



Distance Measures in Mineral Potential Mapping: Squared L2, Shannon and Combination Families

Maysam Abedi *

School of Mining Engineering, Faculty of Engineering, University of Tehran, Iran

Received: 16 July 2020, Revised: 05 October 2020, Accepted: 20 December 2020

© University of Tehran

Abstract

As a complementary work to the first part of the non-Euclidean distance measures in the spatial data decision analysis for mineral potential mapping (MPM), three families of squared L2, Shannon and combination are examined in this research to evaluate the performance of fifteen distances in geospatial data integration. The TOPSIS method as a well-known outranking method in the multi-criteria decision making (MCDM) problem was utilized in the MPM, where its distance kernel was substituted by these new families to investigate the efficiency of different distances in final preparation of the synthesized indicator layers. For consistency with the previous work (i.e. distance measures in spatial data decision analysis), the North Narbaghi porphyry copper deposit in the Saveh, Iran, was revisited as a deposit-scale prospecting zone. The ultimate goal was to prioritize the favorable zones for borehole drilling. The geospatial datasets were derived from a multi-disciplinary survey comprising of geological, geochemical and geophysical criteria. It is indicated that some distances could partially promote the performance of the mineral potential maps in comparison to the traditional Euclidian-based TOPSIS outranking method in the studied case, which is of considerable importance that deserves more investigation in the future.

Keywords: Geospatial Data Integration, Distance Measures, TOPSIS Method, Mineral Potential Mapping, Porphyry Mineralization.

Introduction

Mineral potential mapping (MPM) falls under the umbrella term of a Multi Criteria/Attributes Decision Making (MCDM or MADM) operation which synthesizes several geospatial indicators/attributes to generate a prospectivity map by delimiting favorable mineralized regions. Therefore, it is required to design geospatial datasets (often comprising of geological, geophysical, and geochemical data) at the first stage of the MPM, guided by geoscientist's experts with various disciplines in mineral exploration for indicator preparation, weighting and integration. Among several MPM methodologies, the knowledge-, data-driven and hybrid ones are the most popular categorizes (Yousefi & Nykänen, 2017; Carranza, 2017).

When the prospecting zone has not been well-explored, the knowledge-driven approaches have superiority over the rest methodologies. Among numerous knowledge-driven techniques, the outranking ones are those developed for solving the MCDM problems. They prioritize alternatives of the decision making matrix (equivalent to the geospatial datasets) based on the incorporation of the indicators' weight and score, pairwise comparison of the alternatives and various preference functions (e.g. Pazand et al., 2012; Abedi et al., 2015, 2016, 2017; Pazand & Hezarkhani, 2015; Abedi & Norouzi, 2016). Out of many popular outranking techniques in

* Corresponding author e-mail: maysamabedi@ut.ac.ir

the MCDM problems and operational researches, the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method is actually rather straightforward to be implemented in practice (Opricovic & Tzeng, 2004), and has attracted attention of scholars in MPM (Pazand & Hezarkhani, 2015; Pazand et al., 2012). It works on the basis of the shortest distance from the best alternative as the positive ideal solution (PIS) and the farthest distance from the worst alternative as the negative ideal solution (NIS) to rank alternatives' order. Miscellaneous variants of the TOPSIS approach has been developed to tackle the MCDM problems in varieties of fields (e.g. Tavana & Hatami-Marbini, 2011, Behzadian et al., 2012; Abedi & Norouzi, 2016), in most of which the kernel of algorithms was on the basis of the Euclidian distance measure derived from the Pythagorean metric. In the past decades, substantial efforts have been dedicated to extend the TOPSIS formulation, and to compare (or hybridize) it with other MCDM methods as well (Kuo, 2017).

The performance of fifteen distance measures as the kernel of the TOPSIS method in prioritizing alternatives was profoundly investigated by Abedi (2020) for a multi-disciplinary geospatial dataset, where four distance families of the L_1 , intersection, inner product and fidelity were examined in the MPM. The outputs indicated that some distances compared to the Euclidian one as the default of the TOPSIS method, could outperform the generated maps in terms of correlation with the drilled boreholes for a case study pertaining to a deposit-scale porphyry copper mineralization in the Saveh, Iran. Their excellent performance demonstrated the effectiveness of distance measure replacement. In a similar fashion to the previous work and to complete first part of research about the applicability of distance measures in the MPM as a spatial data decision analysis task (Abedi, 2020), it is imperative to focus on three new families of the squared L_2 , Shannon and combination by investigating another fifteen distances. Our previous datasets were revisited and chosen to compare the results. Indeed, the significance of this study lies in improvement of the performance of the synthesized mineral favorability maps by substitution of the new families of distance measure, which has undergone limited geological research.

The remainder of this work has been organized as follows. The formulation of the distance measures along with the stages of the TOPSIS implementation are presented in the second section. The geological setting of the studied region accompanied with the geospatial datasets from a multi-disciplinary exploratory survey are summarized in the third section. Mineral potential maps were generated in the fourth section, where the MPMs were derived from substituting fifteen distance measures in the conventional TOPSIS method. The performance of the synthesized mineral favorability maps is also discussed at the section fourth. Of note is that some portions of the sections two may bear slightly resemblance to the first part of distance families in the MPM by Abedi (2020) to better assist readers about the geological setting and geospatial datasets of the studied region.

Methodology

The equations of fifteen distance measures selected from three general families of the squared L_2 , Shannon and combination were described concisely in the former subsection. Then, the stages of implementing the conventional TOPSIS outranking method were explained in the later subsection at eight steps, while the distance families must be incorporated in the sixth stage to generate new mineral potential maps.

Squared L_2 , Shannon and combination distance families

Several retrospective studies of the distance measures have been developed to calculate quantitatively how far apart two vectors are in different fields of studies such as feature

selection, pattern recognition, image processing, data mining, MCDM problems, clustering, and so on (Cha & Srihari, 2002; Kumar et al., 2014; Bora & Gupta, 2014; Irani et al., 2016; Anandan & Uthra, 2017a, b). The importance of the distance measures in the MPM has been studied by Abedi (2020), where superiority of some ones were proved over the Euclidian one in geospatial dataset integration. In the following section, new distance families were explained to be incorporated as the kernel of the TOPSIS method for synthesizing various geospatial indicators.

To better explain distance measures in consistency with previous work (Abedi, 2020), uppercase letters present vectors (e.g. A, B, ...), lowercase ones with a subscript are the vector elements (e.g. a_j, b_j, \dots), and m denotes the length of the vector.

Several distance measures constitute the squared L_2 or χ^2 families, where the most reputed ones are the squared Euclidean, Pearson, Neyman, squared, Divergence, Clark and additive symmetric, listed respectively in Table 1 as Eqs. 1-7. Indeed, they embody the squared Euclidean distance (d_{SE}) in their formulation. Among these distances the Pearson and Neyman have asymmetric property ($d(A, B) \neq d(B, A)$), but the rest ones are symmetric.

The Shannon's entropy distance family was indeed derived from the concept of probabilistic uncertainty or "entropy" which was defined as $H(A) = \sum_{j=1}^m a_j \ln a_j$ (Shannon, 1948). The most popular ones are the Kullback-Leibler, Jeffreys, K divergence, Topsøe and Jensen difference distances formulated as Eqs. 8-12 in the Table 1. The Kullback-Leibler distance also known as relative entropy or information deviation, computes the difference between two non-negative vectors (Kullback & Leibler, 1951; Kocher & Savoy, 2017), which has asymmetric property like the K divergence. The other distances listed in this group are symmetric. The likelihood ratio is defined as a_j/b_j in this family (Deza & Deza, 2006).

The last group of the distance measures is categorized as the combination family, where multiple distance measures are combined to define more appropriate distances as Taneja, Kumar-Johnson and average. Equations 13-15 presented their formulations respectively in Table 1 (Cha, 2007). The Taneja distance (Taneja, 1995) employed both arithmetic and geometric means in Eq. 13. Symmetric squared, arithmetic and geometric mean divergence were used in the Kumar-Johnson distance (Kumar & Johnson, 2005). The average of city block (L_1) and Chebyshev (L_∞) distances was also used to introduce the average distance as Eq. 15.

Two dimensions (2D) vector space was assumed as an example to visually present these distances. Imagine that a 2D spaces axes was ranging from 0 to 10 in both directions. A point at center of the 2D space (positioning at $x = 5$ and $y = 5$) was selected, and then the distance to all other points in $[0, 10] \times [0, 10]$ domain was calculated according to all the distance measures listed in the Table 1. The plots shown in Fig.1 indicated the distances from that point, where blue areas are pertained to the points with the lowest distance (i.e. the highest similarity), yellow portions for more distant points, and red regions for the farthest ones with the least similarity.

TOPSIS outranking method

The TOPSIS method, introduced by Hwang and Yoon in 1981, works as a simple ranking method in terms of its conceptual model, which later has come up with a modification in different variants by many scholars to ameliorate its efficiency in prioritizing alternatives (e.g. Chen & Hwang, 1992; Tavana & Hatami-Marbini, 2011). The wide range of the TOPSIS implementation was categorized by Behzadian et al. (2012) into different fields such as "(1) supply chain management and logistics problems, (2) design and manufacturing, (3) business and marketing, (4) health, safety and environment studies, (5) human resources, (6) energy management, (7) engineering studies (e.g. chemical, industrial, metallurgy, mining, and so on), (8) water resources, and (9) other fields". Whereby this technique was widely applied individually or in combination with other approaches in the past decades with satisfactory results to tackle the MCDM problems.

Table 1. Various distance measure families used as the kernel of the TOPSIS outranking method

Squared L₂ family or χ^2 family	
Squared Euclidean	$d_{SE}(A, B) = \sum_{j=1}^m (a_j - b_j)^2 \quad (1)$
Pearson	$d_P(A, B) = \sum_{j=1}^m \frac{(a_j - b_j)^2}{b_j} \quad (2)$
Neyman	$d_N(A, B) = \sum_{j=1}^m \frac{(a_j - b_j)^2}{a_j} \quad (3)$
Squared	$d_{Sq}(A, B) = \sum_{j=1}^m \frac{(a_j - b_j)^2}{a_j + b_j} \quad (4)$
Divergence	$d_{Div}(A, B) = 2 \sum_{j=1}^m \frac{(a_j - b_j)^2}{(a_j + b_j)^2} \quad (5)$
Clark	$d_C(A, B) = \sqrt{\sum_{j=1}^m \left(\frac{ a_j - b_j }{a_j + b_j} \right)^2} \quad (6)$
Additive Symmetric	$d_{AS}(A, B) = \sum_{j=1}^m \frac{(a_j - b_j)^2 (a_j + b_j)}{a_j b_j} \quad (7)$
Shannon's entropy family	
Kullback-Leibler	$d_{KL}(A, B) = \sum_{j=1}^m a_j \ln \frac{a_j}{b_j} \quad (8)$
Jeffreys	$d_{Jef}(A, B) = \sum_{j=1}^m (a_j - b_j) \ln \frac{a_j}{b_j} \quad (9)$
K divergence	$d_{Kd}(A, B) = \sum_{j=1}^m a_j \ln \frac{2a_j}{a_j + b_j} \quad (10)$
Topsøe	$d_{Top}(A, B) = \sum_{j=1}^m \left(a_j \ln \left(\frac{2a_j}{a_j + b_j} \right) + b_j \ln \left(\frac{2b_j}{a_j + b_j} \right) \right) \quad (11)$
Jensen difference	$d_{JD}(A, B) = \sum_{j=1}^m \left[\frac{a_j \ln a_j + b_j \ln b_j}{2} - \left(\frac{a_j + b_j}{2} \right) \ln \left(\frac{a_j + b_j}{2} \right) \right] \quad (12)$
Combination family	
Taneja	$d_T(A, B) = \sum_{j=1}^m \left[\left(\frac{a_j + b_j}{2} \right) \ln \left(\frac{a_j + b_j}{2\sqrt{a_j b_j}} \right) \right] \quad (13)$
Kumar-Johnson	$d_{KJ}(A, B) = \sum_{j=1}^m \left[\frac{(a_j^2 - b_j^2)^2}{2(a_j b_j)^{3/2}} \right] \quad (14)$
Avg (L₁, L_∞)	$d_{Avg}(A, B) = \frac{1}{2} \sum_{j=1}^m a_j - b_j + \max_j a_j - b_j \quad (15)$

It operates on the basis of a predefined Euclidian distance measure, which endeavors to pick up the alternatives that simultaneously have the shortest distance from the positive ideal solution (PIS) and the farthest distance from the negative ideal solution (NIS) for each criterion in the decision matrix of the MCDM/MADM problems. The PIS maximizes the benefit criteria or minimizes the cost criteria, whereas the NIS maximizes the cost criteria or minimizes the benefit criteria (Behzadian et al., 2012). Here, the precise formulations of the conventional TOPSIS method as the most popular version were explained successively in eight steps.

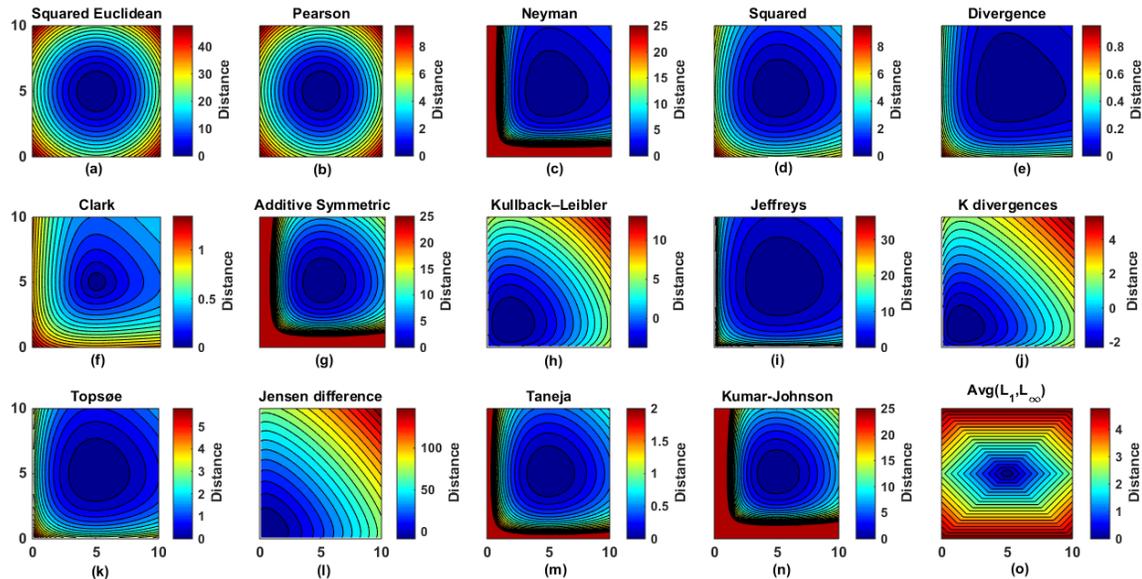


Figure 1. The calculated distances (listed in Table 1) for a synthetic example in 2D space at a point (5, 5)

Before describing the stepwise procedure of the algorithm, let us begin by assuming that $A_i (i = 1, 2, \dots, n)$ and $C_j (j = 1, 2, \dots, m)$ are respectively a set of n alternatives (or the data points in the MPM) and m criteria (mineral indicators). The stages of the TOPSIS method for MPM can be listed in some steps as following (Abedi & Norouzi, 2016),

Step 1: It is required to construct a decision matrix from the geospatial datasets in any MCDM problem, while their matrix elements $X = (x_{ij})_{n \times m}$ are assigned appropriated scores to each alternative i on each criterion j by a group of geoscientist decision makers (GDMs). The size of the decision matrix depends on the number of the data points and geospatial indicators derived from exploratory criteria (e.g. geology, geochemistry, remote sensing, and geophysics surveys).

Step 2: The importance of each geospatial indicator should be determined by assigning a weight (w_j), while it is usually required to differentiate between various indicators. Several knowledge- or data-driven techniques (e.g. Delphi, AHP, Fuzzy AHP, ANP, data-based weighting and so on) have been developed for such task (Tzeng & Huang, 2011; Ishizaka & Nemery, 2013; Abedi et al., 2013; Yousefi & Carranza, 2016) such that;

$$\sum_{j=1}^m w_j = 1, \quad j = 1, 2, \dots, m. \tag{16}$$

Step 3: Since the scaling effects may disturb the MPM result, it is better to tackle such issue by normalizing the decision matrix (r_{ij}) through,

$$r_{ij} = x_{ij} / (\sum_{p=1}^n x_{pj}^2)^{0.5}, \quad i = 1, 2, \dots, n \quad \& \quad j = 1, 2, \dots, m \tag{17}$$

Step 4: The weighted r_{ij} matrix of V_{ij} is calculated in the conventional TOPSIS method as following,

$$V_{ij} = w_j r_{ij}, \quad i = 1, 2, \dots, n \quad \& \quad j = 1, 2, \dots, m \tag{18}$$

Step 5: The PIS and NIS values are determined for each criterion j , respectively, as the best and worst alternatives (or as origin point for calculation of the distance measure),

$$f_j^+ = (v_1^+, v_2^+, \dots, v_j^+, \dots, v_m^+) = \left\{ \left(\max_i \{v_{ij}\} | j \in B \right), \left(\min_i \{v_{ij}\} | j \in C \right) \right\} \tag{19}$$

$$f_j^- = (v_1^-, v_2^-, \dots, v_j^-, \dots, v_m^-) = \left\{ \left(\min\{v_{ij}\} | j \in B \right), \left(\max\{v_{ij}\} | j \in C \right) \right\} \quad (20)$$

Step 6: The Euclidean distance measure computes the separation $M = (S_i^+, S_i^-)$ through the following equations,

$$S_i^+ = \left\{ \sum_{j=1}^m (v_{ij} - v_j^+)^2 \right\}^{0.5}; \quad i = 1, \dots, n \quad (21)$$

$$S_i^- = \left\{ \sum_{j=1}^m (v_{ij} - v_j^-)^2 \right\}^{0.5}; \quad i = 1, \dots, n \quad (22)$$

Of note is that this step has substantial impact on prioritizing the alternatives in the MCDM problem, while the Euclidian distance was chosen as the default kernel in the formulation of conventional and new variants of the TOPSIS method. Incorporating new distance families have been triggered by Abedi (2020) as an avalanche of new developments in TOPSIS method, which helped MPM improvement. Here, the aim is to investigate the applicability of three new families of the distance measures listed in Table 1, which can pave the way for the discovery of new families of distance measures in MPM.

Step 7: The relative closeness coefficient of the alternatives to the ideal solution is subsequently determined as following,

$$T_i^C = \frac{S_i^-}{S_i^+ + S_i^-}; \quad i = 1, \dots, n \quad (23)$$

The set of alternatives (or favorable areas highlighted by the MPM) can be fully ranked according to the descending order of the relative closeness coefficient (T_i^C).

Step 8: The T_i^C values are projected at an interval of [0,1] to normalize the MPM output through Eq. (24),

$$M_i^T = \frac{T_i^C - \min(T_i^C)}{\max(T_i^C) - \min(T_i^C)}; \quad i = 1, \dots, n \quad \& \quad 0 \leq M_i^T \leq 1 \quad (24)$$

where high and low values of the normalized M_i^T are in association with the high and low favorable mineralized zones, respectively (Abedi & Norouzi, 2016).

Background geology and geospatial indicators

This section describes the geological descriptions of the North Narbaghi copper deposit situated in the Saveh prospecting zones, and a summary of geospatial indicators derived from a multi-disciplinary exploratory survey.

Geological survey

A north dipping subduction of the Neo-Tethys Ocean, begun in the Mesozoic era, has affected the Iranian plateau (Stöcklin, 1968; Berberian & Berberian, 1981). Upon maturing the subduction zone and overlying continental magmatic arc, a wide belt consisting of Cenozoic plutonic and volcanic units arose from intense igneous activities- that is known as the Urumieh-Dokhtar magmatic assemblage zone (UDMA) in the Iran structural geology divisions map (Fig. 2a). Such subduction led to generation of the UDMA as a distinct, linear intrusive-extrusive magmatic complex, which has been aligned parallel and between the magmatic-metamorphic Sanandaj-Sirjan zone (SSZ) and the central Iran domain from the NW to SE (Nouri et al., 2018). The constituents of the UDMA zone comprise of thick sequence of volcanic and pyroclastic units with expansion from Sahand at the NW to Bazman at the SE of Iran. Those igneous activities produced these rock units can be traced all along the Alpine-Himalayan orogenic belt (Ghorbani, 2004). Taking the magmatic and mineralogical characteristics especially the copper

into consideration, the UDMA zone has been divided into three districts as southern (Kerman), central (Taft, Anar, and Kashan-Qom), and northern (Tafresh-Takab) portions.

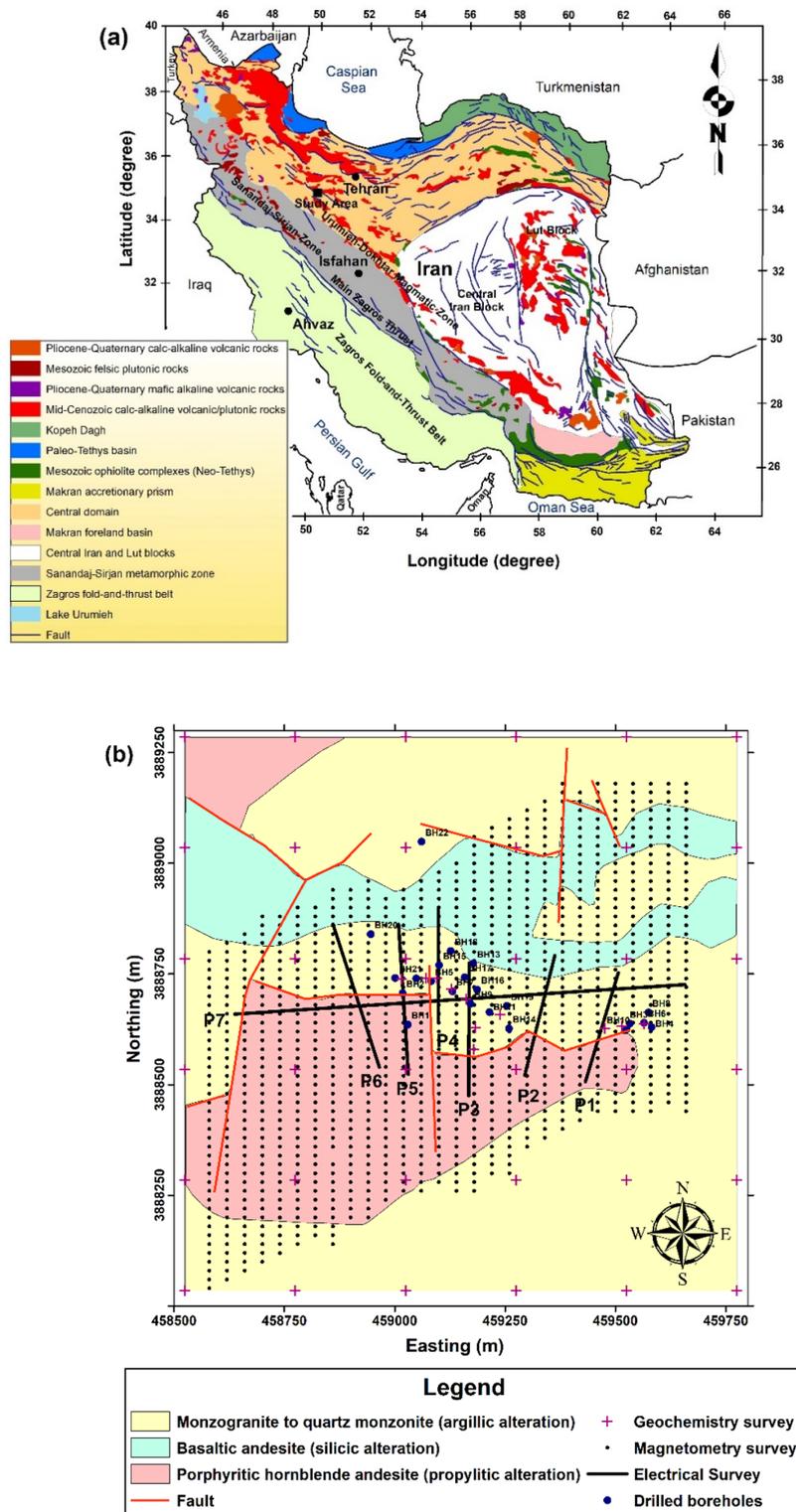


Figure 2. The geological setting of Iran reproduced from Richards et al., 2006 (a), and the simplified geological rock types in the North Narbaghi copper deposit (b)

The intrusive igneous rocks of this zone are characteristically composed of diorites and granites, whereas the extrusive rocks are mostly dominated by basaltic-rhyolitic units. It is worth noting that there are enormous deposits of porphyry-type ore bearing mineralization, magmatic iron, epithermal gold, volcanogenic manganese, and hydrothermal barite, lead and zinc (Ghorbani, 2002, 2004, 2013). This ore-bearing zone defined as a ~50-100 km-wide belt dominated by an Andean-type magmatic arc created at the crust of the Central Iranian Micro-Continent (CIMC) structural unit. The UDMA (also known with other names of the Sahand-Bazman or Tabriz-Bazman zone) is distinguished by the Cenozoic extrusive and intrusive units with an age of Eocene-Quaternary along with the associated volcanoclastic rocks. Intrusive magmatic units in the UDMA comprise of the subvolcanic porphyritic granitoid units of diorite, granite, granodiorite and tonalite rocks (Shahabpour, 2005; Kazemi et al., 2018).

The Saveh prospecting zone, studied region in this work (Fig. 2a), is located at the UDMA zone, as the main host of the porphyry deposits such as Cu, Au and Mo in Iran (Berberian & King, 1981, Rezaei et al., 2015). Reports by the National Iranian Copper Industries Company (NICICO) stated that exploited porphyry deposits over the UDMA zone in terms of economically value contain Cu grades ranging often from 0.15 to 0.8% (Yousefi & Carranza, 2016). In the central segment of the UDMA (Saveh-Yazd porphyry copper belt), porphyry copper deposits formed during middle Miocene time (17-15 Ma), where the well-known porphyry copper deposits in the region include Dalli, Kahang, Ali Abad, and Darreh Zereshk (Aghazadeh et al., 2015). Magmatic intrusions and activities have frequently occurred in Paleogene, where the volcanic rock units in the Saveh area are thicker than 4 km, and involve pyroclastic sequences, lava flows, tuff and ignimbrites (Stöcklin, 1968; Berberian & Berberian, 1981; Alavi, 2007).

The prospect zone in the North Narbaghi copper deposit in the Saveh area indicated a volcano-genetic type of mineralization, located on the volcanic belt of the UDMA zone. The main rock units dominated the studied area have been summarized in three groups as following (Fig. 2b),

(1) Monzogranite to quartz monzonite units hosted the most Cu mineralized regions, severely dominated by argillic alteration (Fig. 3a). The number of magmatic intrusions in the region is partly high, where most portions were exposed in small sizes. The age of these masses is equivalent to the early Oligocene. The structural lineaments of the region have substantially controlled the deployment of such magmatic intrusions since most intrusive sources in adjacency to the Saveh were manifested on the margins of the lineaments. Therefore, such phenomena could be in association with a batholith source as the feeder of the intrusive masses in this region. Two types of hydrothermally alteration were visible in this unit. Phyllic alteration occurred in portions with minerals of pyrite, sericite and quartz. In addition, over areas with depletion of the Cu mineralization, the argillic alteration has severely affected the rocks. Meanwhile, the intrusive units were intact and unaltered in some portions as well.

(2) Basaltic andesite rocks with a distinct silicic alteration. Volcanic activities in the North Narbaghi copper deposit led to the generation of the basaltic andesite unit within the porphyritic hornblende andesite rocks (Fig. 3b), with a dark gray color and a distinct outcrop than the surrounding rocks. These rocks were mostly surrounded by the monzogranite and quartz monzodiorite units.

(3) Porphyritic hornblende andesite units affected by the hydrothermally propylitic alteration have surrounded the mineralized Cu zones (Fig. 3c). Intense alteration of coarse crystals namely chlorite, epidote and carbonate has appeared in the andesite unit, leading to the transformation of plagioclase and amphibole into such minerals. Moreover, the andesine rock as the oldest unit with an Eocene age and the most extensive rock occurred in the south of the North Narbaghi. Chlorite alteration was also sporadically visible in some portions and its intensity has increased in adjacency to the mineralized regions (Ghalamghash, 1998; Ramazi & Jalali, 2015).

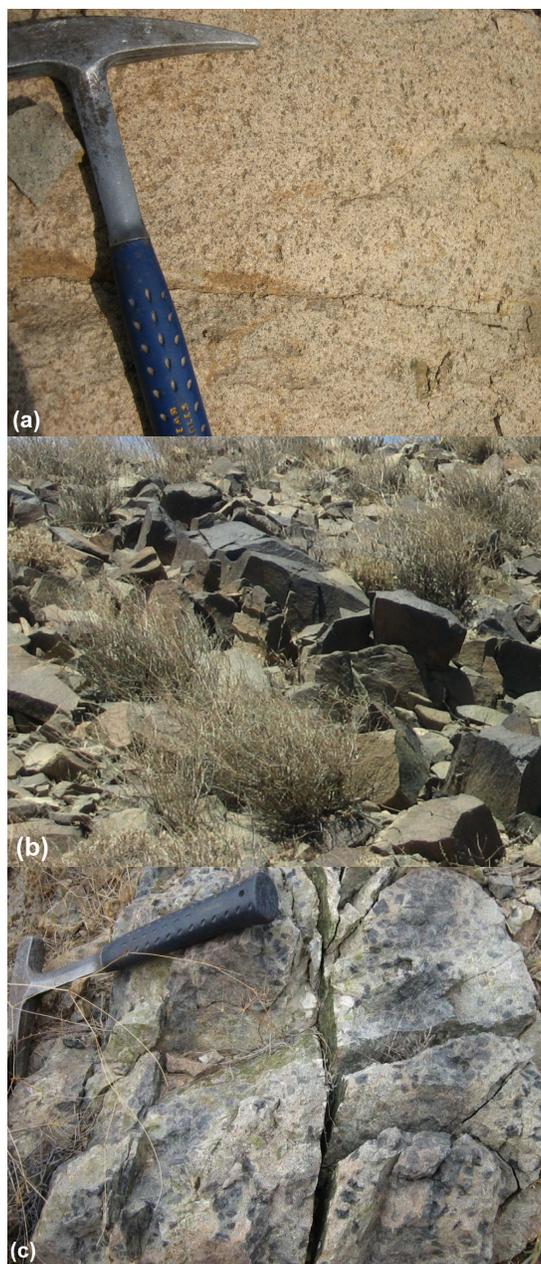


Figure 3. Sample photos of the main rock units in the North Narbaghi, (a) monzogranite to quartz monzonite, (b) basaltic andesite, and (c) porphyritic hornblende andesite (Dehghan Nayeri, 2018)

Structural lineaments (i.e. faults, fractures and contacts) with few traces in the area had low effect on the Cu mineralization. The disseminated type of the Cu mineralization was not in association with the structural lineaments, while the effect of the faults just led to partially Cu enrichment. The largest fault trace was observed in the west portion of the studied region with an approximate north-south trend (Dehghan Nayeri, 2018).

Geospatial indicators

The procedure of extracting seven indicator layers derived from the geological, geochemical and geophysical data have been described in previous work with details (Abedi, 2020). The

MCDM problem was casted in a decision matrix ($X_{9 \times 136 \times 7}$) such that assigning a priority score to each alternative on each criterion was on the basis of a group of GDMs mostly from the geological survey of Iran (GSI). Each indicator layer was scored in a ranging from 0 to 1 which suppresses the scaling effects perturbing the MPM. Figure 4 has portrayed all geospatial indicators. These indicators are rock type, alteration, lithochemical Cu concentration, the main multi-element geochemical principal component, magnetic property, electrical chargeability and resistivity, respectively. Indicators highlight the central portions of the area as favorability regions for probable Cu mineralization.

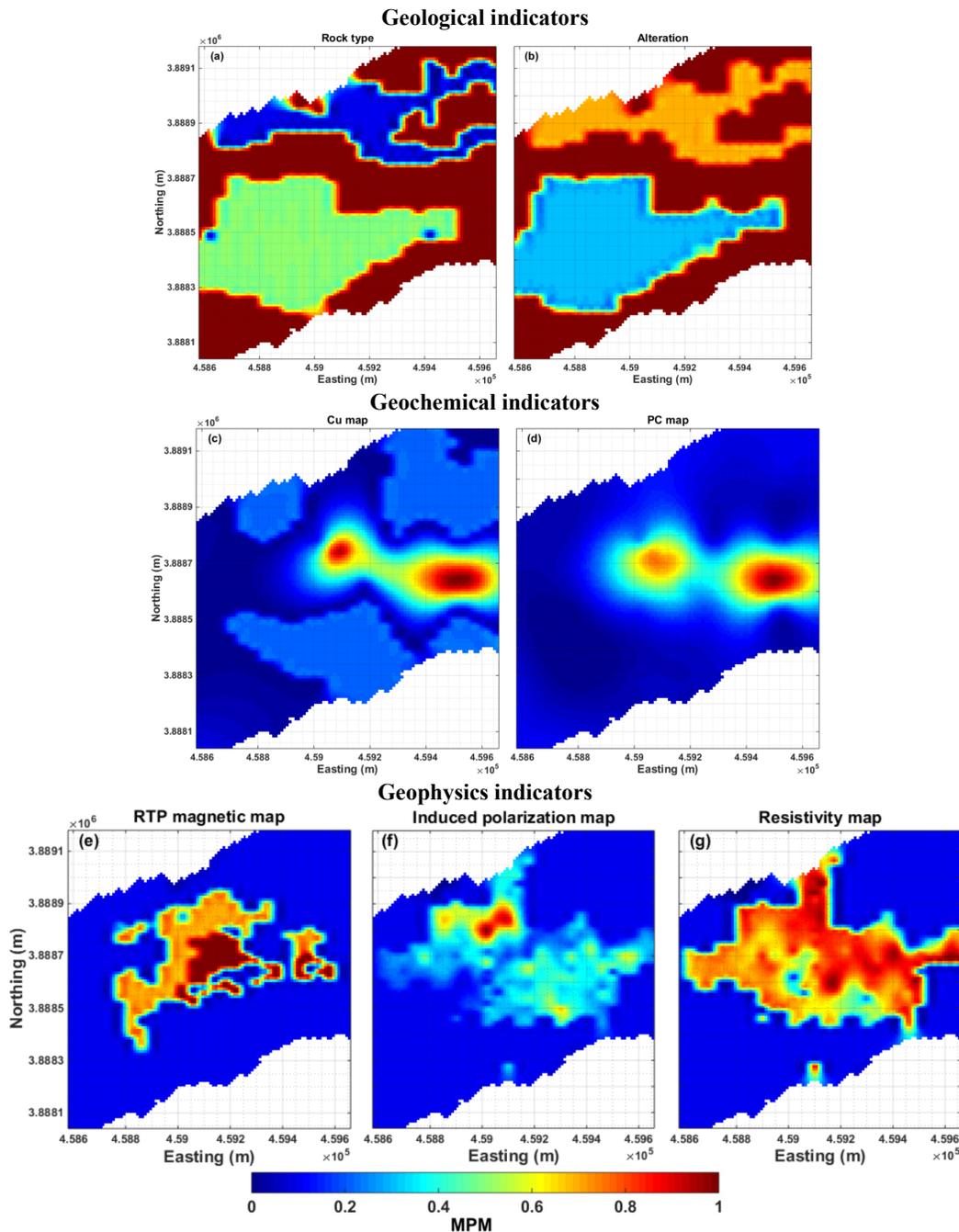


Figure 4. Geospatial indicators, (a) rock type, (b) alteration, (c) Cu concentration, (d) multi-element geochemical component, (e) magnetic, (f) electrical chargeability, and (g) electrical resistivity (Abedi, 2020)

The Delphi method as a popular knowledge-driven weighting method was employed to assign the weights of the indicators (Linstone et al., 1975; Gupta & Clarke, 1996; Skulmoski et al., 2007; Rezapour et al., 2020). To simplify the procedure of determining sub-optimal weights of the indicator layers and criteria, a sample sheet like Table 2 was distributed among all GDMs to fill it by assigning appropriate weight ranging from 0 to 1 to each indicator on the basis of their knowledge in the porphyry Cu exploration. Averaging the assigned weights by the GDMs such that the summation gets one ($\sum_{j=1}^m w_j = 1$), Table 2 has summarized the normalized weigh, where the highest weight equal to 0.27 was assigned to the lithochemical Cu concentration map (Fig. 4c). All indicators were transferred to the raster cells with dimensions of 10 m in both easting and northing directions, while the final geospatial datasets include a database of 9136 rows and 7 indicator columns.

Outranking geospatial indicators

Geospatial decision matrix consisted of three main indicator criteria of the geology (surface studies), geophysics (magnetometry and electrical surveys), and geochemistry. A group of GDMs in the field of porphyry Cu exploration from the GSI with various disciplines in mineral exploration was gathered together to guide the MPM. Seven indicator layers were inserted in the X matrix to implement the TOPSIS method with various distance measures.

The MPMs were produced by the TOPSIS outranking method, in which the conventional Euclidian distance kernel was replaced by three families of distance measures presented in the Table 1. Figure 5 has presented fifteen MPMs, while on most of them a distinct ribbon at the center of each map was manifested. Such promising potential zone located at the monzogranite to quartz monzonite units. The scatter plots of all MPMs versus the conventional distance measure (i.e. the Euclidian), were shown in Fig. 6. It is obvious that each distance measure can yield different potential map by substituting the Euclidian kernel. Hence, the most appropriate ones must be searched to find out the efficient ones in the MPM.

The central portions of the North Narbaghi copper mineralization were drilled by 21 vertically boreholes to investigate its mining prospectivity. The layout map and productivity index of each drilled borehole were plotted in Figs.7a and 7b, respectively, to evaluate the efficiency of the produced maps, and subsequently to search the best output(s). The productivity value was calculated from multiplying Cu concentration (in ppm unit) by its ore thickness (in meter) along each drilled borehole, finally being normalized by the total length of borehole. Indeed, the productivity index has presented the average of the Cu grade along the borehole. Boreholes 10, 13 and 21 were excluded in the MPM efficiency analysis owing to their high uncertainty in grade analysis. Figure 2b has indicated borehole locations.

The scatter plots of the productivities versus the MPM amounts extracted at the locations of drilling were plotted in Fig. 8 for fifteen MPMs. The Pearson's linear correlation coefficient ($\rho_{Pearson}$) for the fitted linear curve was also calculated to evaluate the performance of each distance measure.

Table 2. The normalized weight of indicators through a Delphi analysis by the expert GDMs (Rezapour et al. 2020)

Criteria	Weight	Sub-criteria	Weight	Indicator	Weight	Final weights
Geology	0.25	Surface survey	1	Rock type	0.50	0.12500
				Alteration	0.50	0.12500
Geochemistry	0.45	Lithochemical	1	Cu concentration	0.60	0.27000
				Main component	0.40	0.18000
Geophysics	0.3	Magnetic	0.35	Magnetic	1.00	0.10500
		Electric	0.65	Resistivity	0.45	0.08775
				Chargeability	0.55	0.10725

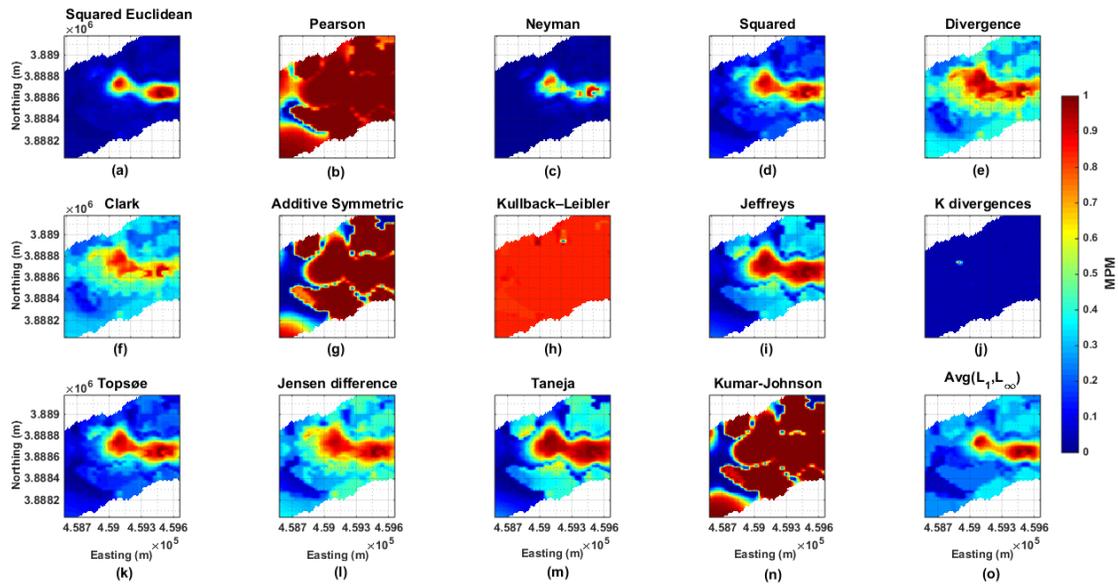


Figure 5. Mineral potential maps through implementing the TOPSIS outranking method with 15 distance measures as the kernel of the algorithm presented in the Table 1. The outputs were normalized at an interval of $[0,1]$, where higher values corresponded to the favorable zones

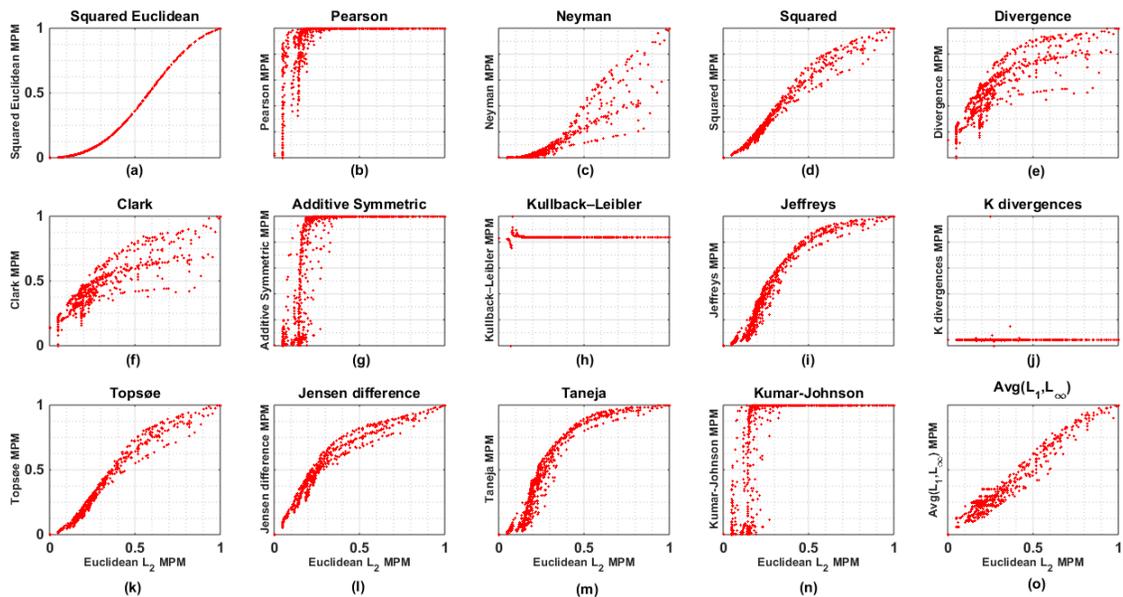


Figure 6. The scatter plots of the 15 MPMs versus the conventional TOPSIS method with the Euclidean distance kernel

It is evident that positive correlation must happen when the MPM output is in consistency with the mineralized zones. Among 15 distances, the TOPSIS outranking via the squared Euclidean measure from the squared L_2 family produced the highest correlation equal to 0.4647 (Figs. 5a, 6a and 8a), which indeed had higher efficiency in detection of the Cu mineralized zones. Meanwhile, the Kullback-Leibler and K divergence generated negative correlation coefficients which led to wrong potential maps. Compared to the Euclidian L_2 distance measure as the default measure in the conventional TOPSIS method with a correlation of 0.4533 (Abedi, 2020), just squared Euclidean and Jensen difference (0.4584) had higher correlation (Figs. 8a and 8l). The squared (0.4116) and average (0.4482) distances had a bit smaller coefficients, respectively shown in Figs. 8d and 8o.

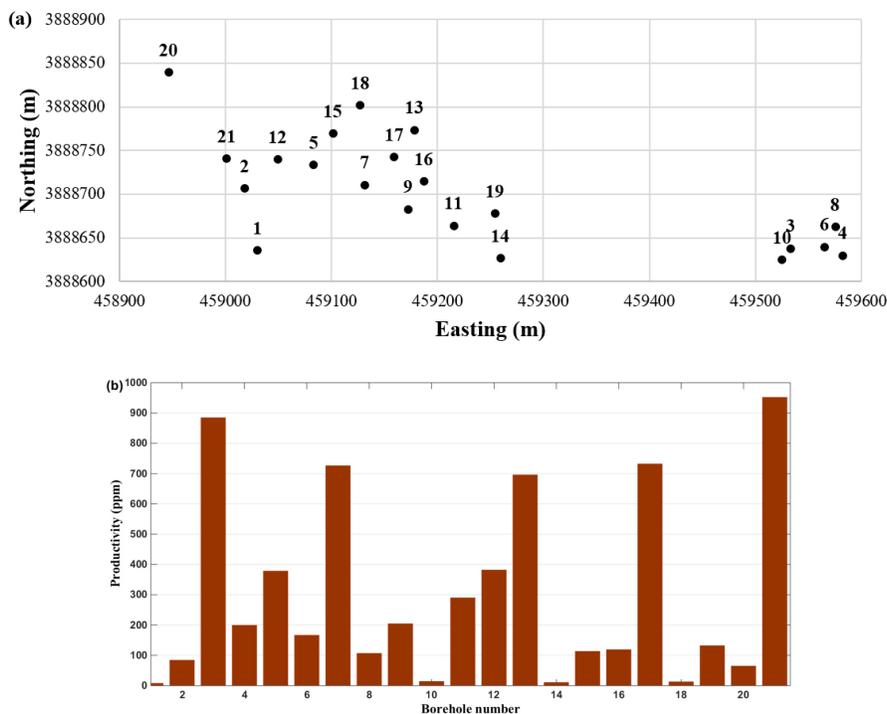


Figure 7. Location map (a) and productivity values of 21 drilled holes (b)

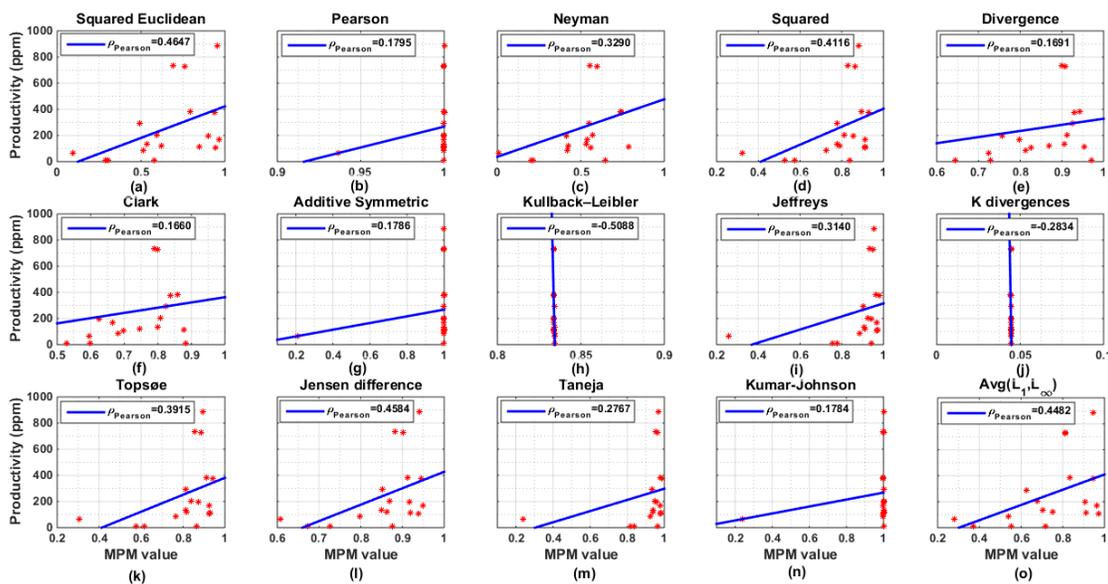


Figure 8. Evaluation of the MPMs by 15 distance measures with the productivity values acquired from the drilling results. The Pearson’s linear correlation coefficient was calculated for each distance measure, where the squared Euclidean distance generated the highest efficiency in the porphyry Cu potential mapping

Since nonlinear correlation may occur for the generated MPMs versus the productivity values at locations of the drilled boreholes, the Spearman’s rank correlation coefficient was also calculated. Figure 9 has presented both the Pearson and Spearman correlation coefficient values for each distance measure, where the squared Euclidian (0.5583), squared (0.5026), Topsøe (0.5026), Jensen difference (0.5624), and Avg (L_1, L_∞) (0.5046) had similarity compared to the Euclidian distance (0.5748). The Kullback-Leibler measures generated the lowest Spearman’s

correlation coefficient value equal to -0.6061. Therefore, it can be inferred that various distance measures should be examined when they are used as the kernel of the MCDM techniques in the MPM procedure. The higher correlation coefficient, the better efficiency in the MPM would be generated. Such research can be extended for regional scale potential mapping where the locations of known deposits are used instead of drilling results to evaluate the performance of the distance measures.

Discussion

To better indicate the importance of the MPM in copper exploration, 3D visualization of the Cu grade for a cut-off value of 500 ppm was shown in Fig. 10. The best map generated by the squared Euclidian distance measure was also superimposed on the plot, where close consistency between the Cu mineralization and the potential map was evident. Since produced MPMs in this work were all 2D for a deposit-scale prospect zone, the importance of drilled boreholes through calculation of the productivity index was projected and mapped on the surface.

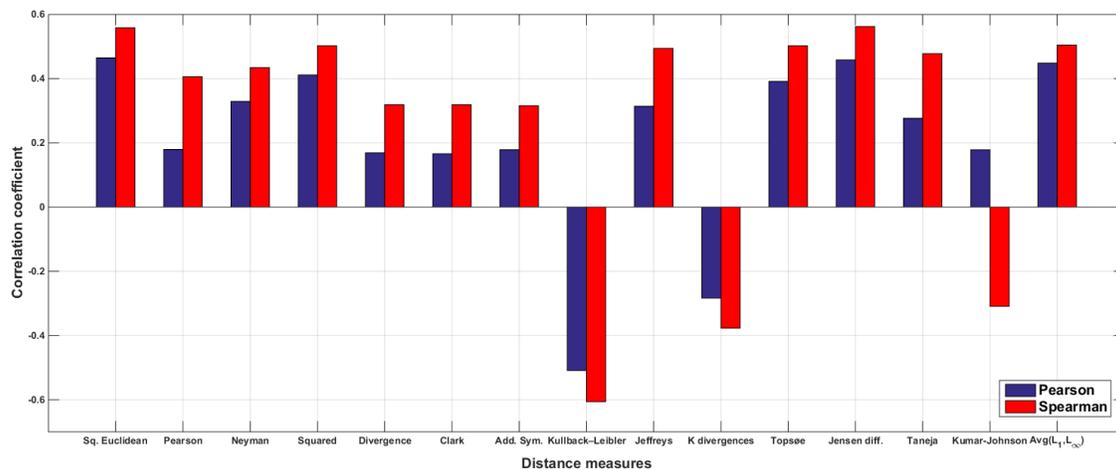


Figure 9. The Pearson's linear correlation coefficient along with the Spearman's rank correlation coefficient were calculated for each distance measure

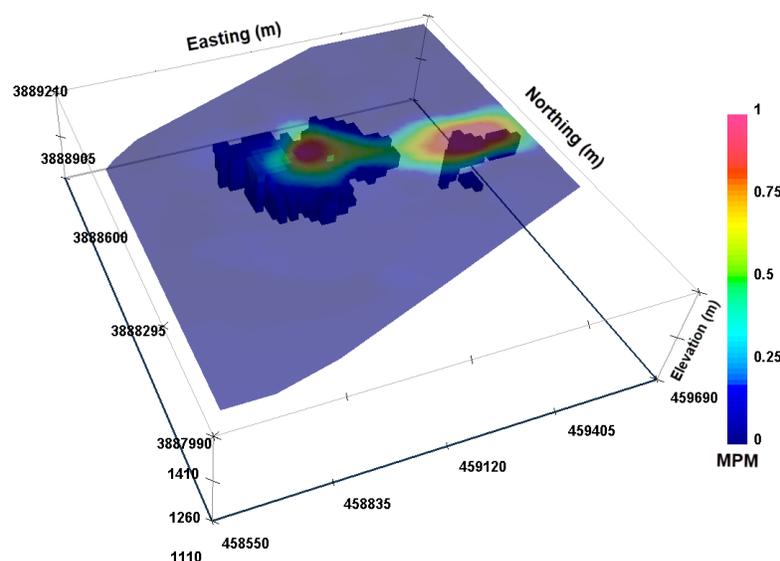


Figure 10. The 3D model of the Cu mineralization for a cut-off value of 500 ppm superimposed by the MPM generated by the squared Euclidian distance

Indeed, it facilitated the evaluation of the performance of all MPMs by comparison to the drilling results. The point should be noted was that some indicator layers such as those derived from the geophysical criteria can provide significant information from deep-seated sources. In other words, the effect of blind targets can be manifested on the surface geophysical data. Therefore, 2D MPMs could provide valuable information about deep targets in case of involving geophysical indicators owing to the limited availability of precise 3D geological data in establishing accurate mineral favorability models.

Distance measure plays a decisive role in various algorithms developed for synthesizing indicators in MPM. In addition to the TOPSIS method, most outranking algorithms are using this measure to rank favorable mineralized zones (Abedi et al., 2017, 2016). Of special interest is the utilization of the data-driven methods like clustering algorithms which incorporate this measure for multivariate data analysis, and reveal significant differences in the results (Pandit & Gupta, 2011; Salarpour & Khotanlou, 2018). Thus, it is of considerable importance that deserves more investigation in the future of MPM.

Conclusion

Three well-known distance measure families of the squared L_2 , Shannon's entropy, and combination were utilized as the kernel of the conventional TOPSIS outranking approach in the mineral potential mapping. Among these families, fifteen distances were selected to integrate seven indicator layers derived from a case study in porphyry copper exploration, in the Saveh, Iran. The geospatial datasets comprising of the geological, geophysical and geochemical data were processed to evaluate those distances in Cu potential mapping. Some distances compared to the Euclidian distance as the default of the conventional TOPSIS method, had superiority in generation of potential maps. Therefore, optimum distance measure should be searched in any MCDM problem, when a distance plays a decisive role as the kernel of the decision making methodology.

Acknowledgements

I would like to express my sincere thanks to the School of Mining Engineering and the Geological Survey of Iran (GSI) for their supports, providing data, and making a decision team. I thank the Editor-in-Chief of the Journal of Geopersia, Prof. Ghasemi-Nejad, and anonymous referees for reviewing the paper precisely and patiently and for their constructive and valuable comments, which helped us to improve the quality of this work.

References

- Abedi, M., 2020. Non-Euclidean distance measures in spatial data decision analysis: investigations for mineral potential mapping. *Annals of Operations Research*, <https://doi.org/10.1007/s10479-020-03681-x>.
- Abedi, M., Mostafavi Kashani, S.B., Norouzi, G.H., Yousefi, M., 2017. A deposit scale mineral prospectivity analysis: A comparison of various knowledge-driven approaches for porphyry copper targeting in Seridune, Iran. *Journal of African Earth Sciences*, 128: 127-146.
- Abedi, M., Mohammadi, R., Norouzi, G.H., Mir Mohammadi, M.S., 2016. A comprehensive VIKOR method for integration of various exploratory data in mineral potential mapping. *Arab J Geosci.*, 9: 482, DOI 10.1007/s12517-016-2512-9.
- Abedi, M., Norouzi, G.H., 2016. A general framework of TOPSIS method for integration of airborne geophysics, satellite imagery, geochemical and geological data. *International Journal of Applied Earth Observation and Geoinformation*, 46: 31-44.
- Abedi, M., Norouzi, G.H., Fathianpour, N., 2015. Mineral potential mapping in Central Iran using fuzzy ordered weighted averaging method. *Geophys. Prospect.*, 63: 461-477.

- 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: 145-164.
- Aghazadeh, M., Hou, Z., Badrzadeh, Z., Zhou, L., 2015. Temporal-spatial distribution and tectonic setting of porphyry copper deposits in Iran: Constraints from zircon U-Pb and molybdenite Re-Os geochronology. *Ore Geology Reviews*, 70: 385-406.
- Alavi, M., 2007. Structures of the Zagros fold thrust belt in Iran. *Am. J. Sci.*, 307: 1064-1095.
- Anandan, V., Uthra, G., 2017a. Extension of TOPSIS using L1 Family of Distance Measures. *Advances in Fuzzy Mathematics*, 12: 897-908.
- Anandan, V., Uthra, G., 2017b. A Hybrid Approach Integrating AHP and Extended TOPSIS by Tanimoto and Jaccard Distance Measure. *International Journal of Pure and Applied Mathematics*, 117: 145-153.
- Behzadian, M., Otaghsara, S.K., Yazdani, M., Ignatius, J., 2012. A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*, 39: 13051-13069.
- Berberian, F., Berberian, M., 1981. Tectono-plutonic episodes in Iran: Zagros Hindu Kush Himalaya geodynamic evolution, edited by H. K. Gupta and F. M. Delany. *Am. Geophys. Union*, 3: 33-69.
- Berberian, M., King, G.C., 1981. Towards a paleogeography and tectonic evolution of Iran. *Canadian Journal of Earth Sciences*, 18: 210-265.
- Bora, D.J., Gupta, A.K., 2014. Effect of Different Distance Measures on the Performance of K-Means Algorithm: An Experimental Study in Matlab. *International Journal of Computer Science and Information Technologies*, 5: 2501-2506.
- 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: 379-410.
- Cha, S.H., 2007. Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions. *International journal of mathematical models and methods in applied sciences*, 1: 300-307.
- Cha, S.H., Srihari, S.N., 2002. On measuring the distance between histograms. *Pattern Recognition*, 35: 1355-1370.
- Chen, S.J., Hwang, C.L., 1992. *Fuzzy Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, Berlin.
- Dehghan Nayeri, R., 2018. Porphyry copper potential mapping in Narbaghi through TOPSIS multi-criteria decision making method. MSc. Thesis in University of Tehran, Iran (published in Persian).
- Deza, E., Deza, M.M., 2006. *Dictionary of Distances*. ELSEVIER, 391 p.
- Ghalamghash, J., 1998. Report and Geological Map of Saveh 1:100000. Geological Survey of Iran, Tehran, Iran (in Persian).
- Ghorbani, M., 2013. *The Economic Geology of Iran, Mineral Deposits and Natural Resources*. Springer Dordrecht Heidelberg New York London, 569 p.
- Ghorbani, M., 2004. *Volcanology basics with a view on Iran volcanoes*. Pars (arian zamin) geology research center, 356 p.
- Ghorbani, M., 2002. *Economic geology of Iran*. Geological Survey of Iran Publication, 700 p.
- Gupta, U.G., Clarke, R.E., 1996. Theory and applications of the Delphi technique: A bibliography (1975-1994). *Technol. Forecasting Social Change*, 53: 185-211.
- Hwang, C.L., Yoon, K., 1981. *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, New York.
- Irani, J., Pise, N., Phatak, M., 2016. Clustering Techniques and the Similarity Measures used in Clustering: A Survey. *International Journal of Computer Applications*, 134: 9-14.
- Ishizaka, A., Nemery, P., 2013. *Multi-Criteria Decision Analysis Methods and Software*. John Wiley & Sons, Ltd, 296 p.
- Kazemi, K., Kananian, A., Xiao, Y., Sarjoughian, F., 2018. Petrogenesis of Middle-Eocene granitoids and their Mafic microgranular enclaves in central Urmia-Dokhtar Magmatic Arc (Iran): Evidence for interaction between felsic and mafic magmas. *Geoscience Frontiers* (in press).
- Kumar, V., Chhabra, J.K., Kumar, D., 2014. Performance Evaluation of Distance Metrics in the Clustering Algorithms. *NFOCOMP*, 13: 38-51.
- Kullback, S., Leibler, R.A., 1951. On information and sufficiency. *Ann. Math. Statist.*, 22: 79-86.
- Kumar, P., Johnson, A., 2005. On a symmetric divergence measure and information inequalities. *Journal of Inequalities in pure and applied Mathematics*, 6: Art 6, 1-13.

- Kocher, M., Savoy, J., 2017. Distance measures in author profiling. *Information Processing and Management*, 53: 1103-1119.
- Kuo, T., 2017. A modified TOPSIS with a different ranking index. *European Journal of Operational Research*, 260: 152-160.
- Linstone, H.A., Turoff, M., Helmer, O., 1975. *The Delphi method: Techniques and applications*, Addison-Wesley Publishing Company, Advanced Book Program, Boston, MA.
- Nouri, F., Azizi, H., Stern, R.J., Asahara, Y., Khodaparast, S., Madanipour, S., Yamamoto, K., 2018. Zircon U-Pb dating, geochemistry and evolution of the Late Eocene Saveh magmatic complex, central Iran: Partial melts of sub-continental lithospheric mantle and magmatic differentiation. *Lithos*, 314-315: 274-292.
- Opricovic, S., Tzeng, G.H., 2004. Compromise solution by MCDM methods: a comparative analysis of VIKOR and TOPSIS. *Eur. J. Oper. Res.*, 156: 445-455.
- Pandit, S., Gupta, S., 2011. A COMPARATIVE STUDY ON DISTANCE MEASURING APPROACHES FOR CLUSTERING. *International Journal of Research in Computer Science*, 2 (1): 29-31.
- 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 Sci. Info.*, 8: 207-220.
- Pazand, K., Hezarkhani, A., Ataei, M., 2012. Using TOPSIS approaches for predictive porphyry Cu potential mapping: a case study in Ahar-Arasbaran (NW-Iran). *Comput. Geosci.*, 49: 62-71.
- Ramazi, H.R., Jalali, M., 2015. Contribution of geophysical inversion theory and geostatistical simulation to determine geoelectrical anomalies. *Stud. Geophys. Geod.*, 59: 97-112.
- Rezaei, S., Lotfi, M., Afzal, P., Jafari, M.R., Meigoony, M.Sh., Khalajmasoumi, M., 2015. Investigation of copper and gold prospects using index overlay integration method and multifractal modeling in Saveh 1:100,000 sheet, Central Iran. *Gospodarka Surowcami Mineralnymi - Mineral Resources Management*, 31: 51-74.
- Rezapour, M.J., Abedi, M., Bahroudi, A., Rahimi, H., 2020. A clustering approach for mineral potential mapping: A deposit-scale porphyry copper exploration targeting. *Geopersia*, 10 (1): 149-163.
- Richards, J., Wilkinson, D., Ullrich, T., 2006. Geology of the Sari Gunay epithermal gold deposit, Northwest Iran. *Econ. Geol.*, 101: 1455-1496.
- Salarpour, A., Khotanlou, H., 2018. An Empirical Comparison of Distance Measures for Multivariate Time Series Clustering. *IJE TRANSACTIONS B: Applications*, 31 (2): 250-262.
- Shahabpour, J., 2005. Tectonic evolution of the orogenic belt in the region located between Kerman and Neyriz. *Journal of Asian Earth Sciences*, 24: 405-417.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Tech. J.*, 27: 379-423, 623-656.
- Skulmoski, G.J., Hartman, F.T., Krahn, J., 2007. The Delphi method for graduate research. *Journal of Information Technology Education*, 6: 1-21.
- Stöcklin, J., 1968. Structural history and tectonics of Iran: a review. *Am. Assoc. Pet. Geol., Bull.*, 52: 1229-1258.
- Taneja, I.J., 1995. *New Developments in Generalized Information Measures*, Chapter in: *Advances in Imaging and Electron Physics*, Ed. P.W. Hawkes, 91, 37-135.
- Tavana, M., Hatami-Marbini, A., 2011. A group AHP-TOPSIS framework for human spaceflight mission planning at NASA. *Expert Syst. Appl.*, 38: 13588-13603.
- Tzeng, G.H., Huang, J.J., 2011. *Multiple Attribute Decision Making Methods and applications*. Taylor & Francis Group, LLC: 335 p.
- Yousefi, M., Nykänen, V., 2017. Introduction to the special issue: GIS-based mineral potential targeting. *Journal of African Earth Sciences*, 128: 1-4.
- Yousefi, M., Carranza, E.J.M., 2016. Data-driven index overlay and Boolean logic mineral prospectivity modeling in greenfields exploration. *Natural Resources Research*, 25: 3-18.

