JJEES

JJEES (2024) 15 (2): 94-106
ISSN 1995-6681

Jordan Journal of Earth and Environmental Sciences

Geochemical Evaluation of a Geothermal Region for the Trace Elements Related to the Subsurface Mineralization Using Machine Learning Methods

Farzad Moradpouri[1*], Hamid Sabeti[2]

[1] Department of Mining Engineering, Faculty of Engineering, Lorestan University, Khoramabad, Iran
[2] Department of Mining Engineering, Faculty of Engineering, Birjand University of Technology), Birjand, Iran

Received August 20, 2023; Accepted December 20, 2023

Abstract

The aim of this paper is to identify the promising geochemical anomalies in the geothermal area of Mianeh in northwestern Iran. This was carried out using 840 stream sediment samples, which were analyzed by ICP-MS for 42 elements for Multivariate Geochemical Analysis (MGA). Due to the high migration and mobility power of the Arsenic element, it was selected as the main geothermal pathfinder. Thus (As) element was determined as a labeled variable, and the other important elements and their relationships with (As) can be identified through MGA. Some of these related elements can then be proposed for future exploration. Firstly, for the implementation of MGA, a Probability Plot Modeling (PPM) on the (As) concentration values was carried out to separate background and anomaly subpopulations and extract threshold values for each subpopulation. Secondly, based on the PPM results for the (As) variable, the Discriminant Function Method (DFM), using their functions of code 0, 1, and 2, was implemented to distinguish different variables belonging to each of those subpopulations (background, mixture, and anomaly). Also, Principal Component Analysis (PCA) and Clustering Analysis (CA) were carried out which lead to almost similar results to the DFM. Finally, the elements of Au, Mo, W, Cu, Be, K, and Rb were recognized as the most important elements related to the (As) element. Finally, ArcGIS mapping shows a geothermal region from the NW to NE of the study area for plans and detailed exploration.

Keywords: *Geothermal pathfinders, Stream sediments, Mineralization, Multivariate analysis, GIS mapping.*

1. Introduction

Geothermal energy is a source of clean energy that can provide reliable base-load power generation, or it can be used as a pathfinder for mineral resources in depth in countries and regions where this resource is available (Barbier, 2002; Lund et al., 2003). The potential for the generation of geothermal energy is found in active seismic areas or with volcanic activity worldwide (Grunsky et al., 2009; Mwangi, 2013). Even some of the geothermal resources are ubiquitous in populated areas and easily accessible; many others are found in the depths of the ocean, in mountainous regions, and under glaciers or cubicles (Bloomfield, 2003; Grunsky et al., 2009; Rybach, 2010).

In the crust, the temperature gradient is typically 30 °C per kilometer, but it can be as high as 150 °C per kilometer in a hot geothermal area. The geothermal resources have been developed in several steps, starting with the exploration of the surface, followed by exploratory drilling to discover and confirm the resource availability (Rowley, 1982). In some countries with many hot springs and surface thermal evidence, it might be a sign of geothermal resources (Openshaw, 1983; Zhu et al., 1989; Ghadimi et al. 2012; Seyedrahimi-Niaraq et al., 2019).

The presence of hot springs, geysers, and fumaroles along with hydrothermal alterations are superficial signs of geothermal activity that may indicate the existence of economical geothermal resources in a region (Lund et al., 2003; Grunsky et al., 2009). According to conducted studies in Iran, many geothermal promising regions in NW of Iran indicate the existence of economic geothermal resources. Using this surface evidence, geological information and geochemical analysis and interpretation can achieve a conceptual model of geothermal resources (Yousefi et al., 2010; Najafi and Ghobadian, 2011; Ma et al., 2016).

Geochemical prospecting plays an important role in the exploration of geothermal resources. In recent years, several researchers used machine learning in geochemical explorations (He, et al., 2022). Pan et al., (2023) carried out geological mapping via a convolutional neural network based on remote sensing and geochemical survey. Zhang et al., (2021) used integration of machine learning algorithms to estimate resources in gold deposits. They also used machine learning-based prediction of trace element concentrations in prospectivity mapping (Zhang et al., 2021a; Zhang et al., 2021b).Geochemical stream sediments can reflect various geochemical characteristics related to the probable subsurface ore bodies better than samples such as soil and groundwater (Anderson, 2006; Ranasinghe et al., 2009; Chandrajit et al., 2001; Gransky et al., 2009; Zuo et al., 2009). In geochemical prospecting, it is not always possible to solve a problem by studying only one element, and it is common

* Corresponding author e-mail: moradpouri.fa@lu.ac.ir

for a set of variables to interact and associate. In this case, it is necessary to study several elements or even the study of contents associated with other variables representative of geological or environmental phenomena (Spadoni et al., 2005; Ranasinghe et al., 2009; Moradpouri et al., 2023). On the other hand, the main problem in the analysis and interpretation of geochemical data (especially stream sediments) is the variety and large size of the data matrix. This problem can be solved using multivariate data analysis that can convert large volumes of the data into simple and visual geochemical information (Peh and Halamić 2010; Gransky et al., 2009; Zuo, 2011; Li et al., 2012; Moradpouri and Ghavami-Riabi, 2020; Tobore, et al., 2023).

In this way, multivariate methods have been an important tool for many years for analyzing a large amount of data so that one can work with many variables simultaneously since it is the set of variables that models a geochemical landscape, not just one in isolation (Anderson, 2006; Chandrajit et al., 2001; Gransky et al., 2009; Al-Momani, et al., 2020; Moradpouri and Hayati, 2021). As the various variables interact to form the final observed picture, some of these interactions and associations sometimes appear clearly in multivariate studies. It is important to mention that the obtained results in the applications of multivariate variables must be compared with the geological information available in the study area for interpretation (Davis, 2002; Dillon and Goldstein, 1984; Moradpouri et al., 2017). For each result of a multivariate application, an association with a geological, metallogenetic, or geochemical process in the area should be sought. Thus, when representing such a result on a map, the behavior of a process that operated in the region can be seen, which can be of great value in understanding its evolution, especially in the case of representing mineralization, hydrothermal alteration, or others (Halfpenny and Mazzucchelli, 1999; Helvoort et al., 2005; Peh and Halamić, 2010; Grunsky et al., 2014, Moradpouri and Ghavami-Riabi, 2020; Moradpouri and Hayati, 2021).

Several researchers used the elements Hg and (As) as indicators for exploring geothermal resources (Matlick and Buseck, 1975; Openshaw, 1983; Shiikawa, 1983; Qian, 2009; Jimoh et al., 2023). The other elements of Sb, Bi, B, and some anions and cations are the key elements in the exploration of geothermal resources (Bingqiu et al., 1988; Zuo et al., 2009; Qian, 2009). Risdianto et al. (2010) presented that the intersections of faults might create permeability in depth and influence the flow of geothermal fluids from reservoirs. Mwangi (2013) used conservative constituents for tracing the origin and flow of geothermal fluids, stable isotopes along with B and Cl as the most important elements. He also used rock-forming constituents of SiO_2, Na, K, Ca, Mg, CO_2, and H_2 to predict subsurface temperatures and potential production problems such as deposition and corrosion (Mwangi, 2013).This paper aims to analyze the stream sediment data of the Mianeh geothermal region in NW of Iran for geothermal pathfinders and new insight into probable subsurface mineralization. The data included 840 stream sediment samples, which were analyzed by ICP-MS methods for 42 geochemical elements. According to the hot spring in the region and field study, the element of (As) was identified as the starting geothermal pathfinder to carry out the Multivariate Geochemical Analysis (MGA).

1.1. Arsenic (As) as Pathfinder

In the early period of geochemical exploration, arsenic (As) and antimony (Sb) were commonly used as so-called pathfinder elements for Au (Boyle, 1979). The name is given because of the common association of gold mineralisation with rocks enriched in arsenopyrite (FeAsS) and/or stibnite (Sb2S3), and because As and Sb were easier to determine analytically due to their higher abundances (Steenfelt, 1996).

Several researchers suggested the use of the identified elements expressly using the strongly correlated elements such as As and Zn to detect hidden anomalies in the complex environment (Dzigbodi-Adjimah, 1993; Kesse, 1985; Arhin et al., 2015).

Yaisamut et al., (2023), presented the interpretation of intricate spatial dispersion patterns and concentration levels of deposit pathfinder elements, specifically arsenic (As), copper (Cu), and zinc (Zn), using a comprehensive array of stream sediment geochemistry data.

2. Geological setting

The Iranian crust is divided into several main and sub-zones based on magmatic, metamorphic, sedimentary, and tectonic occurrences. Mianeh region is located in the NW of Iran (Figure 1). This region has undergone many changes, so that the effects can be monitored from the Precambrian (Mianeh metamorphic unit) to the recent (Sabalan and Sahand volcanism). When looking at the geological map of the Mianeh in Figure 1, it is observed that Tertiary sediments, volcanic rocks, and intrusive masses cover most of it. Most outcrops of the Mianeh region are almost composed of volcanic and sedimentary rocks associated with the Cenozoic by which Eocene and Oligocene volcanic rocks cover the largest area in the region. These rocks often have a combination of rhyolite, ignimbrite, rhyodacite, and trachyte, and the results of the chemical analysis indicate the alkaline nature of this volcanism (Aghanabati, 2004). During the Eocene, the Mianeh region witnessed a lot of volcanic activity, which is often in the form of lava flows with andesitic, basaltic, and andesitic trachea with various combinations of tuff layers. The lowest part of the Eocene volcanic rocks is in the form of pyroclastic deposits, gray acidic tuffs, and acidic lavas followed by andesitic lavas and sandy tuffs (Aghanabati, 2004). A large volume of Eocene volcanic rocks in the form of mega-porphyry andesite is found in the Bozgoosh mountain range. The transition from Eocene to Oligocene volcanism is in the form of ignimbrite, rhyolite, trachyte, and shear tuff rocks that are located on Eocene-related units. Miocene lithology in the regions includes conglomerate, siltstone, sandstone, basalt, andesite, tuff, marl, hornblende, siltstone marls, sandstone conglomerates, shales, and pyroclastic rocks. Miocene deposits are exposed in the southern areas around Mianeh. In addition, Quaternary deposits include old terraces, new terraces, paved limestones, and salt plains that are exposed in the areas around Mianeh (Aghanabati, 2004). The location of the city's border, samples location, Maman hot spring, streams, and river location can also be seen in Figure 1.

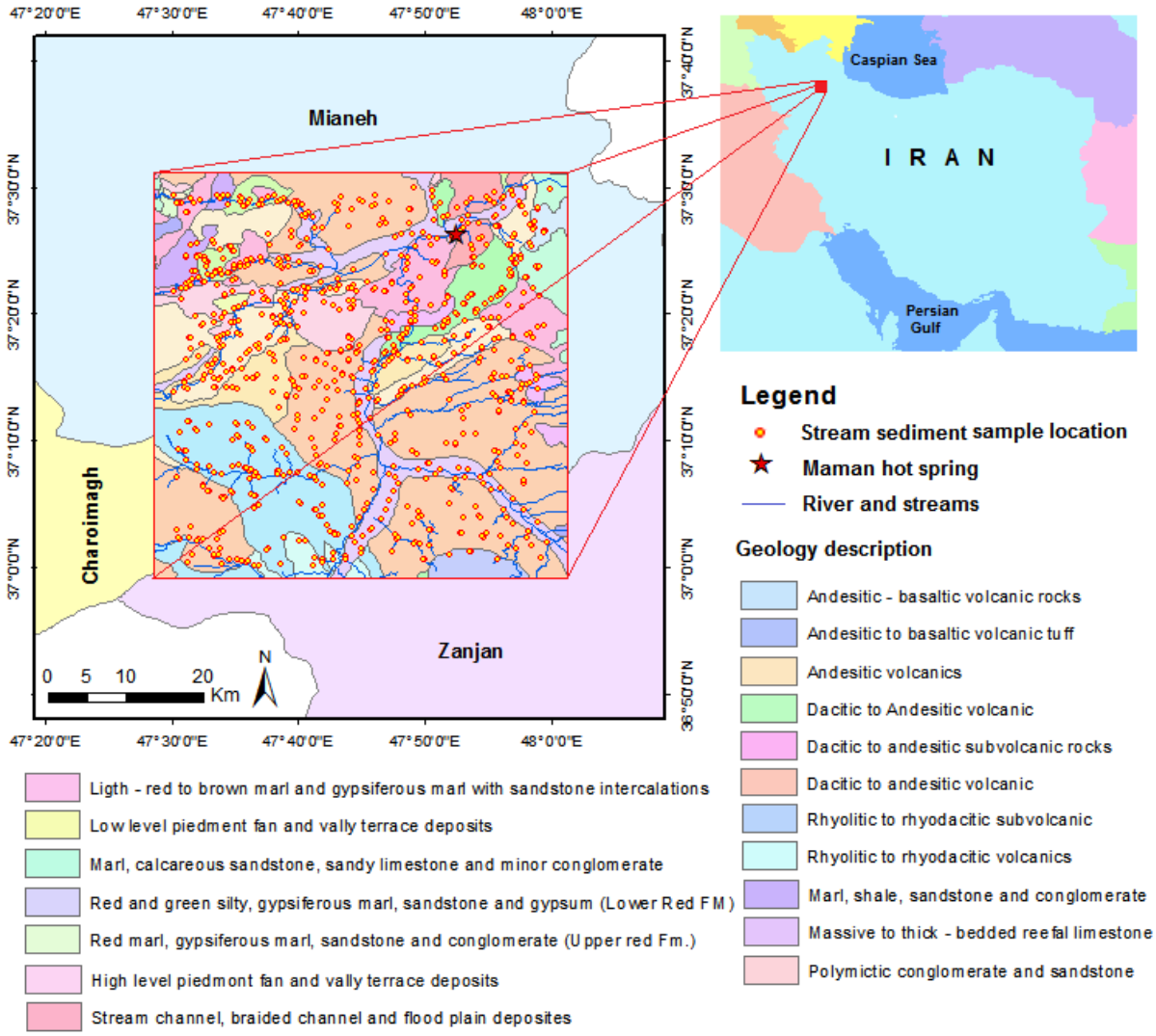


Figure 1. Geological setting and location map of the study area, stream sediment samples, hot springs, rivers, and streams

Rivers, streams, and hot springs: The main rivers in the study area are shown in Figure 2a. Qezel Owzan is one of the longest rivers in Iran. On its track, Zanjanrood and Abharrood Rivers join it, and before the city of Mianeh, the Qaranqo, Shahrchai, and Aidoqmosh Rivers, which flow from west to east join it, and it finally flows into the Caspian Sea (Figure 2a). In addition to the mentioned rivers and due to the topography, the study area includes many streams that are shown in Figure 2b.

The other important phenomena in Eastern Azarbaijan province are hot springs. Some of which are located around Mianeh city. The hot springs of Maman, Sabz, Ivorq, and Gudarq are among the most prominent ones. Sulfur, arsenic, salt, copper, and iron ions are among the most important combinations in these hot waters. Maman hot springs, at a distance of 100 meters from each other, are located in the region of Maman village in the central part of Miyaneh city and the northeast of the study area (Figure 2b). These springs also have therapeutic characteristics of the sulfur type. The water of Maman hot spring is naturally yellow which indicates the presence of sulfur, which can also be seen in the color of the rocks inside and around the hot springs. In addition, many other hot and cold springs in the surrounding regions can be used in geochemical studies to find trace elements. Furthermore, Figure 2b shows a satellite image of the study area including the location map of the stream sediment samples, Maman hot spring, river, and stream sediments which help to interpret the results of the geochemical multivariate analysis.

Figure 2. Location map: (a) rivers (b) streams and stream sediment samples

3. Material and Methods

With the above variety of information (topography, geology, rivers, hot springs), the geochemical analysis and interpretation of stream sediments for probable subsurface mineralization could be a multifaceted work, but the results considerably help us determine the promising areas for future plans and detailed exploration. Therefore, the geochemical sampling of 840 stream sediment samples employed the use of a mattock or shovel to reach appropriate samples from the available river and streams in the study area. Due to the location and nature of the streams, sampling spacing was different. All samples were analyzed for major and trace elements cocentrations by inductively coupled plasma mass spectrometry (ICP-MS) for 42 elments, following a 95°C Aqua Regia digestion. Analytical uncertainties vary from 0.03% to 0.1% for major elements and from 0.1% to 0.55% for trace elements.

3.1. Multivariate Geochemical Analysis (MGA)

A large number of techniques are available that address problems in space with multidimensional approaches to one, two, or more than two populations. Multivariate data are almost all based on the matrix of element concentration. Among the most used methods in geochemistry, principal component analysis, clustering analysis, factor analysis, and discriminant analysis can be mentioned.

Basic statistics:

First, the univariate statistical analysis of the dataset, including maximum, minimum, mean, and standard deviation was carried out for a primary consideration of abnormal variations so that some of the variables show meaningful values (Table 1). Among those elements and based on the reasons that were mentioned, the arsenic (As) element was identified as the geothermal pathfinder in the study area to carry out the Multivariate Geochemical Analysis (MGA) that clarifies which of these elements have more relationships with the As as a pathfinder and which could be a sign of subsurface mineralization. The mean value of (As) in the study area is 20.07 (ppm), and its maximum value is 174.1 (ppm) which may be a meaningful sign of subsurface geothermal activities. In addition, the analysis of histograms, Q-Q and P-P plots, was carried out for the (As) data which indicates a reasonable continuity and lognormal nature in the dataset (Figure 3). Therefore, it was considered for existing sub-populations, thresholds, and background and anomaly separation using PPM.

Table 1. Basic statistics of the stream sediments dataset in the study area (N = 840)

Variable	Min.	Max.	Mean	Std. deviation	Variable	Min.	Max.	Mean	Std. deviation
As	1.460	174.1	20.069	12.942	Mo	0.010	15.61	1.490	1.338
Au	1.000	88	2.338	4.18	Na	1376	73830	13422.8	6550.1
Ag	0.000	2	0.367	0.164	Nb	4.87	69.67	16.435	5.709
Al	23210	103300	64196.7	13703.5	Ni	2.68	74.27	28.366	11.3
Ba	129.6	2516	675.068	208.582	P	20.14	4321	993.05	758.576
Be	0.64	5.82	1.700	0.594	Pb	9.61	172.3	32.307	13.915
Bi	0.01	12.87	1.501	1.255	Rb	12.49	266.9	69.074	30.482
Ca	11240	283200	78684.4	27033.9	S	10	246100	12055.0	27337.8
Cd	0.1	6.48	0.341	0.268	Sb	0.14	342	2.451	11.79
Ce	10.94	111	51.739	12.436	Sc	4.23	25.450	11.781	3.25
Co	4.53	57.83	15.712	5.636	Sn	1.8	7.6	3.204	0.676
Cr	3.61	227.4	63.184	26.408	Sr	117.8	6124	629.82	584.64
Cs	1.0	27.55	5.889	2.535	Te	0.08	1.0	0.156	0.086
Cu	9.090	203.5	34.70	20.98	Th	3.84	83	14.892	5.887
Fe	17480	335750	53608.2	20787.5	Ti	1363	35455	5409.1	2383.8
Hg	0.05	0.6	0.089	0.026	Tl	0.4	2.35	0.918	0.235
K	3434	50800	19424.5	6558.1	U	0.94	17.23	4.437	1.554
La	7.51	75	27.344	7.637	V	41.96	1147	152.98	84.706
Li	11.33	185.8	28.842	12.291	W	0.27	54.3	1.754	2.227
Mg	96.41	32810	9951.3	4237.1	Zn	24.64	334.1	75.95	29.441
Mn	246.7	4431	903.246	283.6	Zr	75.7	911.2	315.02	121.75

Probability Plot Modeling (PPM):

An important issue in the interpretation of various geochemical data with the prospecting perspective is the determination of threshold values that separate the ranges of values defined as representing background and anomalous values. Thus, the probability plot modeling on (As) element as the most important element related to geothermal activities was carried out to de-convolute the subpopulations, which separate the background from the anomalous values. The results of (As) probability diagram modeling are shown in Figure 4. The horizontal axis is the cumulative percent and the vertical axis is the logarithm of concentration values for specific groups of data as open circles (Figure 4a). It should be added that the least-squares method was used to fit the best model to the data. The continuous curve that passes through the groups of data (open circles) is the fitting model that presents the recombination of the sub-populations represented by the sloping straight lines. It is obvious that

three sub-populations can be recognized in the data (Figure 4b). The final step is the improvement of the previous fitting model, which is shown in Figure 4c with the final threshold values. The result of the probability plot modeling, including the details for each subpopulation and the threshold values, is presented in Table 2.

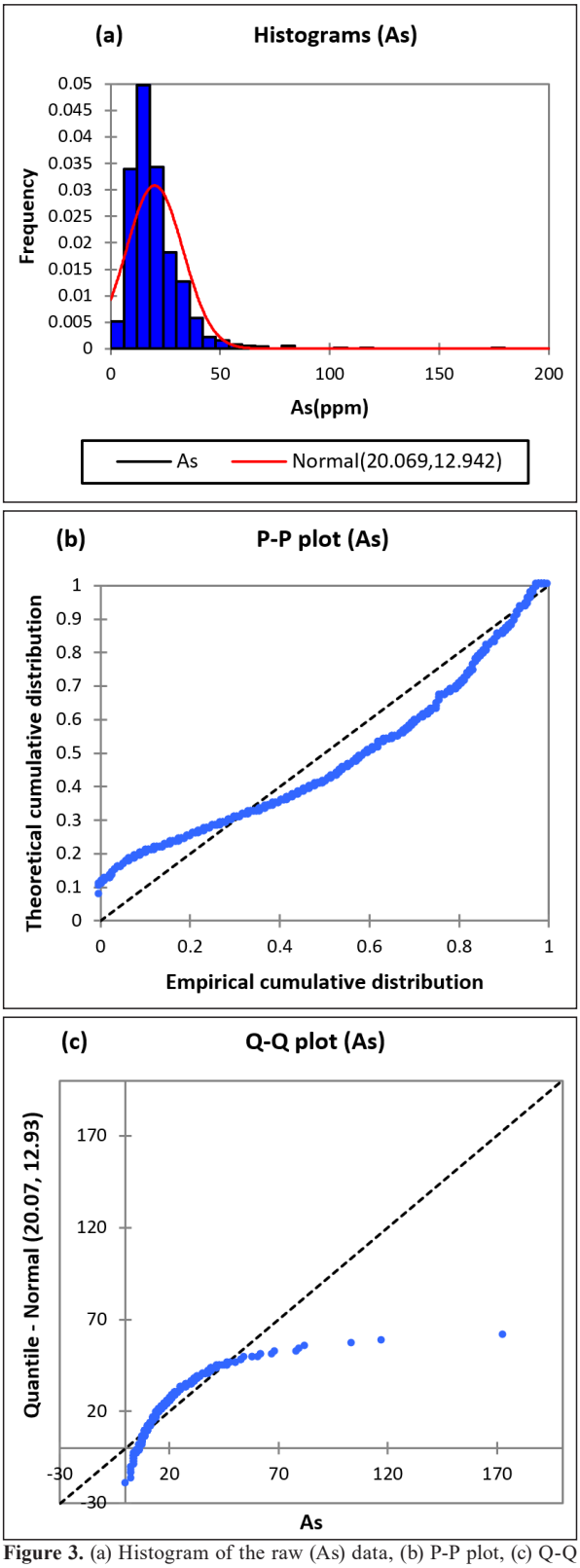


Figure 3. (a) Histogram of the raw (As) data, (b) P-P plot, (c) Q-Q plot.

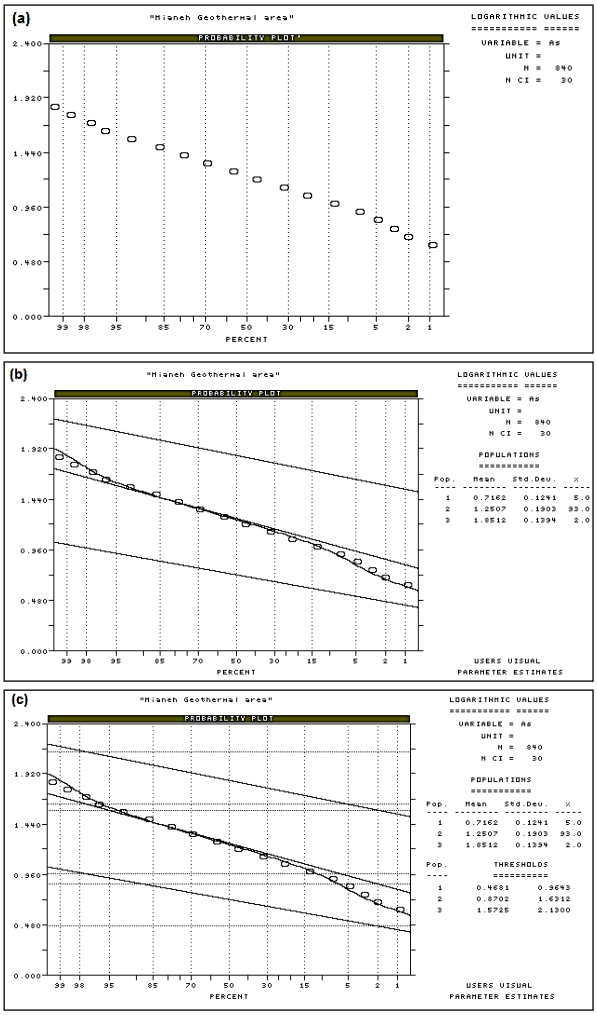


Figure 4. The results of the probability plot modeling. (a) Probability plot of the logarithmic (As) values; (b) First three subpopulation estimates. (c) Results for the revision of the first estimate by three optimization steps to minimize the deviation between data (circles) and model (the fitted curve).

Table 2. Mean, percentage, and threshold values after deconvolution for the three recognized (As) subpopulations

Variable=As	Unit=ppm	N=840	Population	Thresholds
Population	Mean	Percentage	1	2.94-9.21
1	5.20	5.0	2	7.42-42.78
2	17.81	93.0	3	37.37-134.9
3	70.99	2.0	Transform=Logarithmic	
Populations=3 Missing observation=0				

According to the results of probability diagram modeling (in Figure 4 and Table 2), samples with (As) concentration less than 9.21 ppm were identified as background subpopulation, The concentrations values between 9.21 to 42.78 ppm were identified as mixtures of background and anomaly. Finally, samples with concentrations above 42.78 ppm were defined as the anomaly subpopulation. Therefore, the location map of (As) distribution for these three subpopulations is shown in Figure 5. As can be seen, most of the representing anomalous samples are located in the north part of the study area especially along with Qaranqo River and in Maman hot spring. This was somehow expected and showed a promising sign, but it needs to be interpreted considering the

whole geological and geochemical evidence after the proper analysis in the later sections. Thus, first a multivariate discriminant analysis was carried out to determine the most related elements to the (As) content using the PPM threshold results of Table 2. In addition, the multivariate geochemical analysis of PCA and clustering analysis were carried out for comparison.

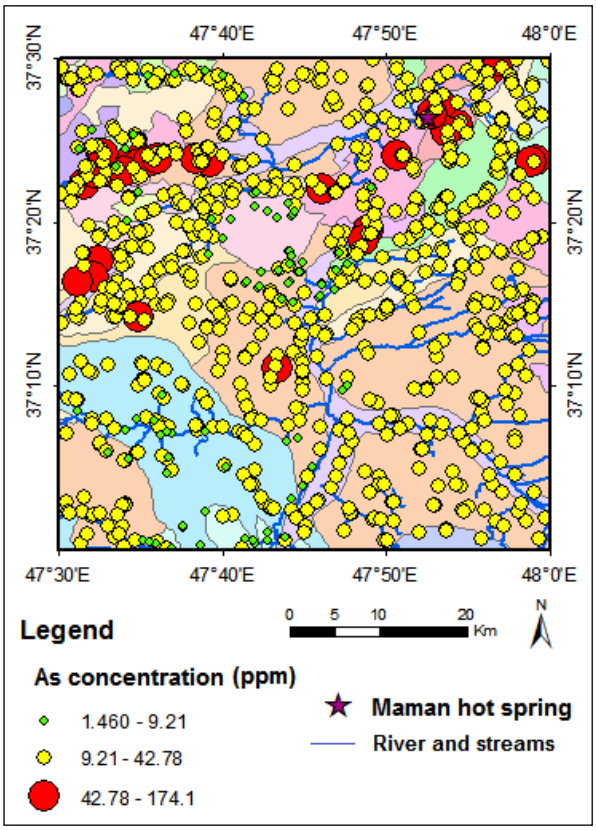


Figure 5. Arsenic (As) content distribution based on the probability plot modeling threshold values

Discriminant Function Method (DFM): DFM is one multivariate statistical data analysis that is used for the evaluation of the dependence of one sample on one of two (or more) known populations. The discriminant function is defined based on known features of those populations. In order to define the discriminant function, a linear combination of all variables is used to create a normally distributed univariate variable. This function is calculated in the direction of the highest discrimination of the populations (Peh and Halamić, 2010).

The coefficients of this linear combination is defined as follows.

$$[a]^T = \left(\left[\bar{X}_1 - \bar{X}_2\right]\right)^T [S]^{-1} \quad (1)$$

where $[a]^T$ is the coefficients matrix, $\bar{X}_1$ and $\bar{X}_2$ are estimated mean vectors of two populations, is the reversed covariance matrix of two populations.

Now, the discriminant function is calculated using the following equation.

$$DF = \left(\left[\bar{X}_1 - \bar{X}_2\right]\right)^T [S]^{-1} [X] \quad (2)$$

where $[X]$ is the vector of the newly observed sample, there is a function that defines the separation boundary of two populations called the critical function as in the following equation.

$$[D]_0 = \left(\left[\bar{X}_1 - \bar{X}_2\right]\right)^T [S]^{-1} \left(\frac{\left[\bar{X}_1 + \bar{X}_2\right]}{2}\right) \quad (3)$$

This function is a vector, which starts from the end of $[\bar{X}_1]$ and finally reaches to the separation boundary of two populations. If then $DF < D_0$ the sample is assigned to population 1 and if then $DF > D_0$ samples are assigned to population 2.

Arsenic mentioned above, (As) element was chosen as the dependent variable versus the other independent variables (elements). Thus, the linear discriminant functions of Code 0, Code 1, and Code 2, were defined based on the probability plot modeling of the (As) variable in Table 2. The analysis was implemented using the dataset of 42 variables related to the 840 stream sediment samples. Based on the threshold values, the (As) values of the background population with a concentration of less than 9.21 ppm were defined as Code 0, the values between 9.21 and 42.78 ppm were defined as Code 1 (a mixture of the background and anomaly), and finally, the values above 42.78 ppm were defined as Code 2 (anomaly). This forms the basis of the linear function that was used for the DFM, which evaluates the relationship between variables through the three defined codes and the two discriminant functions.

Based on the DFM, two functions can be defined which are shown in Table 3 with its parameters. Function 1 describes 90.3% of the variability with a correlation of 84% with the data. In addition, function 2 describes 9.7% of the variability with a 45% correlation. Contrary to the correlation coefficient, the lower the Wilks' Lambda coefficient, the better the discriminant function. The larger the Chi-square coefficient, the better the function could justify the variability.

Table 3. Calculated discriminant functions and validation parameters

Function	Eigenvalue	Variance percentage	Cumulative percentage	Correlation	Wilks' Lambda	Chi-square	df	Sig.
1	2.349	90.3	90.3	0.84	0.239	1169.663	84	0.000
2	0.251	9.7	100.0	0.45	0.799	182.887	41	0.000

In the next step, the samples were classified according to two discriminant functions. In Figure 6, each classified sample was coded and a different color was assigned for visualization purposes. This figure shows that considerable discrimination has occurred between the data related to sub-populations. Then, the variables were discriminated based on the two defined functions between the three disputed sub-populations that are shown in Figure 7, which shows one reality that describes separation boundary based on (As) variable coded values and their functions. In addition, the

other variables' locations in Figure 7, are outputs of the DFM. As can be seen in Figure 7, As, Au, Mo, W, Th, K, Rb, Be, Ce, Fe, and Nb were located in the anomalous part as tracing elements for (As) element and probable mineralization (Code 2). Cs, Co, Ce, Hg, P, Ni, Li, and Zr were located in the mixture part (Code 1), and the rest were located in the background part (Code 0).

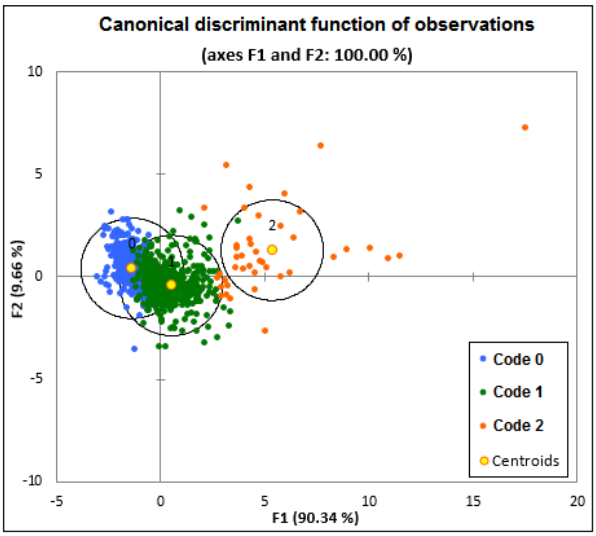


Figure 6. Classified samples based on discriminant function 1 and 2

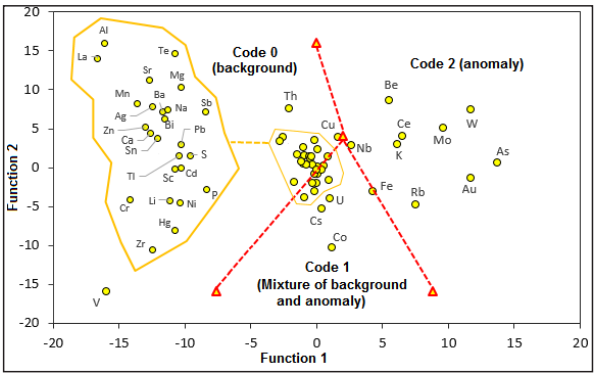


Figure 7. Standardized canonical discriminant function coefficient based on the three discriminant functions.

To evaluate the results of DFM, based on p-p modeling and according to the discriminate functions, 93.2%, 86.6%, and 81% of the samples were correctly classified as code 0, Code 1, and Code 2 respectively when three different groups memberships were used (Table 4). In addition, as can be seen in Table 3, this method could validate that 92% of code 0, 83.5% of code 1, and 71.4% of code 2 were classified properly.

Principal Component Analysis (PCA):

PCA is a way of identifying patterns in data and expressing them in a way that highlights their similarities and differences, with the advantage that once these patterns are identified, we can reduce the number of dimensions without much loss of information. We choose components and form a vector feature, and the eigenvector with the highest corresponding eigenvalue is the principal component of the dataset. Minor components can be ignored, with some loss of information, but if the eigenvalues are small, the loss is not very significant.

Table 4. Classification and validation of the DFM results.

	Code		Predicted Group Membership			Total
			0	1	2	
Original	Count	0	314	23	0	337
		1	61	399	1	461
		2	0	8	34	42
	%	0	93.2	6.8	.0	100.0
		1	13.2	86.6	.2	100.0
		2	.0	19.0	81.0	100.0
Cross-validated	Count	0	310	27	0	337
		1	71	385	5	461
		2	0	12	30	42
	%	0	92.0	8.0	.0	100.0
		1	15.4	83.5	1.1	100.0
		2	.0	28.6	71.4	100.0

In Geology, PCA can represent active geological processes, present lithological changes or events causing mineralization. The dimensionality of the problem is also reduced. In addition, it allows the identification of the variables that have very high or low contributions to the variability of the study area. An identification that allows suggesting the elimination of these variables in future stages of study. In addition, a very common objective in geochemical programs is the classification of observations that make up a given sample (rocks, stream sediments, soils, or water). This classification is often based on the measured concentration. It is very useful in cases where large amounts of observations are available and in cases where there is little prior knowledge of the meaning of the genesis of the constituents.

In the current research, PCA was processed with the raw data of 840 stream sediment samples for 42 variables to explore the elemental associations. There are several techniques for decision-making on the retention of main components, the most accepted and recommended being the one that proposes the retention of main components whose total variability percentages are considered significant, that is, they explain an important proportion of the total variations present. In the original data set. It is evident that, at this point, the participation of the technician who applies PCA is decisive; the non-use of components from a certain level (from the fourth main component, or the fifth, for example) does not, however, cause any problems on the ones used, nor does it affect their interpretation under any circumstances. It can only leave processes represented by the unused components, which have no influence, as stated above, on the interpreted components (and their associated processes). Table 5 presents the variability of each component and its respective eigenvalue for the first component among the whole component. Actually, the first two components explain 45.63% of the variability of the data, while the first seven components explain 66.9%. As can be seen, in Figure 8, after the fifth component, one can see a mild trend in variability, which means that we can ignore the components with the low contribution in introducing the variability in the data.

Table 5. Main components and their respective eigenvalue and variability

	F1	F2	F3	F4	F5	F6	F7
Eigenvalue	13.502	5.664	2.734	1.916	1.544	1.417	1.321
Variability (%)	32.147	13.486	6.510	4.562	3.676	3.374	3.146
Cumulative %	32.147	45.633	52.143	56.705	60.381	63.756	66.901

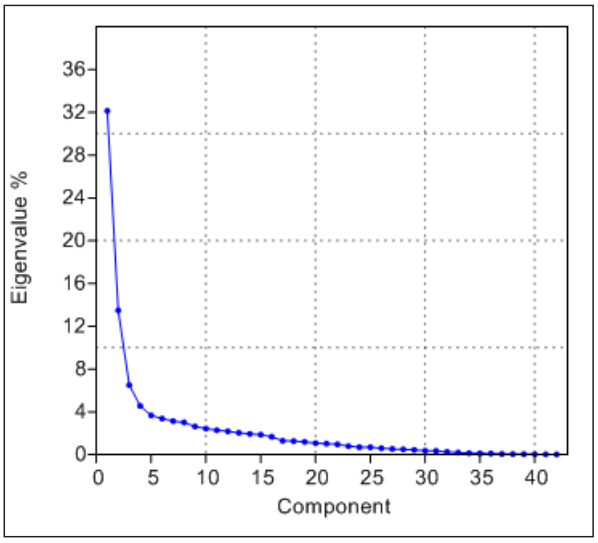


Figure 8. Contribution of variables to the main components

Figure 9a-c shows the correlation between the first three components with the coordinates of the variables. The coordinates are based on correlations and are plotted according to components F1, F2, and F3. As can be seen in Figure 9a, the first group from the loadings of F1 consists of Na, Ni, Tl, Mn, Ba, Nb, Li, Th, Zn, Sn, Cr, Co, Fe, Cs, Ti, V, Hg, Bi, and Tl, and the second group from the loadings of F_2 consists of K, Be, Rb, Au, As, Mo, Sb, P, Cu, W, Ag, Zr, Al, Ce, La, Cd, Pb, and Sc. In addition, in each group, one can interpret the relationship and correlation between elements. For example, the most related elements to (As) as the main pathfinder element in Mianeh geothermal area in F1/F2 map are Au, Sb, Mo, Cu, W, K, Rb, Be, Zr, and P. It is both because of the low angle between their straight lines and rather due to their close distance of (As), Au, Mo, and Sb from the center, we need to obsessively interpret this association and we should look at the relationship between these variables on the other pair of components map. On the other hand, the correlation between the ratio of elements shows that for example for the elements (As), Au, and Cu, there is no correlation between the As/Au ratio with As/Cu, but the short length of the line that connects both Au and Cu to (As) shows the lower variance between each pair of these elements. As an example of a high correlation between two pairs of ratios, the high correlation for the ratio of As/Au to As/Sb can be mentioned. Figure 10a-c shows the coordinates of the samples according to components F1, F2, and F3. Figures 9 and 10 determine the importance degree of each factor based on the location of the different variables and samples' location.

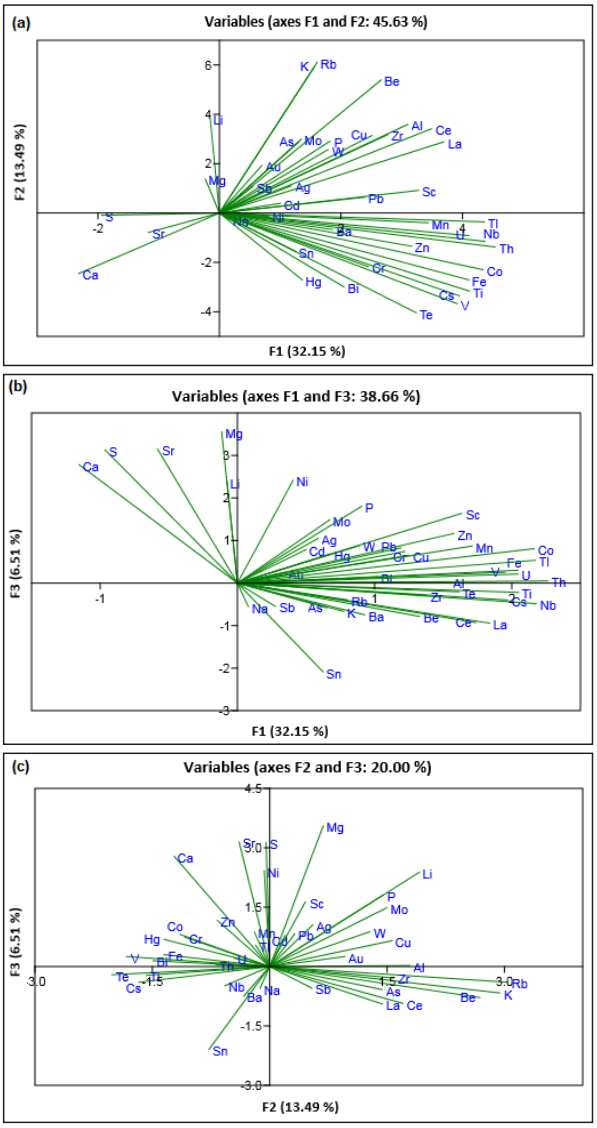


Figure 9. Coordinates of the variables when the axes represent the variable loading values in F1 (principal component 1), and F2 (principal component 2) and F3 (principal component 3). (a) in the F1 and F2 domain, (b) in the F1 and F3 domain, (c) in the F2 and F3 domain.

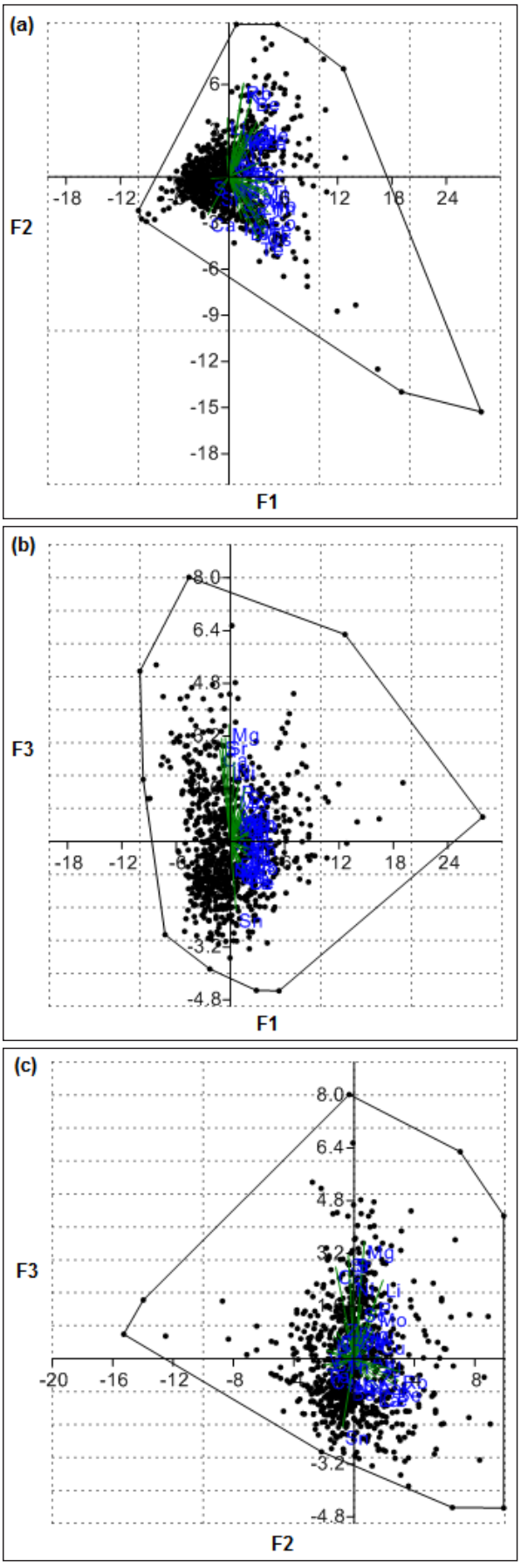


Figure 10. Sample coordinates in PCA. (a) F1 and F2, (b) F1 and F3, (c) F2 and F3.

Clustering Analysis (CA): The application of the CA is very useful as an exploratory tool for data, that is, it can indicate associations /groups that are not previously suspected, and it can often simply reflect some obvious lithological associations. It aims to classify objects (in the case of geochemistry, field samples, or analyzed variables) for similarities and/or dissimilarities, and the groups generated must have high internal homogeneity and high external heterogeneity. Cluster analysis is applied to objects (Q technique), whose variables can be also examined for similarities and groupings (R technique). The different methods for cluster analysis can be framed into four general types: partition methods, methods with arbitrary origin, methods by mutual similarity, and methods by hierarchical groupings.

The method used in this work was hierarchical groupings (the most used in geochemistry) in R mode. The grouping was processed using the method of Ward´s agglomeration (minimum variance), grouped by correlation coefficient. The same raw data used for DFM and PCA was also applied for the CA. In order to evaluate similarities between elements, cluster dendrograms, showing hierarchical variable clustering, are presented in Figure 11. It shows the lines linked according to levels of similarity that group pairs of the variables. The higher the correlation, the shorter the connection distance between variables. From right to left, the first cluster of orange color is largely made up of elements associated with the (As) element as the geothermal pathfinder. This cluster includes the elements Au, As, Mo, W, Be, K, Rb, Zr, Ce, La, Ag, Al, Sc, Cu, and P, so that some of them have a closer connection distance, and some are far away. However, the results are rather similar to the DFM and PCA results, and one can classify a pack of those elements as the geothermal pathfinders of the occurred mineralization based on the geological setting.

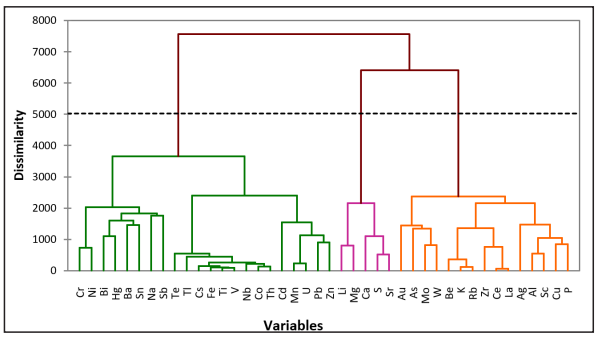
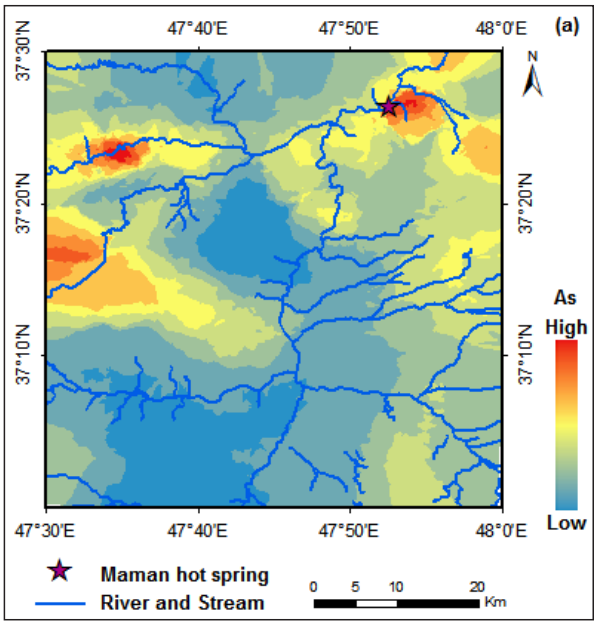
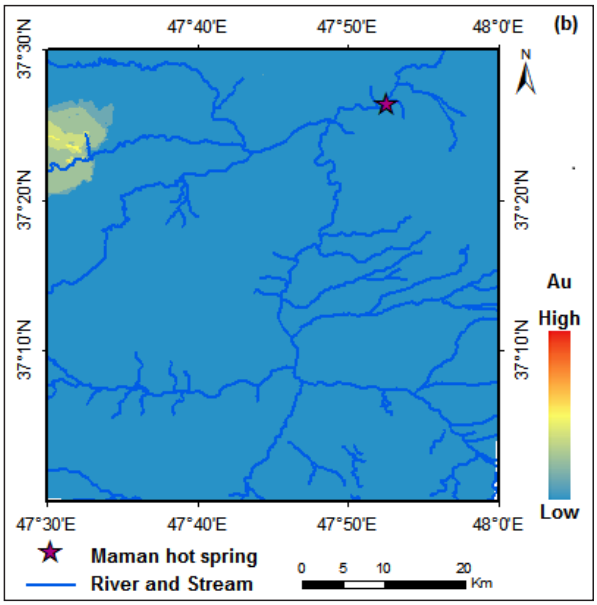


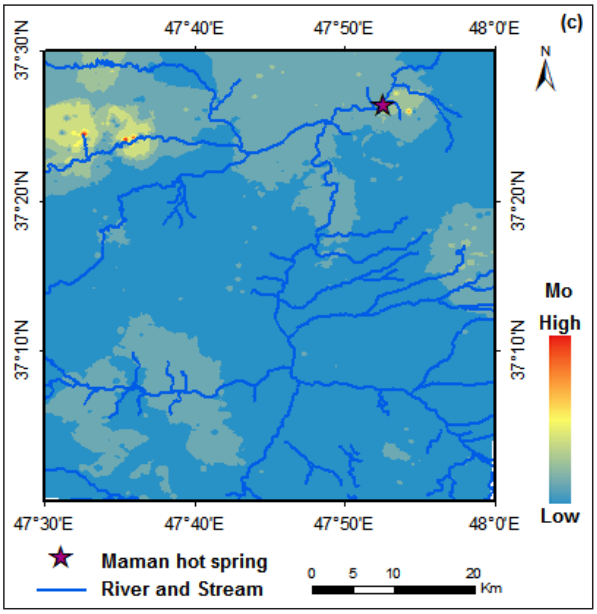
Figure 11. The dendrogram in R mode shows the hierarchical variable clustering by correlation.

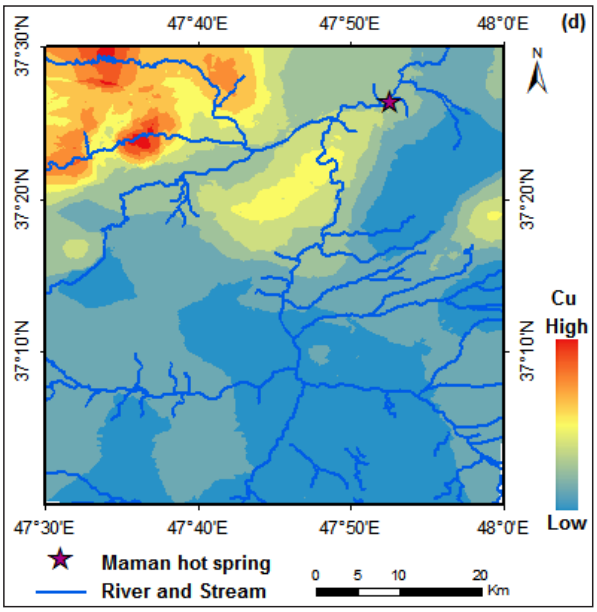
GIS mapping: As discussed in the previous sections, the results obtained from DFM, PCA, and CA indicated that a number of elements were identified as related to the arsenic geothermal pathfinder (e.g. Au, Mo, W, Sb, K, Rb, P, Cs, Be, La, Cu, Zr, Ag, and Al). Among these associations, the elements (variables) of Au, Mo, W, Cu, Be, K, and Rb were recognized as the most important related elements to (As) based on the multivariate geochemical analysis (MGA). Some of them can be interpreted as elements related to mineralization (e.g. Au, Mo, Cu). As, Au, Mo, and Cu reinforce the probability of porphyry deposit. It should be

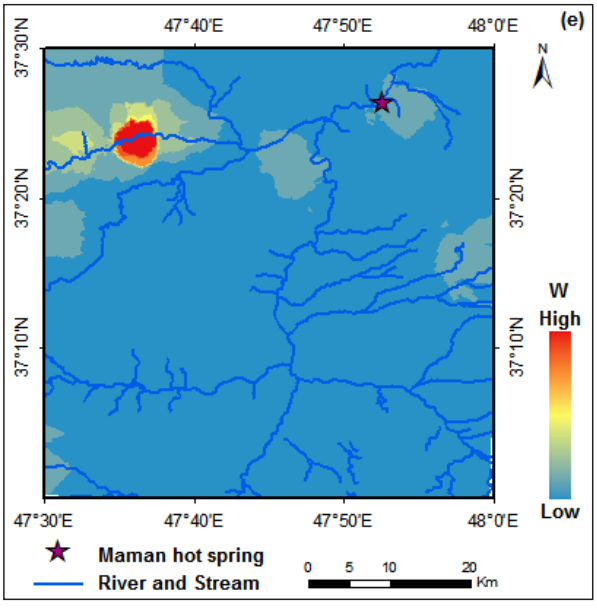
added that (As) has a direct relationship with gold deposits. Among the important geothermal pathfinders such as Sb, Hg, and Bi, only the Sb element showed the relationships with (As) element only in PCA which cannot be seen in DFM or CA. Sb is an element of the nitrogen family (N, As, Sb, P, Bi) and has rather similar chemical properties to Arsenic (As). On the other hand, the elements of K, Rb, and Be are present in alkali feldspar minerals of the volcanic rocks. Distribution maps of the above-related elements were prepared in ArcGIS10.5 software using the kriging method and are shown in Figure 12. In this Figure, the anomalous area of the arsenic geothermal pathfinder and the main related elements of mineralization can be seen. These areas are proposed as promising regions for the later exploration steps.











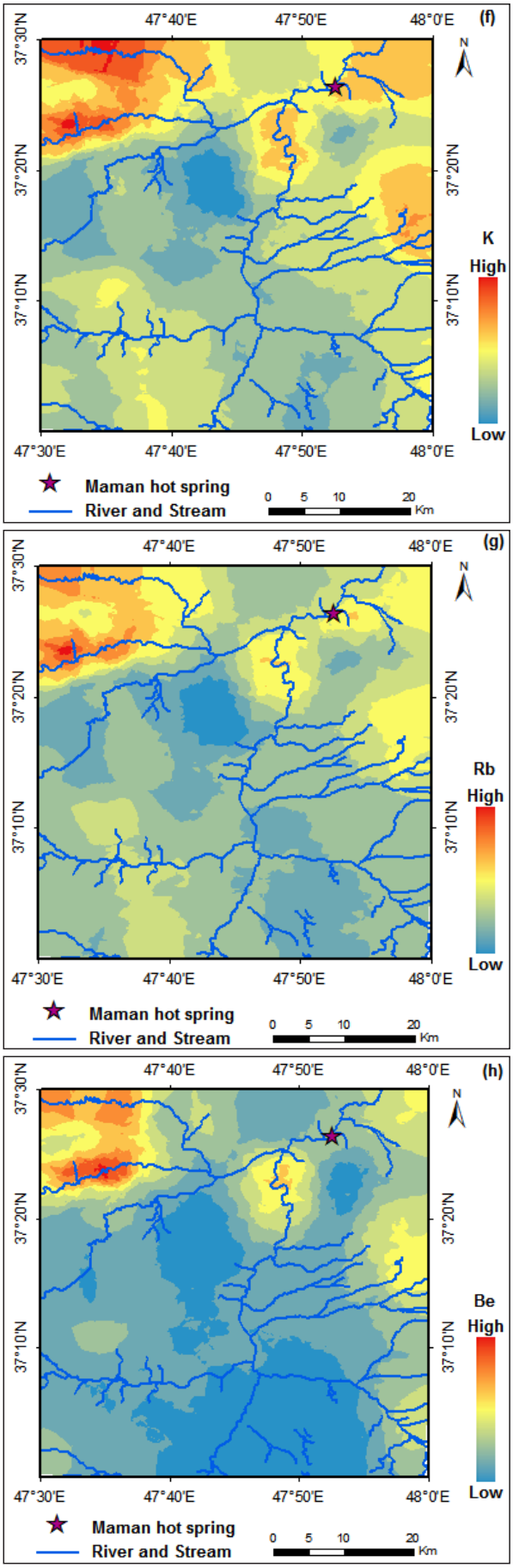


Figure 12. Anomalous area of the geothermal pathfinders and the probable variables of mineralization. (a) As, (b) Au, (c) Mo, (d) Cu, (e) W, (f) K, (g) Rb, (h) Be

4. Conclusion

The results of the present study can be divided into several parts. (a) Due to the selection of arsenic element (As) as a suitable geothermal pathfinder in the study area, the MGA proved that none of the other geothermal pathfinders such as Sb, Hg, and Bi showed a significant relationship with (As) pathfinder. Thus, arsenic plays a key role in the prospecting of geothermal resources in the study area using stream sediments. (b) What can be deduced from the integrated map of the promising areas, geological and MGA indicates that there is an active geothermal region in the northwest of the study area that includes the rivers of Shahrchai, Qaranqu, and Aidoqmosh that originate from volcanic areas of the Sabalan and Sahand mountains and continue to Maman hot spring in the Northeast. These paths often consist of volcanic rocks. (c) Different MGA methods of the present stream sediment data revealed several multi-element associations describing the existence of subsurface probable mineralization in the NW to NE of the study area for plans. Among these associations, the elements of Au, Mo, W, Cu, Be, K, and Rb were recognized as the most important related elements to the (As) variable. The elements of Au, Mo, and Cu can be interpreted as the elements related to the mineralization, which may indicate the probability of a porphyry deposit. The elements of K, Rb, and Be are present in alkali feldspar minerals of the volcanic rocks.

Conflict of Interests

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

Aghanabati, A. (2004). Geology of Iran, Geological Survey and Mineral Exploration of Iran. 586 p. (In Farsi).

Al-Momani, T., Alqudah, M., Ezz-Aldien, M. (2020). Dwairi Mineralogical and Geochemical Characterization of Jarash Kaolinitic Clay, Northern Jordan. Jordan Journal of Earth and Environmental Sciences 11 (4): 272-281.

Anderson, T. (2003). An Introduction to Multivariate Statistical Analysis. 3rd Ed., John Wiley and Sons. 742 pp., ISBN 0-471-36091-0.

Arhin, E., Boadi, S., Esoah, M.C. (2015). Identifying pathfinder elements from termite mound samples for gold exploration in regolith complex terrain of the Lawra belt, NW Ghana. Journal of African Earth Science 109: 143–153.

Barbier, Enrico. (2002). Geothermal energy technology and current status: an overview. Renewable and Sustainable Energy Reviews 6: 3-65.

Bingqiu, Z., Lixin, Z., Changyi, S., Hui, Y. and Gongyuan, W. (1988). Application of geochemical methods in the search for geothermal fields. Journal of Geochemical Exploration 33: 171-183.

Bloomfield, K.K., Moore J.N., Neilson R.N. (2003). Geothermal energy reduces greenhouse gases. Geoth Res Counc Bull 32(2): 77-79.

Davis, J.C. (2002). Statistics and Data Analysis in Geology. 3rd Ed., John Wiley and Sons, NY. ISBN: 978-0-471-17275-8.

Dillon, W.R. and Goldstein, M. (1984). Multivariate Analysis: Methods and Applications", John Wiley and Sons, New York, 587 S. ISBN 0-471-08317-8. https://doi.org/10.1002/bimj.4710290617.

Dzigbodi-Adjimah, K. (1993). Geology and geochemical patterns of the Birimian gold deposits, West African Journal of Geochemical Exploration 47 (1–3): 305-320.

Ghadimi, F., Mirzaei, M., Ghomi, M., Mina, M. (2012). Hydrochemical properties of the thermal waters of Mahalat Abgarm, Iran. Geothermal Resources Council Transactions 36: 1355– 1358.

Grunsky, E.C., Drew, L.J., Sutphin, D.M. (2009). Process recognition in multi-element soil and stream-sediment geochemical data. Applied Geochemistry 24:1602–1616. https://doi.org/10.1016/j.apgeochem.2009.04.024.

Grunsky, E.C., Mueller, U.A., Corrigan, D. (2014). A study of the lake sediment geochemistry of the Melville Peninsula using multivariate methods: applications for predictive geological mapping. Journal of Geochemical Exploration 141: 15–41. https://doi.org/10.1016/j.gexplo.2013.07.013.

Halfpenny, R. and Mazzucchelli, R.H. (1999). Regional multi-element drainage geochemistry in the Himalayan Mountains, northern Pakistan. Journal of Geochemical Exploration 67(1-3): 223-233.

He, Y., Zhou. Y ., Wen, T., Zhang, S., Huang, F., Zou, X., Ma, X., Zhu, Y. (2023). A review of machine learning in geochemistry and cosmochemistry: Method improvements and applications, Applied Geochemistry 140: 105273. https://doi.org/10.1016/j.apgeochem.2022.105273.

Helvoort, P.J., Filzmoser, P., Gaans, P.F.M. (2005). Sequential Factor Analysis as a new approach to multivariate analysis of heterogeneous geochemical datasets: an application to a bulk chemical characterization of fluvial deposits (Rhine–Meuse delta, The Netherlands), Applied Geochemistry 20: 2233–2251. https://doi.org/10.1016/j.apgeochem.2005.08.009.

Kesse, G.O. (1985). The Mineral and Rock Resources of Ghana. A.A. Balkema, Rotterdam, Netherlands, 610 pp. ISBN: 9061915899.

Li, J., Wang, Y., Xie X., Su, C. (2012). Hierarchical cluster analysis of arsenic and fluoride enrichments in groundwater from the Datong basin, Northern China, Journal of Geochemical Exploration 118(1–3): 77–89.

Lund, J.W, Sanner, B., Rybach, L., Curtis, R., Hellstrom, G. (2003). Ground-source heat pumps- A world overview. Renew Energ World 6(l4): 218–227.

Ma, T., Li, C., Lu, Z. (2016). Geographical environment determinism for discovery of mineral deposits. Journal of Geochemical Exploration 168: 163–168. DOI:10.1016/j.gexplo.2016.07.001.

Matlick, J.S. and Buseck, P.R. (1975). Exploration for geothermal areas using mercury: a new geochemical technique. In: Second United Nations Symposium on the Development and use of Geothermal Resources, Volume I: 785-792.

Moradpouri, F. and Ghavami-Riabi, R. (2020). A multivariate geochemical investigation of borehole samples for gold deposits exploration, Geochemistry International 58(1): 40-48. https://doi.org/10.1134/S0016702920010103.

Moradpouri, F. and Hayati, M. (2021). A copper porphyry promising zones mapping based on the exploratory data, multivariate geochemical analysis and GIS integration, Applied Geochemistry 132: 105051, https://doi.org/10.1016/j.apgeochem.2021.105051.

Moradpouri, F., Ahmadi, S.M.H.,Ghaedrahmati, R., and Barani, K. 2023. Determination of the erosion level of a porphyry copper deposit using soil geochemistry. Journal of the Southern African Institute of Mining and Metallurgy 123(2): 103–112.

Moradpouri, F., Moradzadeh, A., Cruz Pestana, R., Soleimani Monfared, M. (2017). An improvement in RTM method to image steep dip petroleum bearing structures and its superiority to other methods. Journal of Mining and Environment 8(4): 573-578.

Mwangi, M.M. (2013). Application of geochemical methods in geothermal exploration in Kenya. Procedia Earth and Planetary Science 7: 602-606. DOI: 10.1016/j.proeps.2013.03.220.

Najafi, G. and Ghobadian, B. (2011). Geothermal resources in Iran: the sustainable future. Renewable and Sustainable Energy Reviews 15(8): 3946-3951. DOI: 10.1016/j.rser.2011.07.032.

Openshaw, R.E. (1983). Hg and As soil geochemistry of the Meager Creak Geothermal Area. Journal of Geochemical Exploration 19: 339-344. https://doi.org/10.1016/0375-6742(83)90027-4.

Pan, T., Zuo, R. and Wang, Z. (2023). Geological Mapping via Convolutional Neural Network Based on Remote Sensing and Geochemical Survey Data in Vegetation Coverage Areas, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing16: 3485-3494.

Peh, Z., Halamić, J. (2010). Discriminant function model as a tool for classification of stratigraphically undefined radiolarian cherts in ophiolite zones, Journal of Geochemical Exploration 107: 30-38. DOI: 10.1016/j.gexplo.2010.06.003.

Ranasinghe, P.N., Fernando, G.W.A.R, Dissanayake, C.B., Rupasinghe, M.S., Witter, D.L. (2009). Statistical evaluation of stream sediment geochemistry in interpreting the river catchment of high-grade metamorphic terrains, Journal of Geochemical Exploration 103: 97–114. https://doi.org/10.1016/j.gexplo.2009.07.003.

Jimoh, R. O., Olatunji, A.S., Ajadi, J., Afolabi, A.O. (2023). Mineralogy and Geochemistry of Beryl-Bearing Prgmatite Dykes from Gbayo, Southwestern Nigeria. Jordan Journal of Earth and Environmental Sciences 14 (2): 91-102.

Risdianto, D. and Kusnadi, D. (2010). The Application of a Probability Graph in Geothermal Exploration. Proceedings of World Geothermal Congress. Bali, Indonesia, 25–29.

Rowley, J.C. (1982). Worldwide geothermal resources. In: Edwards LM et al (eds) Handbook of geothermal energy. Gulf Publishing, Houston, pp 44–176, Chapter 2. ISBN 0-87201-322-7.

Rybach, L. (2010). The future of geothermal energy and its challenges. In: Proceedings world geothermal congress, Bali, Indonesia, 25–29. April 2010.

Seyedrahimi-Niaraq, M., Ardejani, F.D., Noorollahi, Y., Porkhial, S., Itoi, R. and Nasrabadi, S.J. (2019). A three-dimensional numerical model to simulate Iranian NW Sabalan geothermal system. Geothermics 77: 41-62.

Shiikawa, M. (1983). The role of mercury, arsenic and boron as pathfinder elements in geochemical exploration for geothermal energy, Journal of Geochemical Exploration 19: 337-338. https://doi.org/10.1016/0375-6742(83)90026-2.

Spadoni, M., Voltaggio, M. Cavarretta G. (2005). Recognition of areas of anomalous concentration of potentially hazardous elements by means of a subcatchment-based discriminant analysis of stream sediments, Journal of Geochemical Exploration 87: 83–91. https://doi.org/10.1016/j.gexplo.2005.08.001.

Steenfelt, A. (1996). Distribution of gold, arsenic, and antimony in West and South Greenland: a guide to mineral exploration and environmental management. Bulletin Grønland Geologiske Undersøgelse 172: 55-61.

Tobore, A., Senjobi, B., Oyerinde, G., Bamidele, S. (2023). Geospatial Soil Suitability Assessment for Maize (Zea mays) Production in Derived Savanna of Agricultural Research and Training, OYO STATE, Nigeria. Jordan Journal of Earth and Environmental Sciences. 14 (1): 9-18.

Yaisamut, O., Xie, S., Charusiri, P., Dong, J., Wen, W. (2023).

Prediction of Au-Associated Minerals in Eastern Thailand Based on Stream Sediment Geochemical Data Analysis by S-A Multifractal Model. Minerals 13: 1297. https://doi.org/10.3390/min13101297.

Yousefi H., Noorollahi Y., Ehara S., Itoi R., Yousefi A., Fujimitsu Y., Nishijima J., Sasaki K. (2010). Developing the geothermal resources map of Iran. geothermics 39: 140-151.

Zhang, S.E., Nwaila, G.T., Bourdeau, J.E., Ashwal, L.D. (2021a). Machine learning-based prediction of trace element concentrations using data from the Karoo large igneous province and its application in prospectivity mapping. Artif. Intell. Geosci. 2: 60–75. https://doi.org/10.1016/j.aiig.2021.11.002.

Zhang, S.E., Nwaila, G.T., Tolmay, L.C., Frimmel, H.E., Bourdeau, J.E. (2021b). Integration of machine learning algorithms with gompertz curves and kriging to estimate resources in gold deposits. Nat. Resour. Res. 30: 39–56. https://doi.org/ 10.1007/s11053-020-09750-z.

Zhu, B., Zhu, L., Shi, C., Yu, H. Wang, G. (1989). Application of geochemical methods in the search for geothermal fields. In: Xie Xuejing and S.E. Jenness (Editors), Geochemical Exploration in China. J. Geochem. Explor., 33: 171-183. https://doi.org/10.1016/0375-6742(89)90027-7.

Zuo R. (2011). Identifying geochemical anomalies associated with Cu and Pb–Zn skarn mineralization using principal component analysis and spectrum–area fractal modeling in the Gangdese Belt, Tibet (China, Journal of Geochemical Exploration 111: 13–22. DOI: 10.1016/j.gexplo.2011.06.012.

Zuo, R., Cheng, Q., Agterberg, F.P. and Xia, Q. (2009). Application of singularity mapping technique to identify local anomalies using stream sediment geochemical data, a case study from Gangdese, Tibet, western China. Journal of Geochemical Exploration 101: 225–235. DOI: 10.1016/j.gexplo.2008.08.003.