ISSN: 0974-5823 Vol. 6 No. 3 December, 2021

*International Journal of Mechanical Engineering*

# Clustering of Sabzpushan anticline joints by fuzzy method and comparison with classical methods in the geology with the study of the elements in southwestern Iran

**Robabe Shahamiri[1], Kouros Yazdjerdi[2]*, Abdolmajid Asadi[3], Vahid Ahmadi[4] and Mohammadreza Mozaffari[5]**

1,2,3,4 Department of Geology,ISLAMIC Azad University,Shiraz,postcode :74731-71987,Iran.

5 Department of Mathematics,ISLAMIC Azad University,Shiraz, postcode :74731-71987,Iran .

*Corresponding author

Kouros Yazdjerdi

***ABSTRACT***

Sabzpushan anticline is a folded transverse fault zone in the south city of Shiraz, Iran. Joints study of this area and data clustering method was done by using fuzzy k-means clustering algorithm.

The results obtained from the cluster validity graph and clustering algorithms are the best clustering for regions k=4 and for each flank anticline k=3, where k is the number of cluster.Comparison and clustering of classical methods procedure can obtained the same results for each group. There are minor differences that are due to overlapping zone for the algorithm, which lead to a change in calculation of average orientations.

The studied area is covered with Asmari-Jahrom, As, Co , Cr ,Cu , Ni elements are positively correlated with each other and nickel is enriched with Clarke.

*Key words: joints , K-means Algorithm , cluster, iran ,Zagros,element ,Sabzpushan.*

## *1.INTRODUCTION:*

Joints are plan fractures which are without cohesion and cannot be seen slipping. These fractures are structures of various sizes. Usually, they are seen in a certain area. The study of joints is very important. The most common studies are based on the orientation of the joints.

The most common studies of these types of geological structures is base on strike, dip and its dimensions that are expressed as area and not length of a line, is determined. The most common studies are based on strike, these studies are mostly in the vertical joints. A large percentage of the joints are oriented in an area called the joint sets.

These sets which are based on statistical analysis of data concerning the orientation of the joints of a region can be identified. The study of joints is very important in many engineering works (tunnels, bridges, dams, etc.). Checking the condition of joint can be a correct tip for selecting civil works and construction.

The study of joints can be helpful for finding mineral resources, movement of water and studying the water flow in the region. In the science of structural geology, the study of joints can be determined as stress to the trends of a region or part of the rock.

The joints generally have a key role in reshaping the face of the Earth's crust and in identifying trends and distribution and categorizing them to effectively help in identifying suitable channels for the movement of groundwater and surface waters.

The method of studying classified structures is based on classical methods and statistical studies. Most statistical studies use rose diagram, stereographic plots projection and its pole.

Moreover, joints trend, can be used as features together in a classified data, these fractures are studied by plotting graphic stereo images with different programs.

Generally, with general interpretation of these images and rose diagrams, the general trend of joint groups and the number of cluster can be interpreted from images.

The more the number of joints in the data, the more accurate the result achieved. However, there may be an overlap between clusters of images which can be seen in graphics and it will be ignored in classical studies.

There are Outliers joints data and incorrect joints in field observations, which cannot be accurately determined.

In addition to the classical methods that can be used today, mathematical methods can also be used in data clustering.

Clustering is one of the branches of learning without supervision. An automated process whereby samples are similar to clusters whose members are divided into this category are called clusters. The similarity of these clusters can be considered for various criteria and indicators.

Cluster analysis is the art of partitioning a given set of data into similar clusters (groups, subsets, classes),

where the partitions should have the following two properties:

(a) Homogeneity within the clusters,

(b) Heterogeneity between clusters [1].

There are two scopes for clustering: hard clustering and fuzzy clustering. In hard clustering, each point of a data set is assigned to exactly one cluster, while in fuzzy clustering; each point of the data set belongs to several clusters as a matter of degree in [0, 1](Fazel Zarandi et al ,2010). The important issue in clustering is the number of clusters selected .

The use of different algorithms helps in choosing the number of clusters. In some algorithm, there are pre-specified number of clusters while in some algorithm, the algorithm itself decides which data are divided into several clusters.

There is classic and fuzzy clustering algorithm which is used today in various sciences of fuzzy algorithms for clustering.

The classic algorithm for each input sample is only a member of one and only one cluster and the clusters are in fact not overlap in classical algorithm. However, in fuzzy algorithms, a sample can belong to more than one cluster if the amount or degree of membership given to them is between zero and one, but in classical algorithms, members and prominent ones have a zero or one. Using this data fuzzy clustering algorithm in categories with the cluster validation methods, gives better and complete result.

Algorithm used for studying Sabzpushan anticline with cluster validation has been done. K-MEANS algorithm is a clustering regional algorithm. The advantage of this algorithm is speed and implementation. The data space is divided into K non-empty subset. The initial points are selected randomly and the average point clusters are calculated. Each cluster points to the nearest point average than it has been assigned. The calculation of the average cluster returns to the stage again and proceeds until the rest of the new clusters continue.

The aim of the Fuzzy K-means method is to divide the discontinuity sets in K families through the minimization of the distance from the discontinuities and the center of the groups, seeking for regions with a high density of elements (Andre Monteiro Klen, Milene Sabino Lana, 2014).

K-means procedure is one of the methods of clustering discontinuities such as joints as demonstrated by Hammah & Curran (1998).

Statistical tools such as discriminant analysis, regression, decision analysis and cluster analysis exist for the exploratory analysis of data. More recently, artificial intelligence techniques like neural networks have also been developed for the purposes of analyzing and interpreting data. Nevertheless, in an environment where data has to be classified into homogeneous groups of objects in the absence of *a priori* information on the groups, cluster analysis is the tool most suitable for application (Hammah and Curran, 1998).

Since the study of discontinuity in the active fault regions is very important, Sabzpushan anticline and a comparison between classical and fuzzy clustering method of joint studies in the area have been carried out.

## 2.Area position:

Active fault zones include areas that have large discontinuities and fractures, including joints and faults. The study of fractures for determining stress induced main fault and function is very important

and can express activated correct interpretation of the main fault.

Sabzpushan anticline is a folded transverse fault zone in the south city of Shiraz, Iran.This anticline is in Shiraz segment, and this segment is one of the several segments that are present in Sabzpushan active fault zone. Sabzpushan fault zone is a transverse fault with azimut 155- 170 N, one of the transverse-shear faults located in the Zagros folded.


Fig.1. The position of the study area(National Online Project,Google earth),red line is shiraz segment and dehdari segment.The study area is shown With rectangular.

This anticline is an anticlinorium in the south west of Shiraz in the region of Zagros folded anticline. Transverse fault has an important role in the distribution of deformation of the Zagros Mountains. These faults consist of several fault segments that influenced this region and has numerous fractures and folds.

Sabzpushan fault zone is composed of 6 anticlines that have some faults. It forms a ladder from the northwest of the Shiraz city to the southeast and continues to Ghir. This zone is an anticlinorium fault that consists of 6 anticlines. Based on the results of field observations and utilizing satellite imagery, approximate length of fault is 220 km and the width fault is 13 km. Most of their direction is N00-15W for ladder relative to each other (Safari et al., 1999).

Unlike neighboring anticlines, trending Sabzpushan anticline axis does not follow Zagros Mountains general direction and rotates clockwise.

The length of this anticline is 55 km and the highest elevation is 2822 m. The height difference from the valley to the highest point of the anticline is 1200 m. The rate of weathering and erosion in the area of Sabzpushan is very high.

Existence of clastic rocks indicates high temperature difference between night and day. By examining

aerial photographs, the curvature axis of the Sabzpushan anticline was well seen in the region.

Two main segments of Sabzpushan anticline fault are seen and a lot of fractures seen in anticline. The main Sabzpushan anticline axis splits in the north-west and this is broken in different places.

Although the northern half of the North-West is cut by numerous fractures, there is still a simple fold. A large number of springs are present in the north-eastern Sabzpushan mountain flank, which has a high electrical conductivity of same fractures. A study of neotectonic zone shows that the zone has also been active in the quaternary.

Regional seismic survey and evidence such as drainage pattern show that most of the faults and joints are consistent with a fault zone that has a very high fault scarp, tilting quaternary deposits. There exist systematic joints that can be seen in deposits locally and the existence of a lot of debris which indicate active fault zone in Quaternary period.

One sign of activity in the region is the existence of some Dehdari segment abyss that can be properly seen from the satellite imagery.

The anticline of the study area has been studied alongside the southern flank joints of the structure. The formation type of study area is Asmari-Jahrom and Razak.

In North and South flank anticline, the joints are abundant, and because both flanks are adjacent to farmland and southern flank is located in the plain adjacent to the Syakh- Darngun, the plain is watered by the River Ghareaghaj.

Despite these joints and fractures, it is very important to conduct surface water and ground water investigation in the area.

Moreover, this anticline is close to Shiraz metropolis which has increased the importance of studying them.

## 2.STUDY METHOD:

Joint examination of the anticline was carried out using the classic and fuzzy methods; a comparison of the results was carried out at the end. Careful examination of classified joints and cluster in the region is of great importance.

In reviewing classified field observation fractures, joints in the region in both the northern and southern flank anticline were measured.

The fractures in the suitable measuring stations, the number of single joints are to be measured as much as possible regardless of their size.

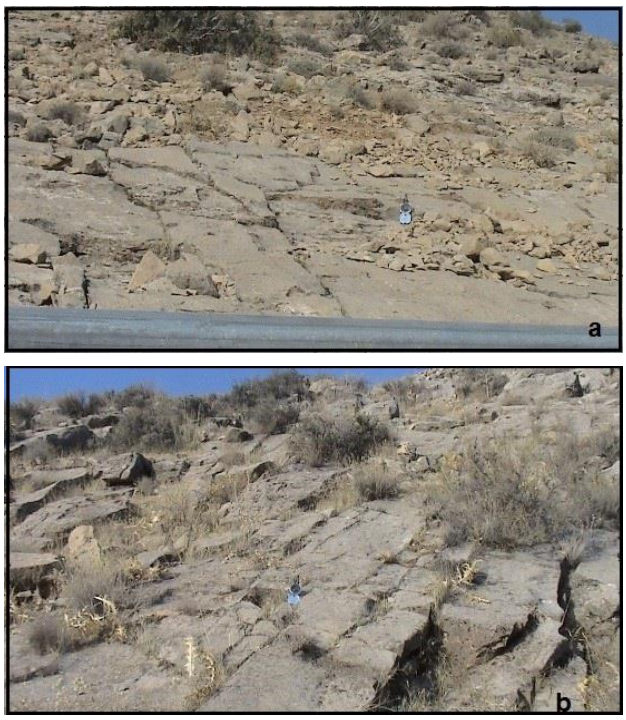


Fig2:North and south limb anticline joints.

In classifying data discontinuity in the region, north and south flank anticline was done separately. The number of clustering of joints will be determined in each domain. Because they are next to each flank domain, on the basis of separate joints, studying the scope of surface water and groundwater is of particular importance for the region. After measuring, the first joints were drawn with joints pole using DIPS program, Thereafter, best classified joints was performed using mathematical methods and fuzzy algorithm.

One of these algorithms, which are k-means algorithms for clustering discontinuities are used especially in geological sciences.

Classification of data in different ways is very high. A branch of learning is unsupervised clustering and an automated process whereby samples are similar to the categories that divide its members into this category is called clusters.

The same can be considered for various criteria clustering from the main problems which depend on the number of clusters.

In some algorithms, some other pre-specified number of clusters algorithm decides which data are divided into several clusters.

There are classic algorithms and fuzzy algorithms, but today the fuzzy algorithm is used in most fields.

First, data obtained were analyzed at several stations and joints poles were drawn with Dips Software for geological data and then at a later stage, work was done separately for each anticline limb.

Thereafter, using mathematical algorithms and validation of algorithms, the K-Means clustering for data was demonstrated.

The results of applying clustering algorithms on a data set according to the parameters of the algorithm can be very different from each other.

The purpose of clusters validation, finding clusters that best fit their desired data, the measure is expressed in various sciences.

Validity indices in a fuzzy environment can be divided into three groups (Fazel Zarandi et al, 2010):

1. Indices involving only the membership values,

2. Indices involving the membership values and data set,

3. Other approaches for fuzzy cluster validity.

However, Hammah & Curran (1998) indicates that for validity, fuzzy algorithms in the joints category in their study must be expressed with indicators expressed in this work; and we tried to see that Sabzpushan has the best clustering in the region.

The Fuzzy K-means requires a priori information about the number of families so that the sets can be partitioned. Nevertheless, in many cases, the information is little or does not exist. This problem is one of the biggest challenges in the cluster analysis, and the appropriate solution is using the clusters validity indices (Andre Montero Klen, Milene Sabino Lana; 2014).

The study of Sabzpushan anticline is used for choosing optimal sets of clustering validation. Measures obtained from the graph are checked using ICC, FS, VMPC, VPBMF, XB, VW.

ICC: This measure of validation (Inter Class Contrast) was developed by Franco (2002), and it takes into consideration the separation and compactness of the formed clusters. Beyond addition, it is shaped to detect very close allocated centers.

$V_{PBMF:}$ This measure of validation was developed by Pakhira et al. (2004) and among those tested by Zhang et al. (2008). VPBMF assumes that the best partitioning of the data set is the one that has the lowest number of groups with the greatest compactness and possible separation.

FS: The validity function proposed by Fukuyaman and Sugeno (1989) measures the compactness and the seperation.

$V_W$: Vw index developed by Zhang et al. (2008) is one of the preferred, since it is based on the concept that the number of clusters should be maintained as small as possible, while the variability is reduced to the maximum.

XB: XB index is developed by Xie and Beni (1991). This measure uses the ratio of the numerator and the number N of elements to estimate compactness of fuzzy partition, while the separation is established by calculating the shortest distance between the centers of the clusters.

In studies carried out by Andre Monteiro Klen on joint for data series, when only two indicators; XB and VMPC were used, Outliers common points with the index stereograms can be seen. The first series of Sabzpushan anticlines data that is collected is around 600 and the different stations are collected in both the north and south limb.

**table.1. *summary* of cluste**r validity.

Criteria for classifying data have been set using information that includes joint orientation azimuth, dip direction and dip. According to Figure 3, pole diagram showing where the joints occur is drawn utilizing DIPS software.

Then using the K-MEANS algorithm, data series is performed for k values between 2 and 7 for each cluster analysis.

| CLUSTER VALIDITY | CRITERIA |
|---|---|
| $XB = \frac{\Sigma_{j=1}^{K}\Sigma_{i=1}^{N} u_{ij}^{m}\left[1-(x_i \cdot v_j)^2\right]}{N\left(\min_{j\neq k}\left[1-(v_j \cdot v_k)^2\right]\right)}$ | MIN |
| $Fs = \sum_{j=1}^{k}\sum_{i=1}^{N} u_{ij}^{m}\left(1-(x_i \cdot v_j)^2\right) - \sum_{j=1}^{K}\sum_{i=1}^{N} u_{ij}^{m}\left(1-(v_j \cdot \bar{v})^2\right)$ | MAX |
| $V_{MPC} = 1 - \frac{K}{K-1}\times\left[1-\left(\frac{1}{N}\sum_{j=1}^{K}\sum_{i=1}^{N} u_{ij}^{2}\right)\right]$ | MAX |
| $ICC = \frac{\Sigma_{j=1}^{k}\Sigma_{1}^{N} u_{ij}\left[1-(V_j \times \bar{V})^2\right]}{N} \times D_{\min} \times \sqrt{k}$ | MAX |
| $v_w = \frac{\left[\Sigma_{j=1}^{K}\Sigma_{i=1}^{N} u_{ij}\left[1-(x_i \cdot v_j)^2\right]/n_j\right]\times\left(\frac{k+1}{k-1}\right)^{1/2}}{\max_{j\neq k}\left[1-(v_j \cdot v_k)^2\right]}$ | MIN |
| $VPBMF = \frac{1}{K}\times\frac{E_1}{J_m}\times D_C$ | MAX |

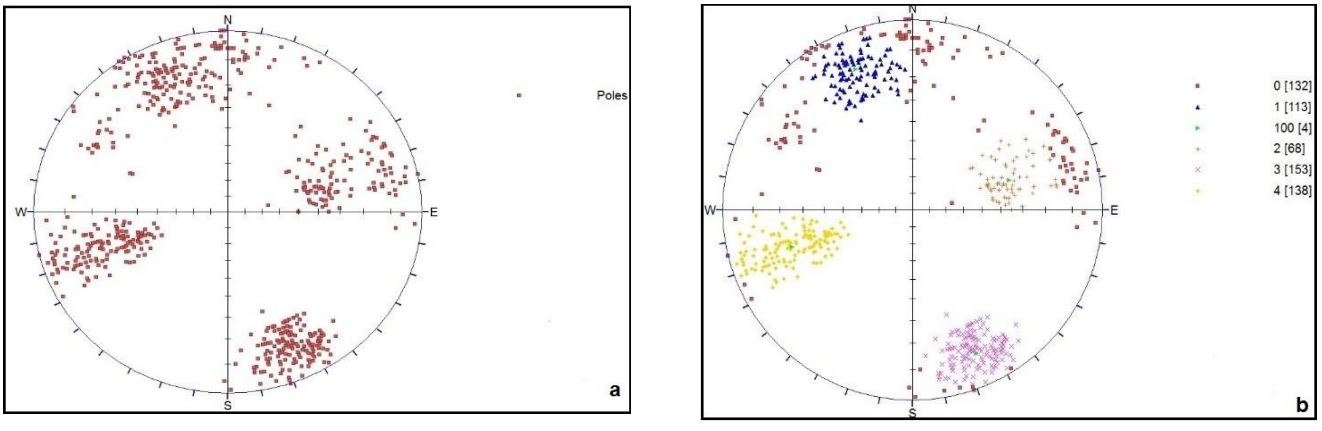


Fig.3. J0ints Pole stereographic projection-classic method(a)&fuzzy method (b)

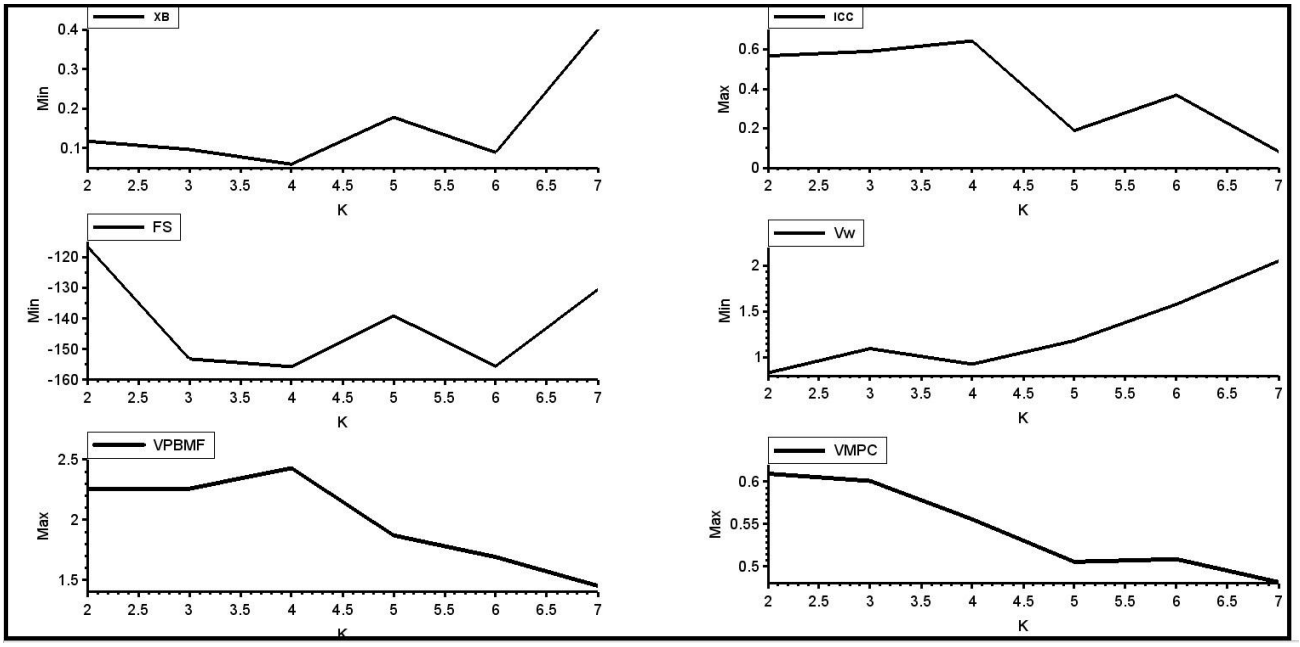

fig.4.Graphe of cluster validity indices for sabzpushan anticline data set.

For each number of clusters, K-MEANS algorithm was run multiple times and performance values were averaged. The values of the validity indices for Sabzpushan anticline indicate the optimal number of joint sets to be 4(fig4).

Using the best clustering, algorithm was done by drawing pole classified data outliers, and data showing the overlapping clusters are identified in Figure (3).

Then, on the edge of the northern and the southern limb, clustering is done separately. First pole joints of each limb flow diagram using the Dips program are shown in Figure (5&6).

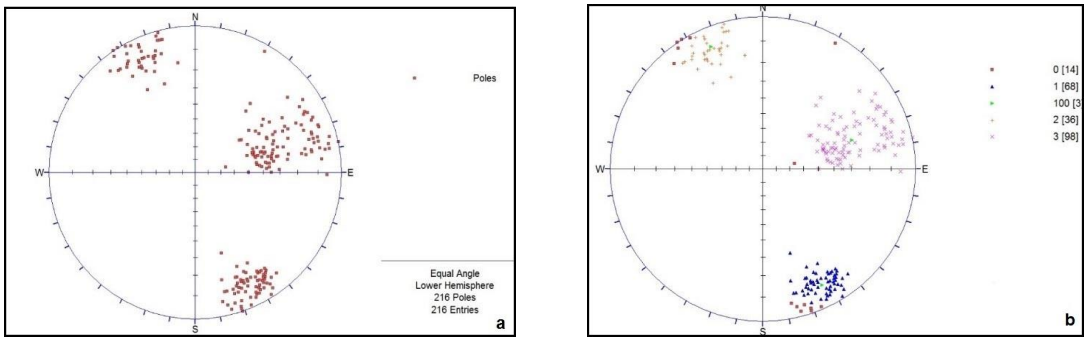

Fig.5. Anticline north limb J0ints Pole stereographic projection-classic method(a)&fuzzy method (b).

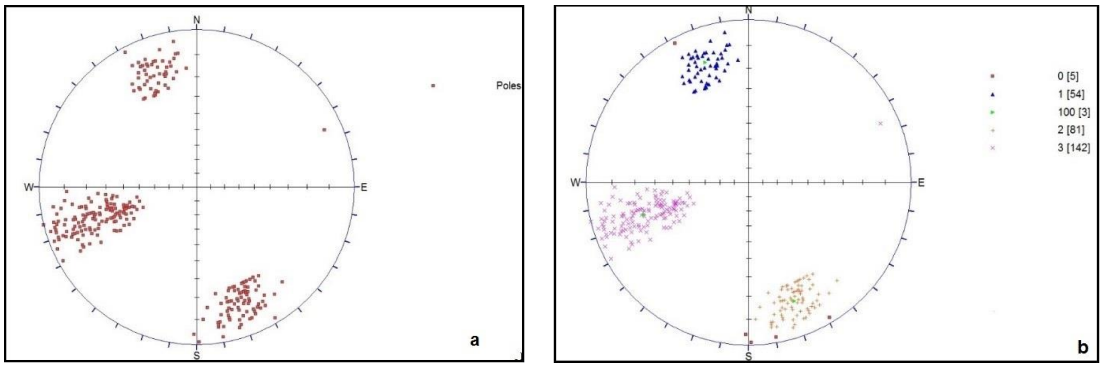

Fig.6. Anticline south limb joints Pole stereographic projection-classic method(a)&fuzzy method (b)

Then an algorithm running is done separately on each limb anticline .With the graphic running the algorithm on each limb, the best clustering in the north and south flank is K=3 (Figures 5 and 6).

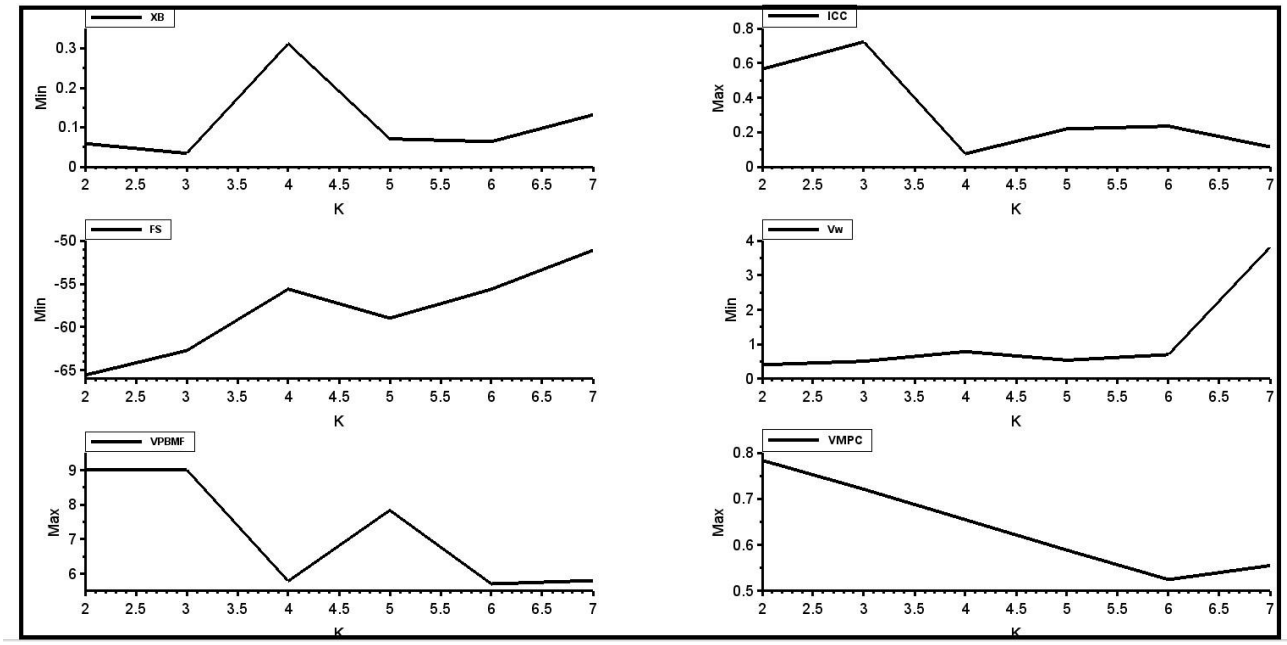


fig.7.Graphe of cluster validity indices for north limb sabzpushan anticline data set.

The values of the validity indices for north Sabzpushan anticline indicate the optimal number of joint sets to be 3 (fig7).

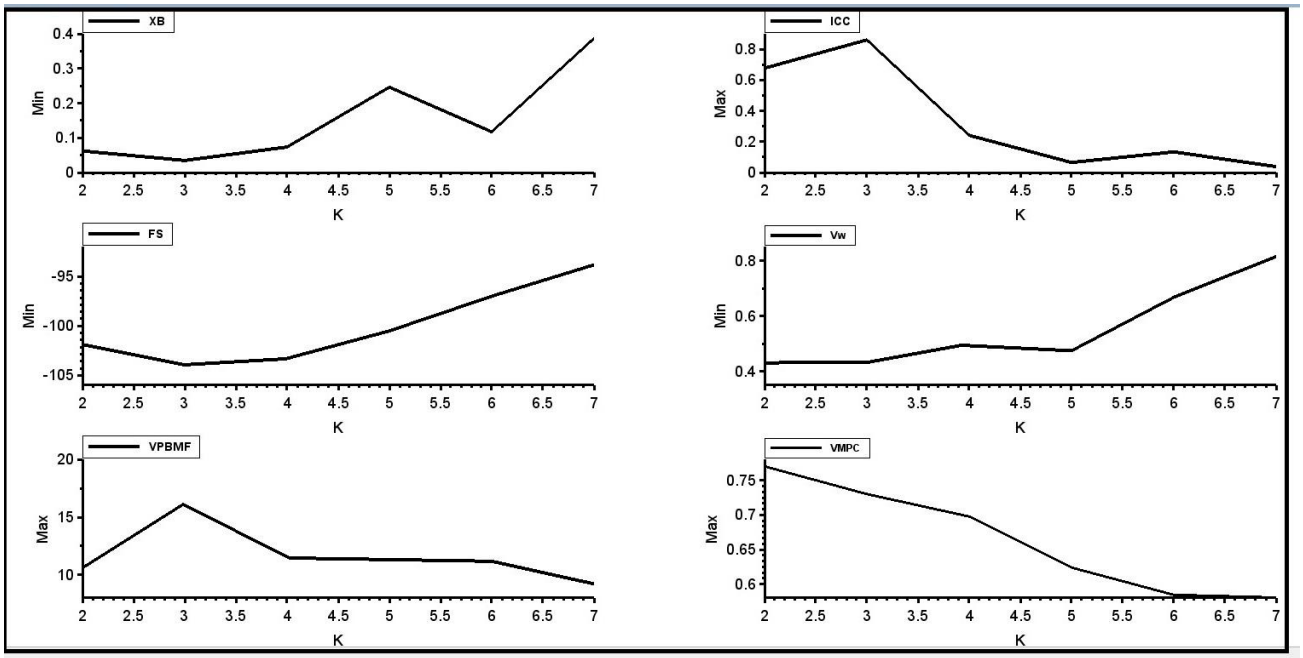


fig.8.Graphe of cluster validity indices for south limb of sabzpushan anticline data set.

The values of the validity indices for South Anticline Sabzpushan indicate the optimal number of joint sets to be 3(fig.8).

After clustering data, pole stereographic projection was drawn. There are 3 joint clusters in each Sabzpushan anticline limb. Overlapping zone and outlier is shown by plotting clustering data stereographic projection in Dips program.

## 3. STUDY GEOCHEMICAL OF ELEMENTS OF THE STUDY AREA :

According to studies conducted by Ansari et al.2012 , the average of elements is higher than the Clark limit in the area in most cases. In this study, the concentration chart of the desired elements was prepared in Shiraz plain by type of formation. The formation is the Asmari-Jahrom in this Anticline , which is based on studies of its nickel content is much higher than other elements .

The degree of accumulation of elements in a soil depends on the type and degree of erosion (Pluto, 2005)

**Table. 2 . Mean elements of soils in Shiraz plain (Ansari et al,2012 )**

| formation | As | Co | Cr | Cu | Ni |
|---|---|---|---|---|---|
| Unit | ppm | ppm | ppm | ppm | ppm |
| DI | 0.5 | 1 | 1 | 1 | 1 |
| Clark | 1.8 | 25 | 100 | 55 | 75 |
| jahrom | 2.1 | 5 | 39 | 16 | 41 |
| Asmari | 8.22 | 12.07 | 141.85 | 17.23 | 84 |
| Asmari-Jahrom | 8.2 | 13.5 | 153 | 13.5 | 161.5 |
| Razak | 5.1 | 14 | 116.25 | 19.5 | 103 |
| Aghajari | 5.1 | 17 | 611 | 21 | 171 |
| Bakhtiari | 4.25 | 14 | 294.5 | 17.5 | 113 |

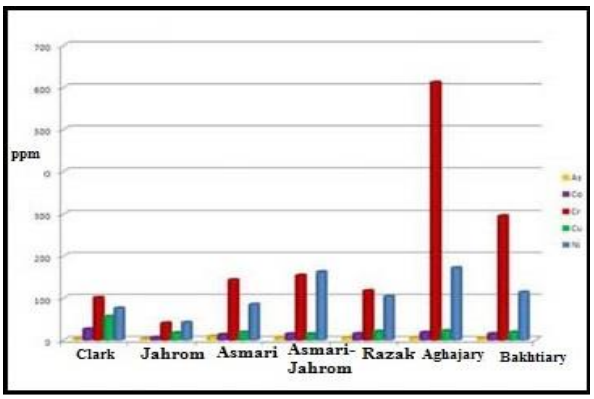


**Figure 9.Bar graph Mean elements of soils in Shiraz plain (Ansari et al,2012 )**

The arsenic element is rich in enrichment than clark in the whole region of the soil.Chromium is enriched from Jahrom, Razak, Aghajari and Bakhtiari formations in soils . Nickel element has enrichment in all soils of the region except the soil from the Jahrom Formation . Finally, two elements of cobalt and copper are not richly enriched in the plains of the plain .

Detection of interconnection and correlation between different elements can help to more accurately assess environmental impacts .Ansari et al., 2012, used the Pearson correlation coefficient to determine the relationship between Heavy Elements in Shiraz Plain (Pearson Correlation) .Determining these relationships can be useful in identifying the source of the element. The high correlation between the elements in the substrate can be explained by several factors: surface absorption in clay minerals and organic matter, the presence of elements in the structure of cinnamins, especially clays, the absorption of elements by oxides and iron oxides and manganese.The correlation coefficients between heavy elements show that the elements of arsenic, cobalt, chromium .They have a positive correlation with each other, which is identical, free, and precipitated in almost identical terms . Nevertheless, the concentration of five precipitates in the Shiraz plain has not exceeded the limit and has Pollution is not human-induced .

## 4.CONCLUSION:

Clustering and studying of joints is very important in many civil works and exploration. Today these fractures and discontinuities are studied in different methods that have surpassed classical and statistical methods.

Performance measures of the partitioning results of fuzzy K-means algorithms allow the data analyst to obtain an idea of which particular clustering (or clusterings) is best or sensible ( Hammah & Curran, 1998).

If doubts were created between two clusters in Clustering Validation, a hundred weathering of other joints characteristics can be obtained such as the joints and more accurate classification can remove doubt and clustering can be optimal.

The use of fuzzy algorithms in the joint study of the Sabzpushan anticline in transverse fault zone is one of the fuzzy methods.

The classical approach in this area was performed with comparison between the two methods.

Clustering of classic and fuzzy method was based on the orientation parameter. When two-way clustering was performed on the whole anticline, 4 joint sets can be seen in the region. The average for each two methods can be seen in Table 1.

In clustering performed separately on both limbs, three cluster joints can be seen in a different direction and the average can be seen for both methods in Table 3.

**Table.3. Clustering results of average orientations**

| Classical | | | | Algorithm | | | |
|---|---|---|---|---|---|---|---|
| Cluster | Total joint | North limb | South limb | signs | Total joint | North limb | South limb |
| 1 | 339/12 | 151/09 | 340/15 | ▲ | 337/13 | 153/09 | 340/14 |
| 2 | 75/37 | 334/06 | 157/15 | + | 74/35 | 334/06 | 159/14 |
| 3 | 155/11 | 77/39 | 252/15 | × | 156/11 | 72/28 | 252/22 |
| 4 | 252/21 | | | ◆ | 252/23 | | |

Comparison is done for the two methods. The fuzzy method can be more accurately clustered using the algorithm, and is characterized by identifying fuzzy method in the overlapping zone.

The use of fuzzy clustering is a powerful tool for the study of joints. This method can identify outliers because they are not in bunch.

The results of area classification of joints should be checked with other results of types of discontinuities and such fractures of the area are to be matched as faults,

Other orientation parameter can be used in clustering the parameters and characteristics of other joints. Whatever it is, more clustering of these features is more accurate.

In discussing the various elements in the samples, it is possible to use different algorithms, such as k-means, to examine fractures. This way, we can use this method to investigate the determination of elements in different samples and clustering them. Like Kayani and Tabatabaei 2017 in surface geochemical samples in one of Iran's deposits .

By using this type of algorithm can be studied in more accurate geochemical paragenesis and elements forming soil layers.

In the study area, an alkaline environment has been caused in the region and the surrounding be cause limestone formation in the anticline, which is causes low mobility of the elements. The formation of elements has caused abnormal increase in heavy metals in the soil but has not exceeded the limit and is not caused pollution yet.

***References***


Andre, M.K. and Milene, S.L., 2014, Fuzzy algorithm of discontinuity sets. International Engineering journal, 67, 439–445.

Ansari,Z., Karimi, M., Esmailbeyg,M,.2012. Geochemical Distribution of Heavy Elements of Shiraz Plains and Their Origins. The 31th Earth Science Conference Geological Survey of Iran .

Farhoudi, G., 1992, A morphotectonic study of environmental impact on ground water in southern Iran and under the Persian Gulf**.** Geologisch Rundschau ,81,581-587 **.**

Fazel Zarandi, M.H., Faraji, M.R., and Karbasian, M., 2010, An exponential cluster validity index for fuzzy clustering with crisp and fuzzy data. International journal of Science an Technology, 17, 95–110.

Hammah , R.E., Curran,J.H.,1998. fuzzy cluster algorithm for the automatic identification of joints sets. International journal of rock Mechanics &Mining science . 35,889-905.

Kiani, F., Tabatabaei, S.H., 2017, Unsupervised classification of surficial Geochemical samples in Zarshuran Gold deposit . The 35th Earth Science Conference Geological Survey of Iran.

Pluta , I., 2005.Mine waters of the Upper Silesian Coal Basin- origin, pollutants and purification, prace Nauk Gyownego Instytutu Gornictiwa 865 (inpolish).

Safari, H. , Ghorayshi, Ghasemi, M. R., 1999. Analyzing Deformation of Sabzpushan shear-

transverse faults zone, south of Shiraz. Geosciences Scientific Quarterly Journal, 8, 31-32.

Shanley, R. J., Mahtab, M. A.,1976 Delineation and analysis of cluster orientation data. Mathematical Geology**.** 8, 9-16.

Zhang, Y., Wang, W., Zhang, X., LI, Y.,2007 A cluster validity for fuzzy clustering. Information Sciences. 178, 1205-1218 .