution. In order to compare the results, all variables were

analyzed in normal distribution (before back-transformation).

As it could be seen in Figure 10, the grade uncertainty does not depend on the closeness to the boreholes as it is in the conventional estimation methods such as kriging. In the high grade zone, according to heteroscedasticity of data, and therefore, the proportional effect, the uncertainty will raise in high grade zones. Thus quantification of grade uncertainty, especially in high-grade blocks, could be a helpful tool for incorporating the uncertainty into the stochastic mine design and short-term/long-term mine planning.
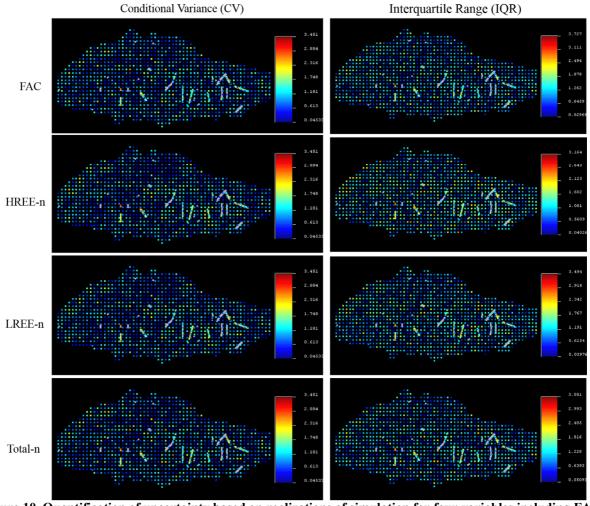

Figure 10. Quantification of uncertainty based on realizations of simulation for four variables including FAC, HREE, LREE and total REE. The right column is based on interquartile range (IQR) and the left one is based on conditional variance (CV).

6.3. Probability of exceeding the thresholds

As stated in the introduction, due to the low-grade REEs in the deposit and high positive skewness of the data, it is necessary to use the geostatistical simulation tools to keep the initial distribution in the simulation results and prevent the smoothing the high grades so that the proportion of the high-concentration parts is prevented. Then in order to analyze the probability that block concentrations exceed the thresholds, the probability maps can be used. These maps determine the probability that block concentration exceeds a threshold and calculate the corresponding probability for each variable based on the count of times the simulated concentration exceeds the threshold. The probability that various REEs exceed different thresholds of 1000 ppm $\left(0.1\%\right)$, 2000 ppm $\left(0.2\%\right)$, and 3000 ppm $\left(0.3\%\right)$ is shown in Figure 11 (left). From the mine design and planning viewpoint, the corresponding grades play an important role that can affect the evaluation of the deposit and totally change the feasibility study of the project. Thus the probability of exceeding different thresholds of factor scores are also shown in Figure 11 (right column).
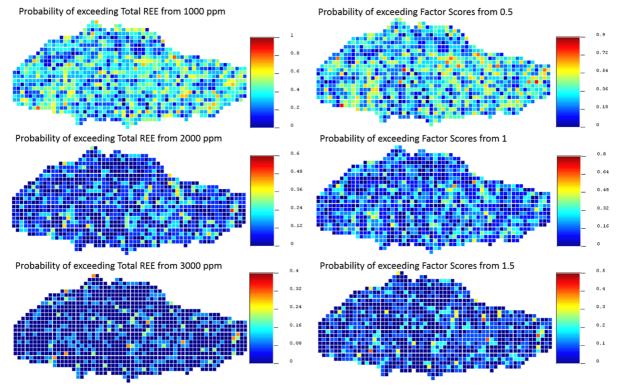
Figure 11. Maps of probability that the total REEs exceed 1000, 2000, and 3000 ppm, and that the corresponding factor scores exceed 0.5, 1, and 1.5.

As it could be seen, with increase in the thresholds, the probability of exceedance decreases. It is worth noting that the difference in the probability of exceedances is based on simulation and estimation. Due to the smoothing effect of the kriging estimation, the exceedance probability will be significantly reduced, and therefore, the economic higher-grade cut-offs are significantly ignored.

7. Conclusions

Exploratory studies carried out on the Gazestan deposit indicate the presence of rare earth elements (REEs) in apatite and monazite minerals but the main issue is evaluating the grade in the whole deposit without any smoothing effect in grade distribution. According to the fractal studies on the variability of the total REEs and its correspondence with different rock units in this deposit, it was determined that the mineralization mainly occurred in areas having magnetite-apatite. On the other hand, based on the staged factor analyses, by elimination of the non-dependent elements during six stages, from among 63 analyzed elements, 18 elements including HREE, LREE, total REE, and also phosphorus remained in the mineralization factor, which again, confirmed the multivariate relationship between REE mineralization and existence of phosphorus minerals. The high correlation between the mineralization factor and phosphorous proves the existence of REEs with magnetite-apatite units.

Due to the importance of reproducing the distribution and tail of high-concentrations of the deposits, conditional simulation was used. Investigating the spatial structure of anisotropy shows the high similarity of the anisotropy structure of light and heavy REEs with the mineralization factor, which, in turn, indicates that there is a high spatial relationship between REEs and the mineralization factor, and also the phosphorus element.

Using the sequential Gaussian simulation, the values of light, heavy, and total REEs were modeled in the 3D space of the ore deposit. The results also show a high correlation between the contents of REEs with the simulated values of the mineralization factor. On the other hand, the probability was investigated, showing that REEs and the mineralization factor exceed various economic thresholds, and due to the reproduction of concentration distribution after simulation, no smoothing or removal of high concentration were observed in the results.

Acknowledgments

authors are grateful to all the personnel of this company.

References

[1]. Rahimi, E., Maghsoudi Gharablagh, A. and Hezarkhani, A. (2016). Geochemical investigation and statistical analysis on rare earth elements in Lakehsiyah deposit, Bafq district, Journal of African Earth Sciences. 124: 139-150.

[2]. Sarparandeh, M. and Hezarkhani, A. (2017). Studying distribution of rare earth elements by classifiers, Se-Chahun iron ore, Central Iran. Acta Geochim. 36 (2): 232-239.

[3]. Humphries, M. (2009). Rare earth elements: The global supply chain, Specialist in energy policy, Congressional research service. USGS. The Global Supply Chain Establish a Stockpile.

[4]. Quigley, M., Dimitrakopoulos, R. and Grammatikopoulos, T. (2017). Risk-resilient mine production schedules with favourable product quality for rare earth element projects. Mining Technology. pp. 1474-9009.

[5]. Dowd, P.A. (1997). Risk in minerals projects: analysis, perception and management. Trans Inst Min Metall Sec A. 106: A9-A18.

[6]. Dimitrakopoulos, R., Farrelly, C.T. and Godoy, M. (2002). Moving forward from traditional optimization: grade uncertainty and risk effects in open-pit design. Mining Technol. 111: 82-88.

[7]. Vallée, M. (2000). Mineral resource + engineering, economic and legal feasibility = ore reserve. CIM Bull. 93: 53-61.

[8]. Journel, A.G. (2007). Roadblocks to the evaluation of ore reserves – the simulation overpass and putting more geology into numerical models of deposits. Spectrum Series 14. 2nd ed. Melbourne: Australasian Institute of Mining and Metallurgy. pp. 29-32.

[9]. Goodfellow, R., Albor Consuegra, F., Dimitrakopoulos, R. and Lloyd, T. (2012). Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada. Comput Geosci. 42: 71-78.

[10]. Journel, A.G. (1979). Geostatistical simulation. 6 methods for exploration and mine planning. Eng Mining J. 180: 6.

[11]. Journel, A.G. and Kyriakidis, P.C. (2004). Evaluation of mineral reserves – a simulation approach. New York (NY): Oxford University Press.

[12]. Rossi, M. and Deutsch, C.V. (2014). Mineral resource estimation. New York (NY): Springer.

[13]. Sadeghi, M., Morris, G.A., Carranza, E.J.M., Ladenberger, A. and Andersson, M. (2013). Rare earth element distribution and mineralization in Sweden: An application of principal component analysis to FOREGS soil geochemistry. Journal of Geochemical Exploration. 133: 160-175.

[14]. Paola Petrosino, P., Sadeghi, M., Albanese, S. and Andersson, M. (2013). REE contents in solid sample media and stream water from different geological contexts: Comparison between Italy and Sweden. Journal of Geochemical Exploration. 133: 176-201.

[15]. Hellman, P.L. and Duncan, R.K. (2014). Evaluation of rare earth element deposits. Applied Earth Science. 123 (2): 107-117.

[16]. Mikhailova, J.A., Kalashnikov, A.O., Sokharev, V.A., Pakhomovsky, Y.A., Konopleva, N.G., Yakovenchuk, V.N., ... and Ivanyuk, G.Y. (2016). 3D mineralogical mapping of the Kovdor phoscorite–carbonatite complex (Russia). Mineralium Deposita. 51 (1): 131-149.

[17]. Zaremotlagh, S. and Hezarkhani, A. (2016). A geochemical modeling to predict the different concentrations of REE and their hidden patterns using several supervised learning methods: Choghart iron deposit, bafq, Iran. Journal of Geochemical Exploration. 165: 35-48.

[18]. Afzal, P., Eskandarnejad Tehrani, M. and Ghaderi, M. (2016). Delineation of supergene enrichment, hypogene and oxidation zones utilizing staged factor analysis and fractal modeling in Takht-e-Gonbad porphyry deposit, SE Iran. Journal of Geochemical Exploration. 161 (2): 119-127.

[19]. Yousefi, M., Kamkar-Rouhani, A. and Carranza, E.J.M. (2014). Application of staged factor analysis and logistic function to create a fuzzy stream sediment geochemical evidence layer for mineral prospectivity mapping. Geochemistry: Exploration, Environmental, Analysis. 14 (1): 45-58.

[20]. Mandelbrot, B.B. (1983). The Fractal Geometry of Nature. W. H. Freeman, San Fransisco. 468 P.

[21]. Cheng, Q., Xu, Y. and Grunsky, E. (2000). Integrated spatial and spectrum method for geochemical anomaly separation. Natural Resources Research. 9 (1): 43-52.

[22]. Li, C., Ma, T. and Shi, J. (2003). Application of a fractal method relating concentrations and distances for separation of geochemical anomalies from background. J. Geochem. Explor. 77: 167-175.

[23]. Afzal, P., Fadakar Alghalandis, Y., Khakzad, A., Moarefvand, P. and Rashidnejad Omran, N. (2011). Delineation of mineralization zones in porphyry Cu deposits by fractal concentration–volume modeling. J. Geochem. Explor. 108: 220-232.

[24]. Soltani, F., Afzal, P. and Asghari, O. (2014). Delineation of alteration zones based on Sequential Gaussian Simulation and concentration–volume fractal modeling in the hypogene zone of Sungun copper deposit, NW Iran. Journal of Geochemical Exploration. 140: 64-76.

[25]. Asghari, O., Soltni, F. and Amnieh, H.B. (2009). The comparison between sequential gaussian simulation (SGS) of Choghart ore deposit and geostatistical estimation through ordinary kriging, Aust J Basic Appl Sci. 3 (1): 330-341.

[26]. Talebi, H., Asghari, O. and Emery, X. (2015). Stochastic rock type modeling in a porphyry copper deposit and its application to copper grade evaluation, Journal of Geochemical Exploration. 157: 162-168.

[27]. Hajsadeghi, S., Asghari, O., Mirmohammadi, M.S. and Meshkani, S.A. (2016). Indirect rock type modeling using geostatistical simulation of independent components in Nohkouhi volcanogenic massive sulfide deposit, Iran. Journal of Geochemical Exploration.168: 137-149.

[28]. Final Report on Advanced Exploration in Mining Area of Gazetan. Parsi Kan Kav Engineering Company. 2015.

[29]. Reimann, C., Filzmoser, P. and Garrett, R.G. (2002). Factor analysis applied to regional geochemical data: 301 problems and possibilities. Applied Geochem. 17: 185-206.

[30]. Yousefi, M. and Carranza, E.J.M. (2015). Prediction- area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. Computer and Geosciences. 79: 69-81.

[31]. Yousefi, M., Kamkar-Rouhani, A. and 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 Geochem Exploration. 115: 24-35.

[32]. Yousefi, M., Kamkar-Rouhani, A. and Carranza, E.J.M. (2013). Application of staged factor analysis and logistic function to create a fuzzy stream sediment geochemical evidence layer for mineral prospectivity mapping. Geochemistry: Exploration, Environmental, Analysis. 14 (1): 45-58.

[33]. Deutsch, C. and Journel, A.G. (1998). GSLIB: Geostatistical Software Library and User's GuideSecond Edition. Oxford University Press, New York. 340 P.

[34]. Goovaerts, P. (1997). Geostatistics for natural resources evaluation. Oxford University Press, New York. 483 P.

[35]. Tehrani, M.M., Asghari, O. and Emery, X. (2013). Simulation of mineral grades and classification of mineral resources by using hard and soft conditioning data: application to Sungun porphyry copper deposit. Arabian Journal of Geosciences. 6 (10): 3773-3781.

[36]. Chilès, J.P. and Delfiner, P. (2012). Geostatistics: Modeling Spatial Uncertainty, 2nd Edition.Wiley. 734 P.

[37]. Barnett, R.M., Manchuk, J.G. and Deutsch, C.V. (2014). Projection pursuit multivariate transformation. Mathematical Geosciences. 46: 337-359.

[38]. Afzali, S. and Nezafati, N. (2011). Petrogenesis and Mineralization at Gazestan Iron Oxid-Apatit Deposit, East of Bafgh, Central Iran. Scientific Quarterly Journal, Geosciences. 24 (93): 77-84.

مدل‌سازی ویژگی‌های عناصر نادر خاکی با استفاده از ترکیب مطالعات آماری چندمتغیره، آنالیز فاکتوری و شبیه‌سازی زمین‌آماری؛ مطالعه موردی کانسار گزستان، ایران مرکزی

فاطمه سلطانی‌[۱]، پرویز معارف‌وند‌[۱]*، فیروز علی نیا‌[۱] و پیمان افضل‌[۲]

۱- دانشکده مهندسی معدن و متالورژی، دانشگاه صنعتی امیرکبیر، ایران

۲- بخش مهندسی نفت و معدن، واحد تهران جنوب، دانشگاه آزاد اسلامی، ایران

* نویسنده مسئول مکاتبات: parvizz@aut.ac.ir

چکیده:

رویکردهای سنتی مدل‌سازی و تخمین در کانسارهای عناصر نادر خاکی به برآورد نادرست منجر شده و مدیریت منبع را با چالش و ریسک مواجه می‌کند. عیار پایین در کانسارهای عناصر نادر خاکی کشور از یک‌سو و اهمیت استراتژیک آن‌ها ضرورت مدل‌سازی چندمتغیره عیار در این کانسارها را دوچندان می‌کند. تغییرات زیاد عیار و ارتباط عیار با واحدهای مختلف سنگی نیز به پیچیدگی مدل‌سازی عناصر نادر خاکی می‌افزاید. در این پژوهش، کانسار مگنتیت- آپاتیت گزستان با استفاده از روش‌های آماری و زمین‌آماری مورد بررسی و مدل‌سازی قرار گرفته است. در این کانسار عناصر نادر خاکی سبک و سنگین در کانی آپاتیت به صورت انکلوزیون‌های ریز مونازیت متمرکز شده است. با استفاده از ۹۰۸ نمونه عیارسنجی شده ۶۴ عنصره شامل عناصر نادر خاکی سبک و سنگین حاصل از مغزه‌های حفاری و استفاده از روش‌های تحلیل فاکتوری مرحله‌ای، مدل‌سازی فرکتالی عیار، تعداد - عیار و نیز انجام شبیه‌سازی زمین‌آماری تلاش شده است تا عیار عناصر نادر خاکی در واحدهای سنگی مختلف بررسی شود. درنهایت بر اساس نتایج حاصل از تحقق‌ها به آنالیز عدم قطعیت عیاری در کانسار پرداخته شد. کلیه مطالعات چندمتغیره، آنالیز ساختار فضایی، شبیه‌سازی و تحلیل واحدهای سنگی، ارتباط فسفر با کانی‌سازی را تائید می‌کنند.

کلمات کلیدی: شبیه‌سازی زمین‌آماری، عناصر نادر خاکی، کانسار گزستان، آنالیز فاکتوری مرحله‌ای، فرکتال، مدل‌سازی عدم قطعیت.