

A hybrid-based clustering algorithm for targeting porphyry copper mineralization at Chahargonbad district in SE Iran

Hossain Rahimi^a, Maysam Abedi^a, Abbas Bahroudi^{a,*}, Soheila Aslani^a, Gholam-Reza Elyasi^a, Mohammad-Javad Rezapour^a

^a Geo-Exploration Targeting Lab (GET-Lab), School of Mining Engineering, College of Engineering, University of Tehran, Iran

Article History:

Received: 07 January 2020,

Revised: 25 September 2020,

Accepted: 28 September 2020.

ABSTRACT

This work presents a hybrid-based clustering approach for mineral potential mapping (MPM) of porphyry-type Cu mineralization at Kerman province in the SE of Iran. Whereby a multidisciplinary geospatial data set was processed and integrated in the Chahargonbad district. Data-driven prediction-area (P-A) plots were drawn for each evidence layer derived from geological, geochemical, geophysical and satellite imagery data. The P-A plots provide insight into the weight of evidence for synthesizing all geospatial layers. Out of many knowledge-driven methods which biasing from experts' opinions, index overlay and fuzzy operators were employed to find out an optimum Cu favorability map through calculating an efficiency index representing the performance of each MPM. A concentration-area (C-A) fractal model was implemented to separate the mineral favorability map into some populations to ensure correct determining the cluster numbers. Clusters number is a prerequisite which must be defined correctly to increase the performance of clustering analysis for generating reliable results in MPM. Such an appropriate number of clusters can be incorporated in running three prevalent groups of clustering methodologies as data-driven approaches in MPM. They are self-organizing map, fuzzy c-means, and k-means algorithms. One of the reasons for this tendency to consider a hybrid-based method is that it overcomes the shortcomings of the both methods (bias of experts' opinions and unknown clusters number) in mineral favorability mapping. The unknown number of clusters was determined through a knowledge-driven method, and then it was passed to an unsupervised data-driven method, i.e. clustering algorithm. This hybrid method produces synthesized maps in close association with known porphyry-Cu mineralization in the Chahargonbad area.

Keywords : Clustering, Hybrid method, Mineral potential mapping, Porphyry copper, Chahargonbad

1. Introduction

Various sources of uncertainty pertaining to a multidisciplinary geospatial data set impact on the mineral exploration task. Whereby such diversities arising from natural conditions of sophisticated geological models need to be handled systematically for targeting sought deposits. It means that it is genuinely a tough task to accurately inform the potential of an investigated area through exploratory data in association with ore mineralization. To tackle the variety of ingredients in multidisciplinary geospatial data, mineral potential mapping (MPM) is a panacea to delimit the sought area in district or deposit scales [1]. Meanwhile such processing ameliorates the risk, time and cost of an exploration program. Hence, to increase the success rate of this task as an active area of research in the mineral exploration community, many researchers have focused their attempts to develop novel algorithms in the MPM, showing great improvement in identifying the locations of the true ore-bearing anomalies.

Among MPM methodologies developed in the last two decades, three categories of data integration are much investigated, namely (1) knowledge-driven, (2) data-driven, and (3) hybrid [2-5]. Of note is that for implementing the supervised versions of the data-driven methods, the locations of known mineral deposits are required as "training points" to computationally find out their spatial relationships with specific

geological, geochemical and geophysical features [2, 4] Their blind relationships are sought to assign the importance weight of each evidence layer [4, 6]. Such evidence layers are ultimately integrated into a single mineral favorability map, showing foremost favorable regions among studied area in association with the sought deposit type [4, 7]. Examples of the supervised data-driven methods are logistic regression [8-10], neural networks [11-16], weights of evidence [17, 18], support vector machine [19, 20], and random forests [21, 22]. Another point worth taking into account concerns to implement data-driven methods in cases of no training points. A possible scenario to strive for MPM is the utilization of the unsupervised versions of the data-driven methods as clustering algorithms, which indeed divide multi-dimensional feature (or evidence layer) space into some clusters [23-27].

Another group of MPM proposed in the literature is the knowledge-driven methods that are on a basis of geoscientists' opinions [4]. This kind of processing is sometimes time consuming and somewhat arbitrary. Main well-known approaches of this group are Boolean logic [18, 28, 29], index overlay [28-33], fuzzy logic [4, 33-37], outranking methods [28, 38-40], and evidential belief functions [41, 42]. Hybrid algorithms are a combinatory of knowledge- and data-driven methods for tackling the weakness of each group of MPM when running individually [15, 43-45].

In cases of lack of accessing to training points for running supervised versions of the data-driven methods, MPM falls under the

* Corresponding author. E-mail address: bahroudi@ut.ac.ir (A. Bahroudi).

umbrella term of unsupervised algorithms of clustering. Clustering is a process for organizing data in groups whose members are similar. It is applied in a variety of fields, including image processing, data mining, pattern recognition and machine learning, with abilities to manage a large amount of information by categorizing them into several clusters. Clustering can reduce the feature vectors' dimensions. The clustering process includes several steps of feature selection, proximity measure, clustering criterion, clustering algorithm, validation of the results, and its interpretation [23, 46]. Unknown number of clusters which has critical impact on the synthesized evidence layers, should be selected appropriately to divide prospect zone into some populations in association with geological setting of the studied region. An integral part of any clustering analysis in MPM is the validation of the quality of the obtained clusters, where the determination of clusters number from mathematical indices may bear little resemblance to the true geology of a prospect zone. Correct determination of the clusters number has been investigated in multiple fields of engineering often inferred from mathematical indices [47-49]. Following the same line of thought as Rezapour et al. (2020), this work has examined a knowledge-guided clustering methodology, where the optimum number of clusters is geologically determined through implementing a fractal analysis of synthesized evidence layers by index overlay and fuzzy gamma operators [50]. Porphyry-Cu favorability maps were plotted into four geologically meaningful clusters by running fuzzy C-means (FCM), K-means (KM) and self-organizing map (SOM) algorithms. Whereby, the desired cluster in association with porphyry-Cu mineralization was extracted.

The FCM clustering algorithm was developed by Dunn [60] and corrected by Bezdek [61]. This algorithm is often used in pattern recognition. The FCM algorithm allows data to belong to two or more clusters. FCM is sensitive to noise, outlier values and initial conditions, and requires a long computational time [62]. The KM clustering algorithm is the simplest and most common algorithm that uses mean squared quantization error. The above algorithm first randomly determines the cluster centers. These centers should be carefully selected because different initial centers produce different results. The main advantage of the method is its simple implementation, but there are problems that should always be taken into consideration. The KM algorithm has a high sensitivity to randomly selected cluster centers. This means that there is a possibility of stopping the algorithm at a local minimum. This algorithm can be implemented several times to reduce this effect [63].

The SOM clustering is a kind of artificial neural network that introduced by Kohonen [64]. This neural network is trained by unsupervised learning [23]. The self-organizing map algorithm can convert nonlinear statistical relationships between input data into simple geometric relationships. This indicates a nonparametric recursive regression relation, so that regression is performed recursively with each instance. Accordingly, it can be claimed that the SOM can check and correct the error rate [65, 66]. Indeed, it includes two-dimensional relationships as a network of map units that connect to neighborhoods by a neighborhood relationship. The number of map units, which often varies from several to several thousand, determines the accuracy and capability of SOM generalization. Each neuron is represented by a pre-sample weighted vector containing the input vector. In the training phase, the SOM creates a network that blends the mass of input data together. Data that are located close to each other in the input area, are mapped into map units. So, SOM is a topology map that allows the display, interpretation, and arrangement of clustering and is able to map the data input space to a two-dimensional network of map units.

The remainder of this work has been prepared as following sections. Workflow of this study has been summarized in the second section. Geological setting of the Chahargonbad district is presented in the section third. Geospatial data set is constructed in the section fourth, where a multidisciplinary database is designed from geophysical (magnetometry and radiometry data), geological, geochemical and satellite imagery data. In the fifth section, evidence layers are integrated through a hybrid clustering algorithm. After determining the weight of each evidence through plotting a prediction-area (P-A) curve [45],

optimized knowledge-driven MPM is generated. An efficiency index is used to distinguish the most efficient favorability map over the rest ones. A concertation-area (C-A) fractal curve [51] is subsequently plotted for the synthesized evidences to acquire the number of clusters. Then, it passes to the clustering methodologies that are SOM, FCM, and KM, where they plot evidence layers into some clusters. Then in the sixth section, the performance and quality of clustered mineral favorability maps are discussed by comparison to the Porphyry-Cu occurrences in the prospect region. Finally, the main achievements of this hybrid method are summarized in the conclusion.

2. Workflow of this study

Since knowledge-based methods suffer from bias weighting to each evidence layer, data-driven methods can resolve this issue in MPM. This study follows a hybrid-based algorithm to implement three clustering methods in two steps (Fig.1). In step 1, after constructing a geospatial database from a multidisciplinary data set, the P-A plot as a data-driven approach is portrayed to determine the weight of each evidence layer. These weights are incorporated in running knowledge-based methods of index overlay and fuzzy operator to find out an optimum synthesized mineral favorability map, where an efficiency index criterion is used to evaluate the efficiency of each MPM. To end the first step, a C-A fractal method is applied to the best mineral favorability map, extracting the number of populations that is indeed the number of clusters for the second step. For running any clustering methodology in the second step, such a clustering number is assumed. Then, each generated cluster is evaluated to match it with the background geological setting of the studied area, leading to introducing the main cluster in association with the porphyry-Cu mineralization. Clustering output suffers less from experts' bias weighting to evidence layers and can be more reliable rather than the knowledge-based outputs.

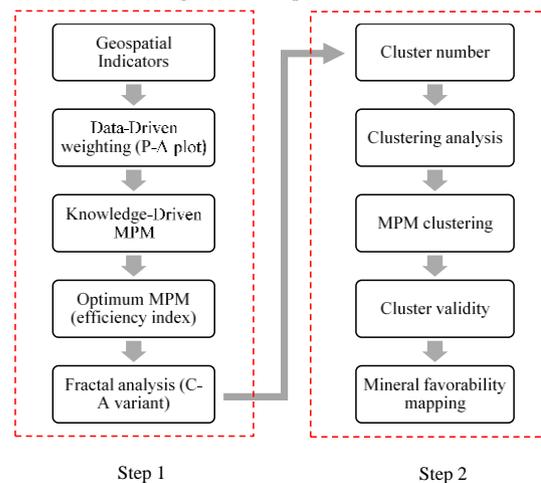


Fig. 1 The proposed diagram for the hybrid-based clustering algorithm in MPM.

3. Geological setting of the Chahargonbad district

Urumieh-Dokhtar magmatic arc (UDMA) is associated with the subduction of the Neo-Tethys oceanic plate and then the Arabic plate below Iran during the middle Miocene (about 13 million years ago) [52]. UDMA is the most important Iranian metallogeny belt, and indeed matches on the Alpine-Himalayan orogenic universal belt. This magmatic belt is affected by alpine orogeny phases from the time of the Mesozoic to the late Cenozoic. UDMA is the main host of the porphyry Cu deposits in Iran (Fig.2a), along with other deposits related such as Au and Mo pertaining to this geodynamic origin. This ore-bearing belt is arising from the developments caused by the closure of the Neo-Tethys ocean, or in other words the digestion of the Neo-Tethys oceanic crust through its subduction to the continental subfloor. Most porphyry Cu deposits of Iran, such as Sarcheshmeh, Miduk and Sungun copper,

are occurred along this belt [53, 54]. UDMA includes two major mineralization areas, the Chahargonbad region to the south and the Sungun region to the NW. Dominant mineralization in this region is porphyry-Cu associated with the Eocene-Miocene plutonic bodies and volcanic rocks [54].

The Chahargonbad area is situated at the NE of the city of Sirjan, in Kerman province of Iran and in the mountainous region (Fig. 2a). The largest outcrops of rock units in the Chahargonbad district are the Eocene volcano-sedimentary rocks. At the end of the Miocene, tectonic movements have caused the folding of the rock masses with a NW- SE trend [55]. It's worth pointing out that magmatic bodies have intruded older rocks in the region at five phases as follows,

(1) The colored melange complex includes a set of mafic and ultra-mafic intrusive rocks, spilitic rocks, volcanic breccia and calcareous sediments. It is generally believed that this structure concerns to the oceanic crust that was formed during the Cretaceous.

(2) Andesitic-trachytic lavas with tuffs and other pyroclastic rocks related to the Eocene.

(3) Granite-granodiorite rocks of the Eocene.

(4) Intrusive bodies of quartz diorite-granodiorite related to lower Miocene. The aforementioned rocks are of particular importance due to the fact that they are an important factor in the formation of copper ores in the region.

(5) Hypabyssal rocks of the trachy-basalt related to the Pliocene. At this phase, dykes and many sills with an irregular trend have cut off the older rocks [55, 56].

Plutonic masses have intruded at two main phases. Older masses include post Eocene-Oligocene granite and granodiorite extensively in the region, and younger masses are Miocene quartz diorite combination mainly spread in some parts of the Chahargonbad area. The utmost importance rock type in the region in association with the porphyry Cu mineralization is quartz diorite. Therefore, these bodies have more priority for porphyry Cu exploration. The general trend of these masses is along the NW-SE direction. Table 1 presents the geological characteristics of 28 deposits/prospects in the Chahargonbad district (Fig. 2b), where they are mostly controlled by the intrusive magmatic and hydrothermally units.

Table 1. The geological descriptive summaries of 28 deposits/prospects in the studied region.

#	Name	Ore type	#	Name	Ore type
1	Sarbagh	Intrusive Porphyry	15	Roode Tangoo	Intrusive Hydrothermal
2	Gah Dij	-	16	Roode Shelang	Volcanic Hydrothermal
3	Band Bagh	Volcanic Hydrothermal	17	Takht (Soltan Hosein)	Skarn
4	Koohpanje(Bande Mozafar)	Intrusive Porphyry	18	Chahargonbad	Intrusive Vein
5	Chehel Shomali	Volcanic Hydrothermal	19	Takhte Chahargonbad	-
6	Koohpanje1	-	20	Takht Bonie	Skarn
7	Koohpanje2	-	21	Takht Gonbad Sirjan	Intrusive Porphyry
8	Koohpanje3	-	22	Takht Gonbad Sirjan 1	Intrusive Porphyry
9	Koohpanje4	-	23	Takht Gonbad Sirjan 2	Intrusive Porphyry
10	North Ab Talkhon	Intrusive Vein	24	Bolboli(Soltan Hosein)	Intrusive Vein
11	Central Ab Talkhon	Intrusive Vein	25	Parsan	-
12	West Ab Talkhon	Intrusive Vein	26	Zangu	-
13	South Ab Talkhon	Intrusive Vein	27	Bab Zanguee1	-
14	Chehel Tone Jonoobi	Volcanic Hydrothermal	28	Bab Zanguee2	-

4. Geospatial data sets

In this study, eight geospatial evidence layers are used for MPM. They are extracted from geological map, satellite images, stream sediment geochemical samples, and airborne geophysics (radiometric and magnetic data) to construct a multidisciplinary database.

4.1. Geological layers

Three evidential layers were extracted by expert decision makers from the geology map. Miocene quartz diorite was selected as the main host rock for copper mineralization in the study area [54, 57]. Hence, four 125-m-interval buffers were considered around this unit in Fig. 3a to highlight the importance of adjacent regions. The alteration layer, depicting advanced argillic, phyllic, iron oxide (jarosite) and argillic which were distinguished from the processing of satellite imagery data (ASTER and ETM data), was prepared in Fig. 3b [57]. The faulted and lineament features were extracted from the fault traces derived from geological field survey, and hidden faults from processing airborne magnetic data. Then, four 100-m-interval buffers were considered around these lineaments to capture the importance of those features (Fig. 3c). As can be seen from evidential layers in Fig. 3, deposits and prospects are located in proximity to those favorable regions, while fault evidence could closely localize most of these targets in Fig. 3c.

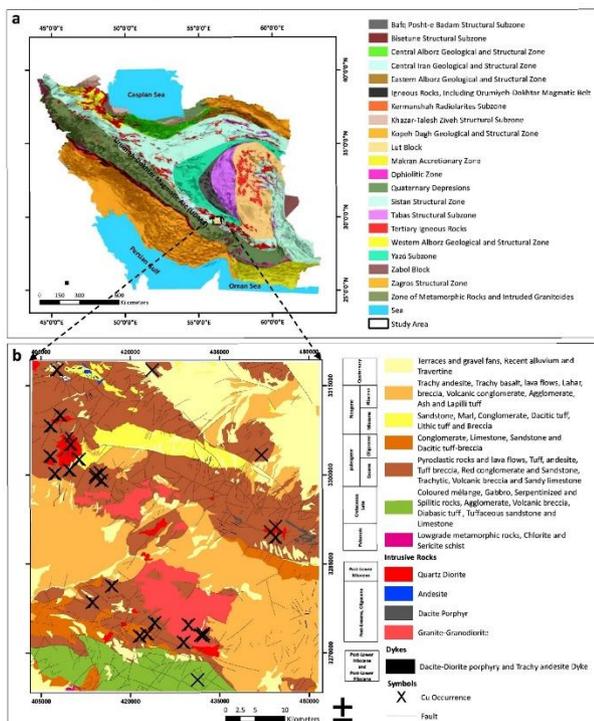
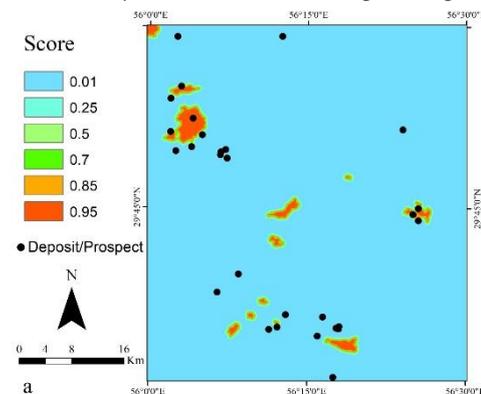


Fig. 2 (a) The location of the studied area on a map of tectono-sedimentary zones of Iran, and (b) a simplified geological map of the Chahargonbad area [55, 57].



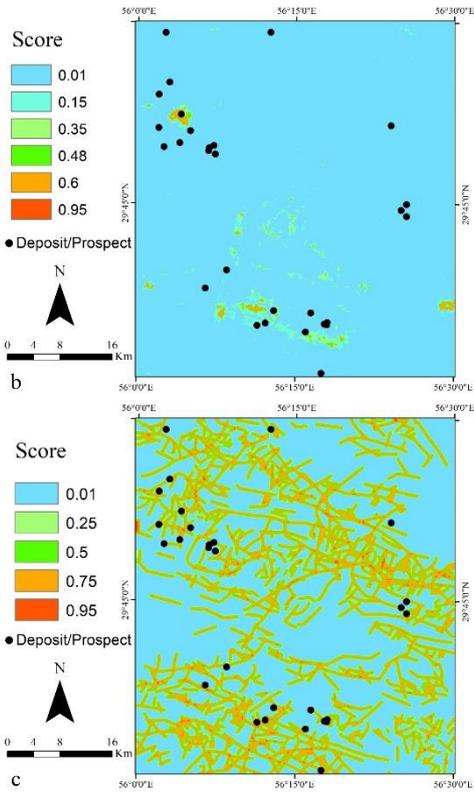


Fig. 3 The geological evidence layers, (a) rock type, (b) alteration, and (c) fault.

4.2. Geochemical layers

Geochemical data used in this study were collected by an experienced group under the supervision of the Geological Survey of Iran (GSI) at a scale of 1:100,000. They were analyzed during a systematic exploration project through the ICP-MS instrument. Totally 846 samples were collected from the stream sediments in the Chahargonbad district, while only elements of Zn, Pb, Cu, Mo and Ag have been processed in this study. It's worth pointing out that a logarithmic conversion was carried out to normalize the distribution of elements which indeed can amplify spatial correlation among footprint elements of Cu-bearing targets. After statistical analysis of samples, two elements of Pb and Zn had meaningful spatially correlation with the Cu element in this region. The statistical characteristics of these elements are presented in Table 2. In Table 3, the Pearson's linear correlation coefficients of these elements have been tabulated. Three geochemical evidences were respectively prepared for Cu, Pb and Zn in Figs. 4a, 4b and 4c. In the NW of the area, deposits/prospects are much correlated with the Cu evidence layer, while in southern portions Pb and Zn evidences have better consistency with the Cu occurrences.

Table 2. Statistical summaries of the main geochemical elements (in ppm unit).

Parameters	Cu	Mo	Pb	Zn	Ag
Mean	65.0686	1.7293	23.8168	113.8073	.0270
Median	58.0000	2.0000	17.0000	94.0000	.0200
Mode	52.00	2.00	14.00	70.00	.02
Std. Deviation	53.44284	1.03376	25.29473	80.37311	.08229
Variance	2856.137	1.069	639.823	6459.836	.007
Skewness	13.289	4.142	6.124	2.572	11.769
Kurtosis	257.100	29.814	57.273	11.581	136.822
Range	1198.00	11.00	319.00	750.00	.98
Minimum	2.00	1.00	2.00	2.00	.02
Maximum	1200.00	12.00	321.00	752.00	1.00

Table 3. The Pearson's linear correlation coefficient between concentration of five elements.

Cu	1				
Mo	-0.005	1			
Pb	0.428	-0.171	1		
Zn	0.411	-0.148	0.529	1	
Ag	0.094	-0.061	0.107	0.177	1
	Cu	Mo	Pb	Zn	Ag

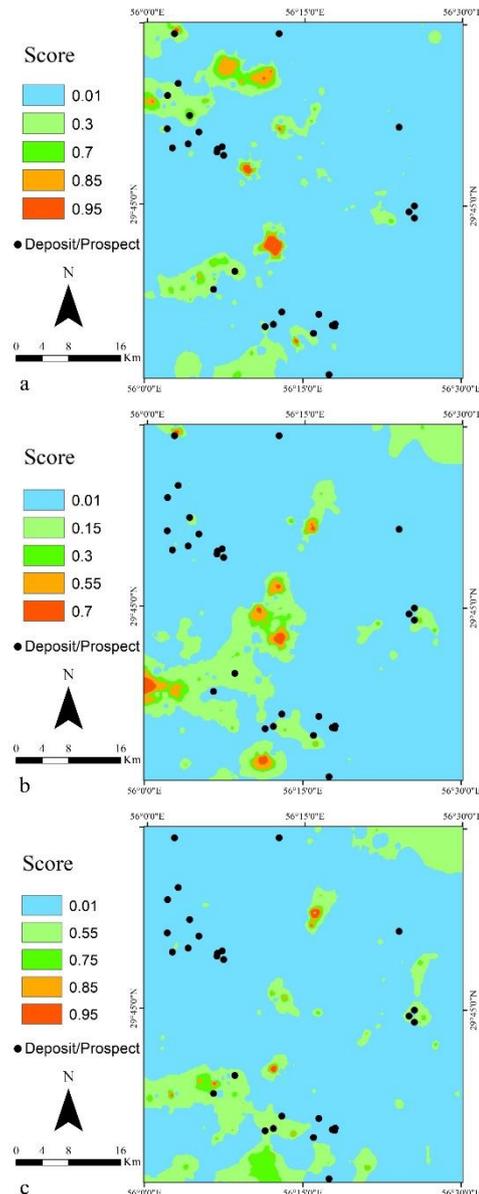


Fig. 4 The geochemical evidence layers, (a) Cu, (b) Pb, and (c) Zn.

4.3. Geophysical layers

Airborne geophysical survey (magnetometry and radiometry) was carried out in 1977 under the supervision of the Atomic Energy Organization of Iran. Flight line spacing and altitude of the survey were chosen 500 and 150 m, respectively. Directional derivatives of both the reduced-to-pole magnetometry data (RTP) along with the upward continued data were calculated to enhance respectively the borders of shallow and deep-seated magmatic sources in association with the porphyry-type Cu-bearing mineralization. As firstly stated by Nabighian [58], the amplitude of the directional derivatives can enhance the

borders and the main traces of bodies responsible for magnetic anomalies. Figure 5a presents the evidence map of intrusive bodies buffered with some narrow rings to highlight the importance of adjacent areas. The ratio of K/Th was also calculated to delineate regions controlled by potassic and phyllic alteration along with rock units responsible for probable Cu occurrences [59]. Figure 5b shows the radiometric evidence layer more compatible with deposits/prospects in comparison to the intrusive bodies.

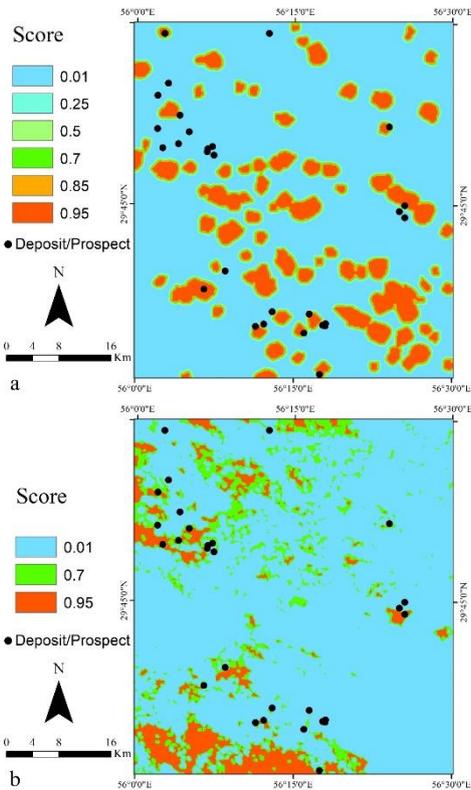


Fig. 5 The geophysical evidence layers, (a) intrusive bodies from magnetometry data, and (b) K/Th ratio.

5. Hybrid-based clustering methodology

Exploratory geospatial data set consists of eight evidence layers that were derived from three criteria of geology (rock type, alteration and fault), geochemistry (Cu, Pb and Zn), and geophysics (intrusive bodies and K/Th). Figure 6 presents a decision tree network for inferring final mineral favorability map.

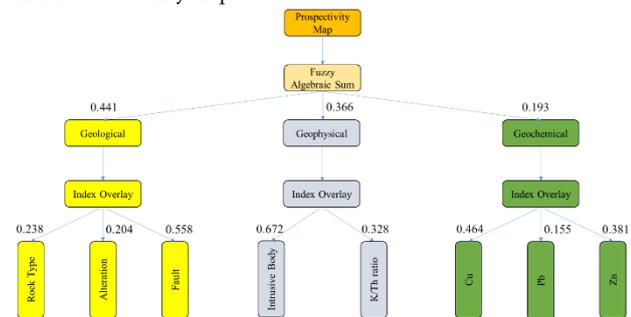


Fig.6 Decision tree network for inferring final mineral favorability map and the weight extracted from the P-A plot.

5.1. Knowledge-driven method

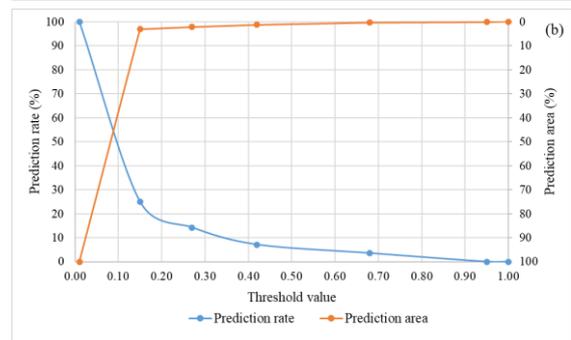
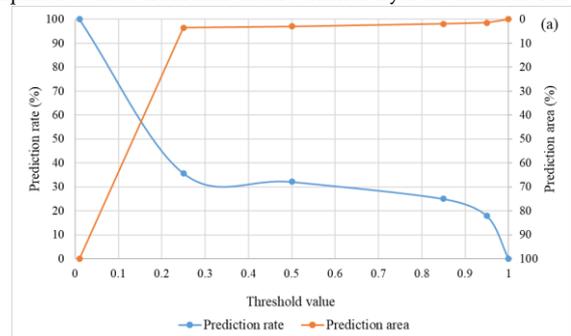
Since evidence layers have different weights of importance in synthesizing all layers, the P-A plots are being drawn to determine each layer's weight on the basis of simultaneous consideration of the ore prediction rate and the area of favorability. For a more rigorous and in-

depth introduction to the P-A plot, readers can refer to Yousefi and Carranza [45]. To determine the weight of each layer, first, the value of the ore prediction rate at the point of intersection is divided by its prediction area, and then the normalized density is obtained. The natural logarithm of the normalized density is calculated to obtain the weight of each layer. Table 4 has summarized the procedure of generating a data-driven approach for assigning weights. The locations of known deposits/prospects listed in Table 1, were taken into consideration for depicting all the P-A plots. Figure 7 presents all the P-A plots derived from three main groups of criteria that were geological, geochemical and geophysical features. This analysis indicates that the utmost important layers are faults and intrusive bodies. It reveals the close connection between the Cu deposition systems with the faults and intrusive bodies. Intrusive bodies as the main source of ore-forming fluids and faults as transfer structures to the ground level, always are key factors in the constitution of the porphyry deposits. The results of this research demonstrate that the geochemical layers have lower impact on ore formation, and the cause of this result is that geochemical samples were surveyed from the ground level. Thus, they illustrate elements distribution in shallow level with lower information from depth.

Table 4. The data-driven weight extraction through the P-A plots of eight evidence layers.

Layers	Prediction area (%)	Prediction rate (%)	Normalized density	Normalized weight
Rock Type	43	57	1.33	0.105
Alteration	44	56	1.27	0.090
Fault	34	66	1.94	0.246
Cu	44	56	1.27	0.090
Pb	48	52	1.08	0.030
Zn	45	55	1.22	0.074
Intrusive Body	34	66	1.94	0.246
K/Th Ratio	42	58	1.38	0.120

Upon determining the weight of evidential layers through the P-A plots, two knowledge-based method of index overlay (IO) and fuzzy gamma operator (FGO) were utilized to integrate all layers in a single mineral favorability map. Figure 8b presents a potential map derived from the IO method which has been reclassified into four populations on the basis of its C-A fractal curve shown in Fig. 8a. The P-A plot of this favorability map shows that the amounts of ore prediction rate and potential area at the intersection point are equal to 72 % and 28%, respectively (Fig.8c). Note that these values present better results compared to those from individual evidence layer tabulated in Table 4.



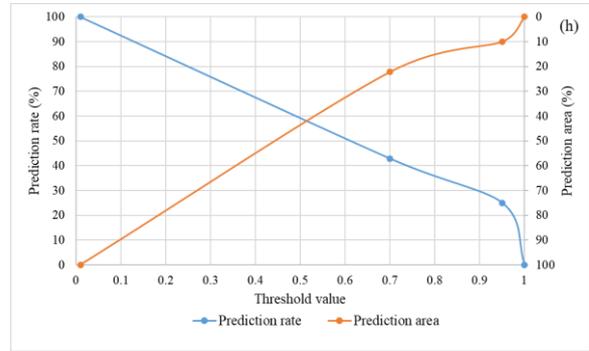
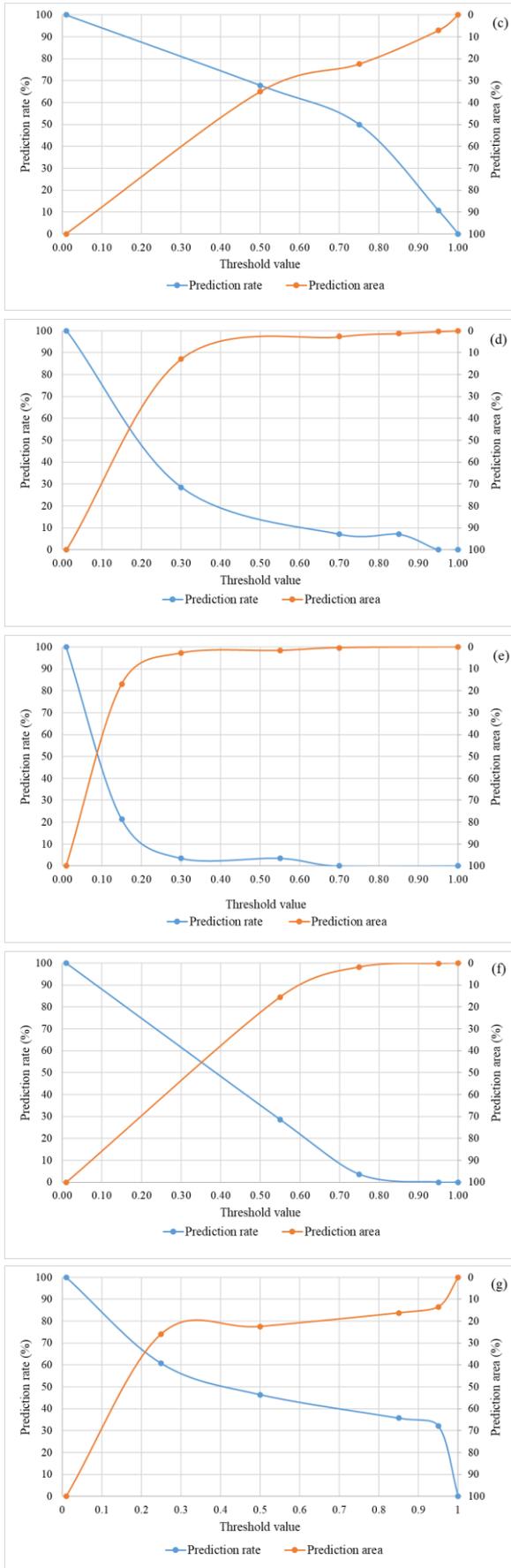


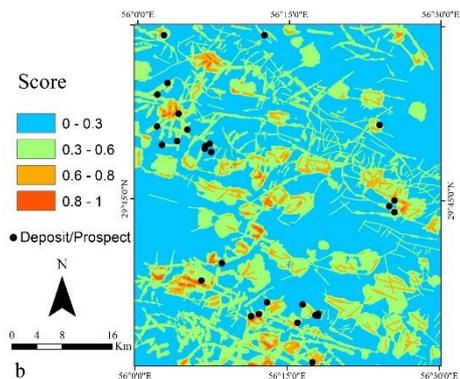
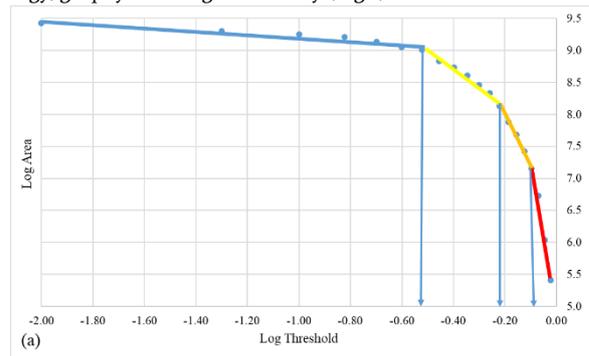
Fig.7 The P-A plots for eight evidence layers of (a) rock type, (b) alteration, (c) fault, (d) Cu, (e) Pb, (f) Zn, (g) intrusive bodies, and (h) K/Th. The locations of 28 deposits/prospects listed in Table 1 were used to generate the P-A plots.

Various gamma amounts can be taken into account to run a FGO in MPM. Therefore, it is required to find out the most appropriate one in generating mineral favorability map. Examining several amounts of gamma revealed that the FGO with a value of 1 generates the most productive map (Fig. 9) with a higher efficiency index (EI). The MPM EI was introduced as Eq. (1) [67],

$$\text{MPM Efficiency Index (\%)} = w_1(100 - \text{predicted area \%}) + w_2(\text{ore prediction rate \%}) \quad (1)$$

where $\sum_{i=1}^2 w_i = 1$ and w_i expresses the importance of each criterion. For unbiased weighting of each term defined in Eq. (1) that are the ore prediction rate and the predicted area, equal weight of 0.5 is usually assumed [67]. Higher values of this index present potential map with higher ore prediction rate and lower area as favorability zone.

Assuming different thresholds for each MPM, the MPM EI can be calculated. Figure 9 presents the EI curve for several values of gamma accompanied with the IO output. Highest efficiency value equal to 74.2% was generated for gamma value equal to one at an MPM threshold of 0.45. Fuzzy operator with a gamma value of one is similar to a fuzzy algebraic sum operator (FAS). Figure 10 indicates the maximum and average of the EI for all knowledge-driven methods, depicting higher efficiency of the FAS operator compared to the other FGO and the IO outputs. Note that the FGO was implement for three main criteria of geology, geophysics and geochemistry (Fig.6).



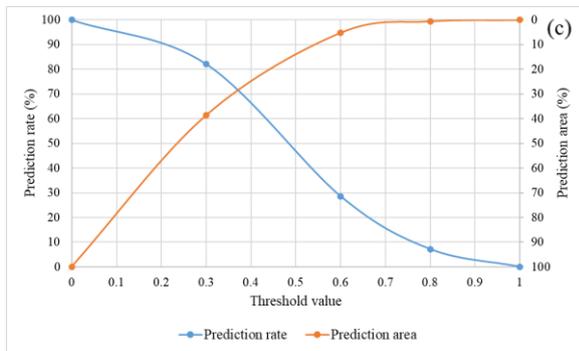


Fig. 8 The index overlay output, (a) the C-A fractal curve, (b) the MPM, and (c) the P-A plot.

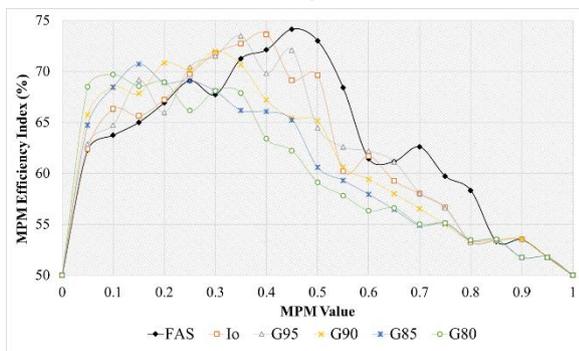


Fig. 9 The curve of the efficiency index values for various MPMs generated by the fuzzy gamma operator and index overly techniques, when different MPM thresholds were taken into consideration.

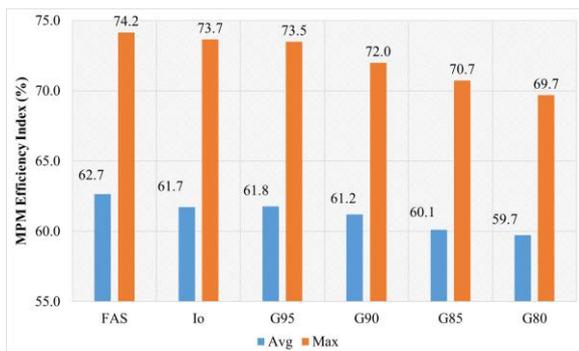


Fig. 10 The efficiency index values for various MPMs generated by the fuzzy gamma operator and index overly techniques, where the most efficient one was obtained for gamma 1.

The optimum knowledge-driven mineral favorability map has been presented in Fig. 11b, while it has been reclassified into four populations on the basis of its fractal model in Fig. 11a. The P-A plot of the FAS map shown in Fig. 11c indicates an ore prediction rate of 72% and occupied about 28% of area. Since this map is the most efficient Cu favorability map, it has been divided into four populations. Thus this number was fixed as the optimum number of clusters.

5.2. Clustering mapping

Assuming four clusters, clustering algorithms were run to map eight evidence layers into four clusters. Note that hexagonal topology with four neurons was used in running the SOM clustering. Figure 12 presents clustering results respectively for the FCM, KM and SOM, in all of which the fourth cluster was in association with more favorable zone in association with porphyry-Cu mineralization. Table 5 presents the amounts of the ore prediction rate and the area of cluster number four for each clustering methodology, showing that efficiency index for the FCM, KM and SOM is respectively 66%, 70.5%, and 75%. KM has the best ore prediction rate but predict great area that decreases its

efficiency index. FCM with the worst ore prediction rate and the middle amount of area has the least efficiency index. SOM predicts 61% of the known deposits/prospects whereas cluster number four occupies 11% of the study area that demonstrates its effectiveness to delimit the Cu-forming zones.

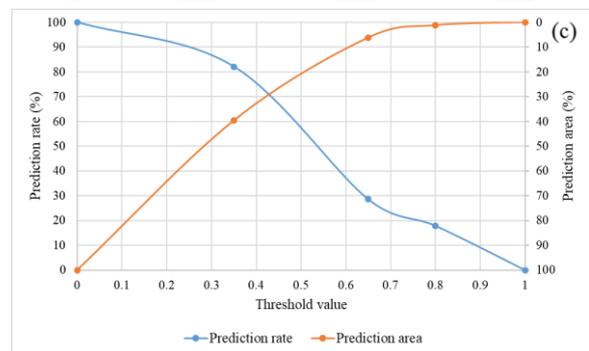
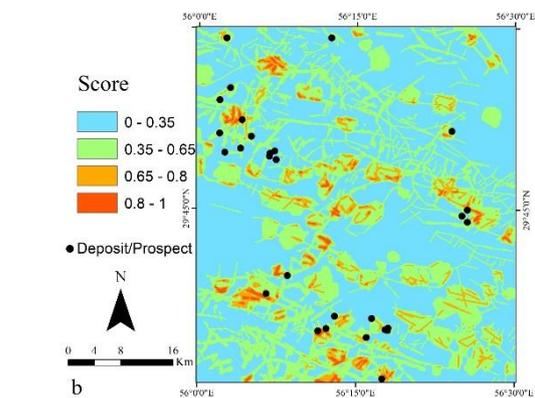
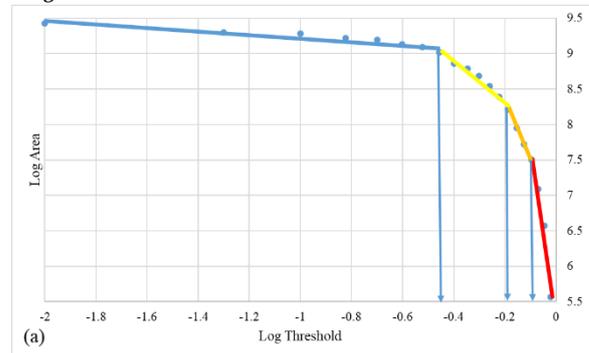


Fig. 11 The optimum hybrid MPM (fuzzy algebraic sum), (a) the C-A fractal curve, (b) the MPM, and (c) the prediction-area (P-A) plot.

Table 5. The amounts of the ore prediction rate and area for cluster number four, and knowledge-driven methods at the threshold values with the highest efficiency.

Method	Ore prediction rate (%)	Area (%)	Efficiency Index (%)
FCM	54	22	66
KM	68	27	70.5
SOM	61	11	75
IO	67.9	20.5	73.7
FAS	71.4	23.1	74.2

6. Discussion

To apply a hybrid-based MPM, eight evidence layers were designed in an exploratory geospatial data set. In other words, the feature vectors are eight dimensional in this case. After clustering, this dimension reduces to four clusters. This is the most important advantage of the clustering approach. Cluster number one for three methods concerns to

the background geological setting that has low potential for mineralization. This area is compatible with the Quaternary recent alluvium and ash tuff, lahar and breccia related mainly to Pliocene. The second cluster of the FCM and KM is more match with the K/Th layer. Colored melange related to Cretaceous and Eocene andesitic rocks are relevant to this class. This cluster in SOM map is similar to the fault layer to some extent. Cluster number three for three methods delineates intrusive bodies. This class locates in the UDMA zone. The fourth cluster of FCM and KM is matched to the fault layer, and in SOM is accordant with overlap of the intrusive bodies and faults. These areas have high potential for Cu mineralization, and can be powerful footprints in the Chahargonbad district. Among the above-mentioned methods, results of SOM is highly compatible with the FAS operator.

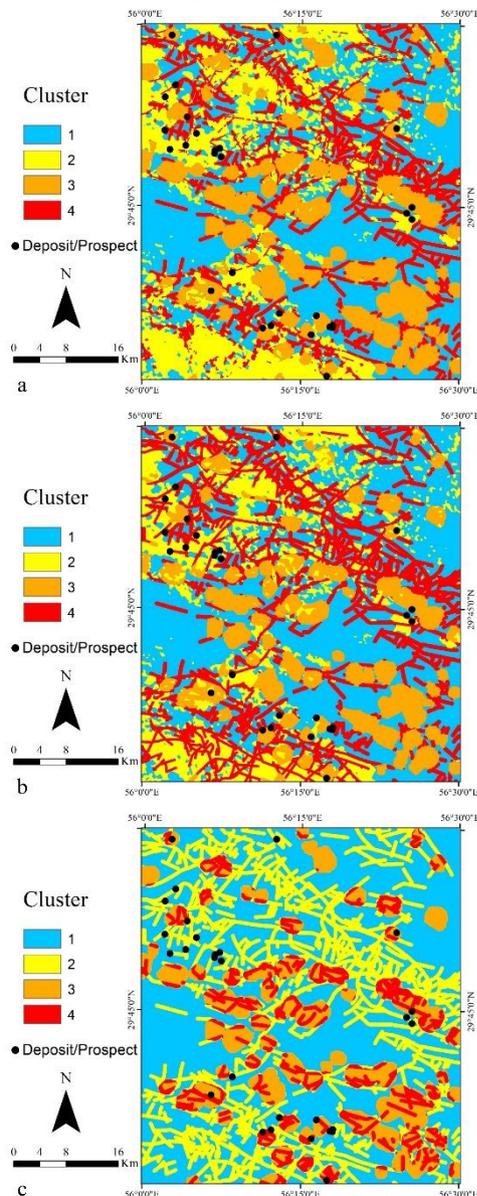


Fig. 12 The hybrid-based clustering outputs, (a) FCM, (b) KM, and (c) SOM, where the optimum cluster number was estimated from the fractal analysis of the optimum gamma operator.

7. Conclusion

In this study, the integration of knowledge-based and clustering algorithms was used as a hybrid method to prepare mineral favorability map for porphyry Cu exploration in the Kerman province of Iran. The Chahargonbad district situated in this region, was selected for this study.

To implement the hybrid method, eight evidential layers were prepared as input criteria from processing geological, geochemical, geophysical and satellite imagery data. Among the knowledge-driven outputs, the fuzzy algebraic sum operator has the highest efficiency in mineral potential mapping. On the basis of the optimum clusters number obtained from the fuzzy sum operator, the clustering algorithms were applied to eight geospatial layers to reduce the dimensions of the feature vectors into four clusters. The promising areas that obtained from the SOM clustering are appropriately more consistent with the desired ore-bearing targets, showing the superiority of the hybrid method over the conventional methods.

Acknowledgements

The authors gratefully acknowledge the support provided by the School of Mining Engineering, University of Tehran. We also express our sincere thanks to the National Iranian Copper Industries Company (NICICO) for providing required data.

REFERENCES

- [1] Yousefi, M., & Kamkar Rouhani, A. (2010). Principle of mineral potential modeling (Vol. 1). Amir Kabir University, Tehran.
- [2] Carranza, E.J.M. (2008). Geochemical anomaly and mineral prospectivity mapping in GIS (Vol. 11). Elsevier.
- [3] Carranza, E.J.M. (2017). Natural resources research publications on geochemical anomaly and mineral potential mapping, and introduction to the special issue of papers in these fields. *Natural Resources Research*, 26(4), 379-410.
- [4] Kashani, S.B.M., Abedi, M., & Norouzi, G.H. (2016). Fuzzy logic mineral potential mapping for copper exploration using multi-disciplinary geo-datasets, a case study in seridune deposit, Iran. *Earth Science Informatics*, 9(2), 167-181.
- [5] Pan, G. & Harris, D. (2000). Information synthesis for mineral exploration: Oxford Univ. Press, New York.
- [6] Carranza, E.J.M. & Hale, M. (2002). Where porphyry copper deposits are spatially localized? A case study in Benguet province, Philippines. *Natural Resources Research*, 11(1), 45-59.
- [7] Nykänen, V., & Salmirinne, H. (2007). Prospectivity analysis of gold using regional geophysical and geochemical data from the Central Lapland Greenstone Belt, Finland. *Geological Survey of Finland*, 44, 251-269.
- [8] Agterberg, F., & Bonham-Carter, G.F. (1999). Logistic regression and weights of evidence modeling in mineral exploration. in *Proceedings of the 28th International Symposium on Applications of Computer in the Mineral Industry (APCOM)*, Golden, Colorado.
- [9] Carranza, E.J.M., & Hale, M. (2001). Logistic regression for geologically constrained mapping of gold potential, Baguio district, Philippines. *Exploration and Mining Geology*, 10(3), 165-175.
- [10] Mejía-Herrera, P., Royer, J.J., Caumon, G., & Cheilletz, A. (2015). Curvature attribute from surface-restoration as predictor variable in Kupferschiefer copper potentials. *Natural Resources Research*, 24(3), 275-290.
- [11] Abedi, M., & Norouzi, G.H. (2012). Integration of various geophysical data with geological and geochemical data to determine additional drilling for copper exploration. *Journal of Applied Geophysics*, 83, 35-45.
- [12] Nykänen, V. (2008). Radial basis functional link nets used as a prospectivity mapping tool for orogenic gold deposits within the Central Lapland Greenstone Belt, Northern Fennoscandian Shield.

- Natural Resources Research, 17(1), 29-48.
- [13] Harris, D., Zurcher, L., Stanley, M., Marlow, J., & Pan, G. (2003). A comparative analysis of favorability mappings by weights of evidence, probabilistic neural networks, discriminant analysis, and logistic regression. *Natural Resources Research*, 12(4), 241-255.
- [14] Porwal, A., Carranza, E.J.M., & Hale, M. (2003). Artificial neural networks for mineral-potential mapping: a case study from Aravalli Province, Western India. *Natural resources research*, 12(3), 155-171.
- [15] Porwal, A., Carranza, E.J.M., & Hale, M. (2004). A hybrid neuro-fuzzy model for mineral potential mapping. *Mathematical Geology*, 36(7), 803-826.
- [16] 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(8), 1017-1023.
- [17] Agterberg, F., Bonham-Carter, G.F., & Wright, D. (1990). Statistical pattern integration for mineral exploration, in *Computer applications in resource estimation*. Elsevier, 1-21.
- [18] Bonham-Carter, G.F. (1989). Weights of evidence modeling: a new approach to mapping mineral potential. *Statistical applications in the earth sciences*, 171-183.
- [19] Abedi, M., Norouzi, G.H., & Bahroudi, A. (2012). Support vector machine for multi-classification of mineral prospectivity areas. *Computers & Geosciences*, 46, 272-283.
- [20] Shabankareh, M., & Hezarkhani, A. (2017). Application of support vector machines for copper potential mapping in Kerman region, Iran. *Journal of African Earth Sciences*, 128, 116-126.
- [21] Carranza, E.J.M., & Laborte, A.G. (2016). Data-driven predictive modeling of mineral prospectivity using random forests: a case study in Catanduanes Island (Philippines). *Natural Resources Research*, 25(1), 35-50.
- [22] Zhang, Z., Zuo, R., & Xiong, Y. (2016). A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. *Science China Earth Sciences*, 59(3), 556-572.
- [23] Abedi, M., Norouzi, G.H., & Torabi, S.A. (2013). Clustering of mineral prospectivity area as an unsupervised classification approach to explore copper deposit. *Arabian Journal of Geosciences*, 6(10), 3601-3613.
- [24] Eberle, D.G., & Paasche, H. (2012). Integrated data analysis for mineral exploration: A case study of clustering satellite imagery, airborne gamma-ray, and regional geochemical data suites. *Geophysics*, 77(4), 167-176.
- [25] Paasche, H., & Eberle, D.G. (2009). Rapid integration of large airborne geophysical data suites using a fuzzy partitioning cluster algorithm: a tool for geological mapping and mineral exploration targeting. *Exploration Geophysics*, 40(3), 277-287.
- [26] Ghezelbash, R., Maghsoudi, A., & Carranza, E.J.M. (2019). Mapping of single-and multi-element geochemical indicators based on catchment basin analysis: Application of fractal method and unsupervised clustering models. *Journal of Geochemical Exploration*, 199, 90-104.
- [27] Daviran, M., Maghsoudi, A., Cohen, D., Ghezelbash, R., & Yilmaz, H. (2019). Assessment of Various Fuzzy c-Mean Clustering Validation Indices for Mapping Mineral Prospectivity: Combination of Multifractal Geochemical Model and Mineralization Processes. *Natural Resources Research*, 29(1), 229-246.
- [28] Abedi, M., Torabi, S.A., Norouzi, G.H., Hamzeh, M., & Elyasi, G.R. (2012). PROMETHEE II: a knowledge-driven method for copper exploration. *Computers & Geosciences*, 46, 255-263.
- [29] Bonham-Carter, G.F., (1994). *Geographic information systems for geoscientists-modeling with GIS*. *Computer methods in the geoscientists*, 13, 398.
- [30] Carranza, E.J.M., Mangaoang, J.C., & Hale, M., (1999). Application of mineral exploration models and GIS to generate mineral potential maps as input for optimum land-use planning in the Philippines. *Natural Resources Research*, 8(2), 165-173.
- [31] Mirzaei, M., Afzal, P., Adib, A., Khalajmasoumi, M., & Zia Zarifi, A. (2014). Prospection of Iron and Manganese Using Index Overlay and Fuzzy Logic Methods in Balvard 1: 100,000 Sheet, SE Iran. *Iranian Journal of Earth Sciences*, 6(1), 1-11.
- [32] Sadeghi, B., Khalajmasoumi, M., Afzal, P., & Moarefvand, P. (2014). Discrimination of iron high potential zones at the zaghia iron ore deposit, bafq, using index overlay GIS method. *Iran J Earth Sci*, 6, 91-98.
- [33] Sadeghi, B., & Khalajmasoumi, M. (2015). A futuristic review for evaluation of geothermal potentials using fuzzy logic and binary index overlay in GIS environment. *Renewable and Sustainable Energy Reviews*, 43, 818-831.
- [34] Abedi, M., Torabi, S.A., & Norouzi, G.H. (2013). Application of fuzzy AHP method to integrate geophysical data in a prospect scale, a case study: Seridune copper deposit. *Bollettino di Geofisica Teorica ed Applicata*, 54(2).
- [35] An, P., Moon, W.M., & Rencz, A. (1991). Application of fuzzy set theory for integration of geological, geophysical and remote sensing data. *Canadian Journal of Exploration Geophysics*, 27(1), 1-11.
- [36] Chung, C.J.F., & Moon, W.M. (1991). Combination rules of spatial geoscience data for mineral exploration. *Geoinformatics*, 1991, 2(2): p. 159-169.
- [37] Moradi, M., Basiri, S., Kananian, A., & Kabiri, K. (2015). Fuzzy logic modeling for hydrothermal gold mineralization mapping using geochemical, geological, ASTER imageries and other geo-data, a case study in Central Alborz, Iran. *Earth Science Informatics*, 8(1), 197-205.
- [38] Abedi, M., Norouzi, G.H., & Fathianpour, N. (2013). Fuzzy outranking approach: a knowledge-driven method for mineral prospectivity mapping. *International Journal of Applied Earth Observation and Geoinformation*, 21, 556-567.
- [39] Abedi, M., Norouzi, G.H., & Fathianpour, N. (2015). Fuzzy ordered weighted averaging method: a knowledge-driven approach for mineral potential mapping. *Geophys Prospect*, 63, 461-477.
- [40] Abedi, M., Torabi, S.A., Nourozi, G.H., & Hamzeh, M. (2012). ELECTRE III: A knowledge-driven method for integration of geophysical data with geological and geochemical data in mineral prospectivity mapping. *Journal of applied geophysics*, 87, 9-18.
- [41] Moon, W.M. (1990). Integration of geophysical and geological data using evidential belief function. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4), 711-720.
- [42] Tangestani, M.H., & Moore, F. (2002). The use of Dempster-Shafer model and GIS in integration of geoscientific data for porphyry copper potential mapping, north of Shahr-e-Babak, Iran. *International Journal of Applied Earth Observation and Geoinformation*, 4(1), 65-74.
- [43] Pazand, K., & Hezarkhani, A. (2015). Porphyry Cu potential area selection using the combine AHP-TOPSIS methods: a case study in Siahrud area (NW, Iran). *Earth Science Informatics*, 8(1), 207-

- 220.
- [44] Yousefi, M., & Carranza, E.J.M. (2016) Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. *Natural Resources Research*, 25(1), 3-18.
- [45] Yousefi, M., & 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. *Computers & Geosciences*, 79, 69-81.
- [46] Theodoridis, S., & Koutroumbas, K. (2009). *Pattern recognition*. Elsevier Inc.
- [47] Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411-423.
- [48] Saha, S., & Bandyopadhyay, S. (2012). Some connectivity based cluster validity indices. *Applied Soft Computing*, 12(5), 1555-1565.
- [49] Ghezlbash, R., Maghsoudi, A., & Carranza, E.J.M. (2020). Optimization of geochemical anomaly detection using a novel genetic K-means clustering (GKMC) algorithm. *Computers & Geosciences*, 134, 104335.
- [50] Rezapour, M.J., Abedi, M., Bahroudi, A., & Rahimi, H. (2019). A clustering approach for mineral potential mapping: A deposit-scale porphyry copper exploration targeting. *Geopersia*, 10(1).
- [51] Cheng, Q., Agterberg, F., & Ballantyne, S. (1994). The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration*, 51(2), 109-130.
- [52] Alavi, M. (1994). Tectonics of the Zagros orogenic belt of Iran: new data and interpretations. *Tectonophysics*, 229(3-4), 211-238.
- [53] Forster, H. (1976). Continental drift in Iran in relation to the Afar structure. *Afar between continental and oceanic rifting (VII)*, 182-190.
- [54] Elyasi, G.R. (2009). Mineral potential mapping in detailed stage using GIS in one of exploration prospects of Kerman Province. Master of Science Thesis, University of Tehran (published in Persian).
- [55] Khan Nazer, N., Emami, M., & Ghaforie, M. (1995). Geological map of Chahargonbad. Geological survey of Iran.
- [56] Yazdi, Z., Jafari Rad, A.R., & Kheyrollahi, H. (2015). Recognition of geological features and alteration zones related to porphyry copper mineralization using airborne geophysical data a case study: Chahargonbad1: 100000 geological map, Kerman province, central iran.
- [57] Elyasi, G.R., Bahroudi, A., Abedi, M., & Rahimi, H. (2020). Weighted Photolineaments Factor (WPF): An Enhanced Method to Generate a Predictive Structural Evidential Map with Low Uncertainty, a Case Study in Chahargonbad Area, Iran. *Natural Resources Research*, 1-33.
- [58] Nabighian, M.N. (1974). Additional comments on the analytic signal of two-dimensional magnetic bodies with polygonal cross-section. *Geophysics*, 39(1), 85-92.
- [59] Ford, K., Keating, P., & Thomas, M. (2007). Overview of geophysical signatures associated with Canadian ore deposits. Geological Association of Canada. Mineral Deposits Division, Special Publication, (5), 939-970.
- [60] Dunn, J.C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters (Vol. 3). *Journal of Cybernetics*.
- [61] Bezdek, J. (1981). *Pattern recognition with fuzzy objective function algorithms-NY*. Plenum Press.
- [62] Barbakh, W.A., Wu, Y., & C. Fyfe, C. (2009). Non-standard parameter adaptation for exploratory data analysis (Vol. 249). Springer.
- [63] Jain, A.K., Murty, M.N., & Flynn, P.J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3), 264-323.
- [64] Kohonen, T. (2012). *Self-organization and associative memory (Vol. 8)*. Springer Science & Business Media.
- [65] Kohonen, T., & Somervuo, P. (1998). Self-organizing maps of symbol strings. *Neurocomputing*, 21(1-3), 19-30.
- [66] Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on neural networks*, 11(3), 586-600.
- [67] Abedi, M., Mohammadi, R., Nourozi, G.H., & Mir Mohammadi, M.S. (2016). A comprehensive VIKOR method for integration of various exploratory data in mineral potential mapping. *Arabian Journal of Geosciences*, 9(6), 482.