

FFKM: A SOFTWARE FOR AUTOMATIC DISCONTINUITY SET IDENTIFICATION AND STEREOGRAPHIC PROJECTION ANALYSIS

André Monteiro KLEN^{1*}, Stefano BONDUÀ²,
Sara KASMAEYAZDI³, Danielle Aparecida DE MENEZES⁴

¹Federal Institute of Education, Science and Technology of Minas Gerais, Ouro Preto, Brazil, andre.klen@ifmg.edu.br

²Department of Civil, Chemical, Environmental and Materials Engineering, University of Bologna, Bologna, Italy, stefano.bondua@unibo.it

³Department of Civil, Chemical, Environmental and Materials Engineering, University of Bologna, Bologna, Italy, sara.kasmaeeyazdi2@unibo.it

⁴Federal University of Ouro Preto, Morro do Cruzeiro, Ouro Preto, Brazil, danielle.menezes@aluno.ufop.edu.br

DOI: 10.2478/minrv-2025-0018

received – 24.01.2025 / revised – 10.02.2025

accepted – 18.02.2025 / published – 18.06.2025

Abstract: Discontinuities significantly influence the mechanical and hydraulic behavior of rock masses. Identifying and classifying discontinuities into sets with similar orientations is a crucial task in geotechnical engineering and mining. The most common method for grouping discontinuities involves the visual interpretation of stereographic projection maps. However, this method is limited by subjective factors such as personal experience, the size of the counting circle, variations in the frequency of discontinuity sets, outliers, and overlap. To overcome these limitations, we propose applying the Fuzzy K-Means algorithm based on the Fisher distribution (FFKM), implemented in Scilab as open-source software. This software automatically identifies discontinuity sets and calculates their statistical characteristics. Additionally, it determines the optimal number of sets and generates stereographic projections of the results. To demonstrate the effectiveness of FFKM, it was employed to identify fracture sets in an operational slope in an open-pit iron mine located in the Quadrilátero Ferrífero region in Brazil. The FFKM consistently identified the number of fracture sets and classified them according to the field data, providing reliable and objective results.

Keywords: Fuzzy K-Means algorithm, Discontinuity set identification, Stereographic projection analysis, Open-pit iron mine, Orientation data analysis, Clustering algorithms.

1. Introduction

The term discontinuity generally refers to the collective term for all joints, bedding planes, contacts, and faults [1, 2]. Discontinuities tend to follow patterns, occurring in sets with similar orientations, appearances, and mechanical properties. Parameters such as orientation, spacing, persistence, roughness, filling, aperture, and the number of sets, among others, are important characteristics that influence rock mass stability and are used to group discontinuity sets.

The conventional method for grouping information about discontinuity sets typically relies on orientation and requires visual interpretation of density contours of discontinuity poles on stereographic projections. This process is subjective because the results depend on factors such as personal experience, the size of the counting circle, outliers, the frequency of discontinuities in each set, and overlap among the sets [2,3].

Thus, this conventional method is not satisfactory in some cases, and machine learning techniques, such as cluster analysis, have gradually been adopted to identify discontinuity sets in rock masses [2-10].

Related studies in the literature on algorithms for identifying discontinuity sets are primarily based on clustering approaches, such as Fuzzy K-Means (FKM) and K-Means [5].

The Fuzzy K-Means (FKM) clustering algorithm, introduced by [6] for detecting rock discontinuity sets, represents a significant advancement, as it developed a methodology for calculating both the distances between discontinuities and cluster validity indices. As a result, FKM and K-Means have become widely adopted for clustering discontinuity sets.

* Corresponding author: André Monteiro Klen, Ph.D., Associate professor, Federal Institute of Education, Science and Technology of Minas Gerais, Ouro Preto, Brazil, e-mail address: andre.klen@ifmg.edu.br

Several studies have explored strategies for combining FKM and K-Means with intelligent optimization algorithms. For instance, [7, 8] integrated FKM with genetic algorithms, while [9] applied K-Means in conjunction with swarm intelligence algorithms. In [3], FKM was incorporated with a mutative scale chaos optimization algorithm. These algorithms focus on the automatic characterization of discontinuity sets and the determination of the optimal number of clusters. A detailed literature review of algorithms related to identifying rock discontinuity sets can be found in [1-10].

In this paper, we present the application of the Fuzzy K-Means algorithm based on the Fisher distribution (FFKM) for identifying rock discontinuity sets. Originally developed by [2], the algorithm was implemented in Scilab, an open-source platform designed for scientific and engineering applications. The FFKM was validated by [2] using the methods proposed by [6, 8, 9] as benchmark. It accurately identified the number of discontinuity sets and produced results consistent with those of other methods.

A new functionality was added to generate stereographic projections in the equal-area and upper hemisphere. Additionally, the original version was revised to resolve certain issues. The main features of the algorithm are outlined and demonstrated through a case study using data collected from an operational slope in an open-pit iron mine.

2. Software Description

The FFKM is an unsupervised classification algorithm designed to detect regions with a high density of discontinuities. It uses the Fisher distribution to improve the initialization method of the FKM proposed by [6], eliminating the need for parameter adjustments or additional intelligent optimization algorithms. FFKM retains the fundamental features of FKM and requires only one input parameter: the specification of the maximum number of sets into which the data set should be divided (Fig. 1). The software limits this value to a range of 2 to 8, reflecting the common range of discontinuity sets observed in engineering applications [2-7].

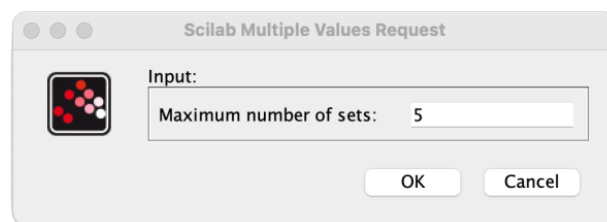


Fig.1. Input window for specifying the maximum number of discontinuity sets

The user selects the data set through an input window, where the discontinuity orientations are provided in a TXT file format (Fig. 2). Additionally, all output files can be saved directly using Scilab, with the added advantage that the optimal clustering results are stored in a CSV file format. This allows the clustering results to be exported and analyzed using different software.

```
// Example of input data file
// Format: Dip Angle, Dip Direction

dip    direction
20     240
32     175
32     180
28     180
38     150
19     210
10     130
38     190
32     180
48     180
25     155
42     155
46     180
24     150
```

Fig.2. Example of an input data set in TXT format

The software automatically identifies the discontinuity sets and generates the following outputs: (i) stereographic projections in equal-area and upper hemisphere; (ii) statistical analysis of the discontinuity sets, such as mean orientation, minimum and maximum dip angles and dip directions, and the number of discontinuities in each set; and (iii) the determination of the optimal number of sets. Compatible with both Windows OS and macOS, the software requires Scilab version 2023.1.0 or higher and is freely available for download at: <https://github.com/andreklen/FFKM-algorithm.git>.

2.1. Software Development Methodology

The FFKM evaluates clustering performance using a fuzzy objective function, where lower values indicate better clustering results. Cluster centers are then recalculated based on the clustering outcomes and membership degrees. This iterative process continues until the cluster centers stabilize, at which point the objective function reaches its minimum value [2, 10].

The application of the FFKM algorithm for clustering rock discontinuity sets is summarized below, with additional details on the algorithm's steps provided in [2]. To aid understanding, a list of key symbols and parameters is included:

N : Total number of discontinuities

K : Number of discontinuity set

K_{best} : Optimal number of discontinuity sets determined by the algorithm

K_{min}, K_{max} : Minimum and maximum number of discontinuity sets

$j = 1, 2, \dots, K$

$i = 1, 2, \dots, N$

V_j : Center of the j^{th} discontinuity set

$e_i = (x_i, y_i, z_i)$: Unit normal vector of the i^{th} discontinuity (equal-area and upper hemisphere)

α : Dip Direction ($^\circ$)

β : Dip angle ($^\circ$)

J : Objective function

S_j : Orientation matrix of the j^{th} discontinuity set

u_{ij} : Membership degree of the i^{th} discontinuity in the j^{th} set

m : Fuzziness indexes defined as 2

XB : Xie-Beni index

$d^2()$: distance metric using the sine-squared measure

x_a, y_a : Coordinates of the poles in the stereographic projection (equal-area and upper hemisphere)

R : stereonet radius is defined as 1

2.2. Main steps of the FFKM algorithm

Input the orientation data of the discontinuities as dip angle, dip direction (Fig.2), and convert them into a unit normal vectors based on Eq.(1) and Fig.3.

$$e_i = [x_i, y_i, z_i] = [\sin(\beta)\sin(\alpha), \sin(\beta)\cos(\alpha), \cos(\beta)] \quad (1)$$

Based on the specified number of K , randomly generate K unit normal vectors to serve as the initial centers of the discontinuity sets, V_j . Refer to the initialization method outlined in [2].

For each e_i , compute its distance to each V_j , using the sine-squared measure defined in Eq.(2).

$$d^2(e_i, V_j) = 1 - (e_i \cdot V_j)^2 \quad (2)$$

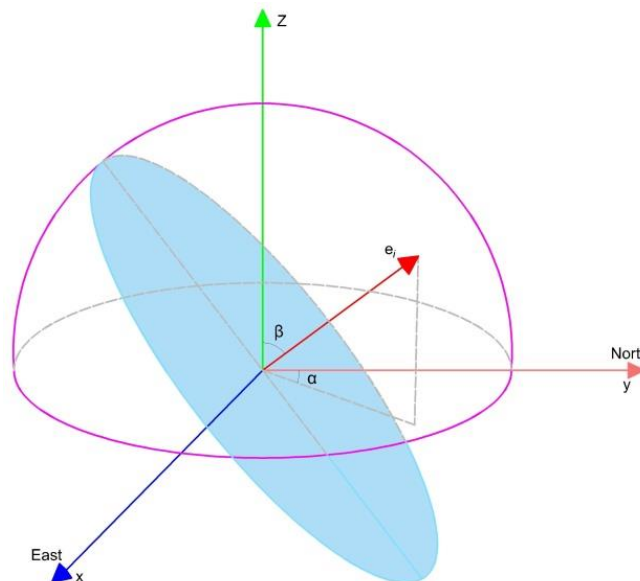


Fig.3. Representation of a rock discontinuity orientation in a counter-clockwise direction

Compute the membership degree, u_{ij} , of each e_i with respect to V_j employing Eq. (3).

$$u_{ij} = \left[\left(\frac{1}{d^2(e_i, V_j)} \right)^{1/m-1} \right] \left[\sum_{t=1}^K \left(\frac{1}{d^2(e_i, V_t)} \right)^{1/m-1} \right]^{-1} \quad (3)$$

Update V_j by performing eigen-analysis method on the orientation matrix S_j , (Eq.4). The eigenvector associated with the largest eigenvalue represents the update center.

$$S_j = \begin{bmatrix} \sum_{i=1}^N (u_{ij})^m x_i^2 & \sum_{i=1}^N (u_{ij})^m x_i y_i & \sum_{i=1}^N (u_{ij})^m x_i z_i \\ \sum_{i=1}^N (u_{ij})^m y_i x_i & \sum_{i=1}^N (u_{ij})^m y_i^2 & \sum_{i=1}^N (u_{ij})^m y_i z_i \\ \sum_{i=1}^N (u_{ij})^m z_i x_i & \sum_{i=1}^N (u_{ij})^m z_i y_i & \sum_{i=1}^N (u_{ij})^m z_i^2 \end{bmatrix} \quad (4)$$

Repeat steps 3-5 iteratively until the maximum change in the membership degree between two consecutive iterations is less than the threshold ($\epsilon = 0.001$).

$$\max_{ij} \left[\left| u_{ij} - \hat{u}_{ij} \right| \right] < \epsilon \quad (5)$$

where u_{ij} and $\hat{u}_{ij}$ are the membership degree of iterations t and $t + 1$, respectively.

Assign each discontinuity to the set based on its highest membership degree, u_{ij} .

Compute the objective function Eq.(6) and apply the Xie-Beni index (Eq.7) to determine the optimal number of discontinuity sets, K_{best} .

$$J = \sum_{i=1}^N \sum_{j=1}^K (u_{ij})^m d^2(e_i, V_j) \quad K \leq N \quad (6)$$

$$XB = \frac{J}{N \times \min_{l \neq m} [d^2(V_l, V_m)]}, \forall l, m \in \{1, 2, \dots, K\} \quad (7)$$

Outputs: Discontinuity sets for K_{min} to K_{max} , including stereographic projections, statistical analysis, XB indices and K_{best} . The pole coordinates, $x_a - y_a$, in stereographic projections in equal-area, upper hemisphere are computed using Eq.(8) as derived from [11].

$$x_a = R\sqrt{2} \left(\frac{\beta}{2} \right) \sin(\alpha) \quad \text{and} \quad y_a = R\sqrt{2} \left(\frac{\beta}{2} \right) \cos(\alpha) \quad (8)$$

The algorithm's flowchart (Fig.4) clarifies the steps involved in the process.

3. Illustrative Example

In this section, the capabilities and functionalities of the software are demonstrated through an example from an operational slope in an open-pit iron mine located along the western edge of the Quadrilátero Ferrífero (QFe) in the central part of Minas Gerais state, Brazil.

The Quadrilátero Ferrífero (QFe) covers an area of approximately 7,000 km² and is one of the world's foremost mineral provinces. With a rich history of over 300 years of continuous exploration, it is particularly renowned for its significant deposits of iron ore, manganese, and gold [11]. The iron reserves in the region are primarily composed of itabirites (Banded Iron Formations) with layers of quartz and iron oxides (mainly hematite and magnetite), with an average iron content of 43.7% [12, 13].

The region's geology is highly intricate but can be broadly divided into three major rock unit groups: the metamorphic complexes of Archean age, the metavolcanic-metasedimentary sequences of the Archean greenstone belt (corresponding to the Rio das Velhas Supergroup), and the Paleoproterozoic supracrustal rocks of the Minas Supergroup [12-15].

In the intermediate portion of the Minas Supergroup lies the Itabira Group, which is subdivided into the lower Cauê Formation and the upper Gandarela Formation. The mine is situated within the Cauê Formation, which is primarily composed of itabirites with alternating layers of quartz and hematite. This formation is of significant importance to the QFe, as it concentrates the majority of the region's iron ore reserves. Approximately 48 iron ore deposits have been mapped in this region [12-15].

As described by [15], the geological distribution of the mine consists predominantly of siliceous itabirites in compact, semi-compact, and friable forms, along with itabiritic breccias and hematites. It is estimated that more than 80% of the deposit is composed of compact itabirite, which is characterized by hard rock with low iron content. Fig. 5 represents a summary geological map of the QFe and the location of the mine.

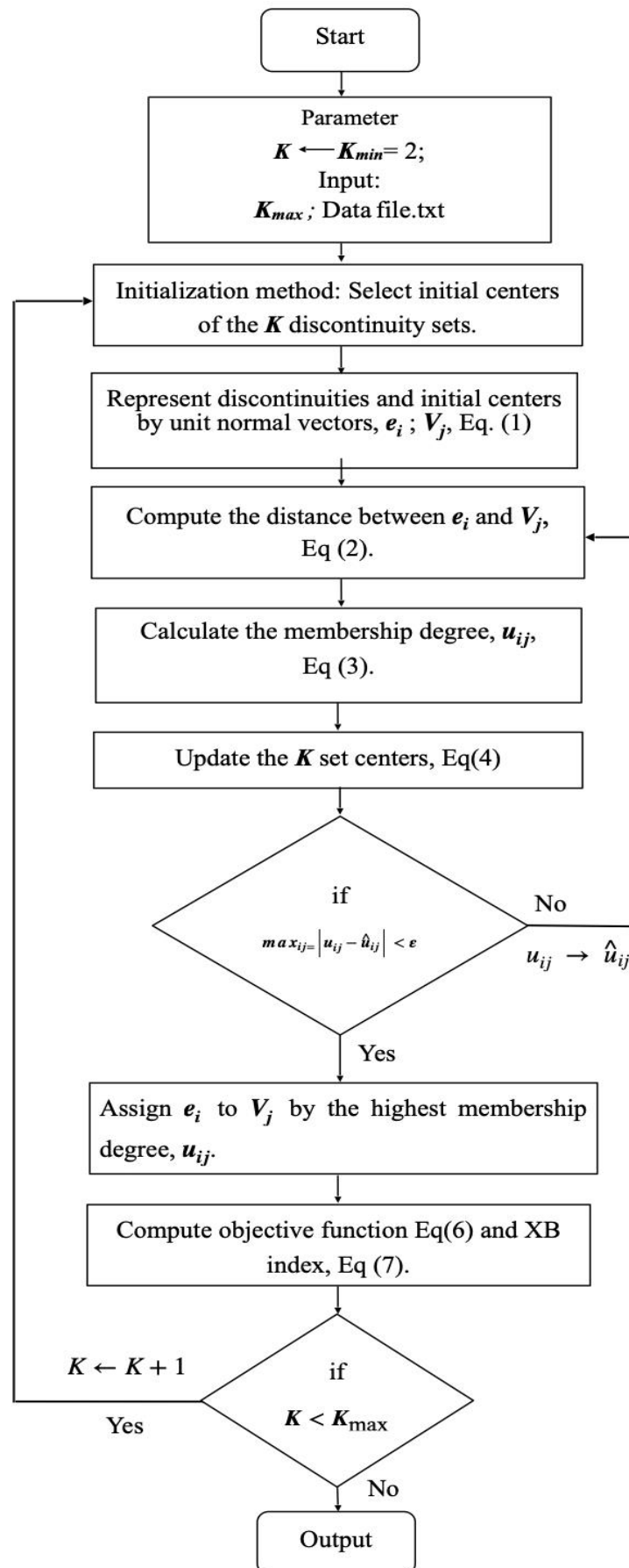


Fig.4. Flowchart of the FFKM algorithm

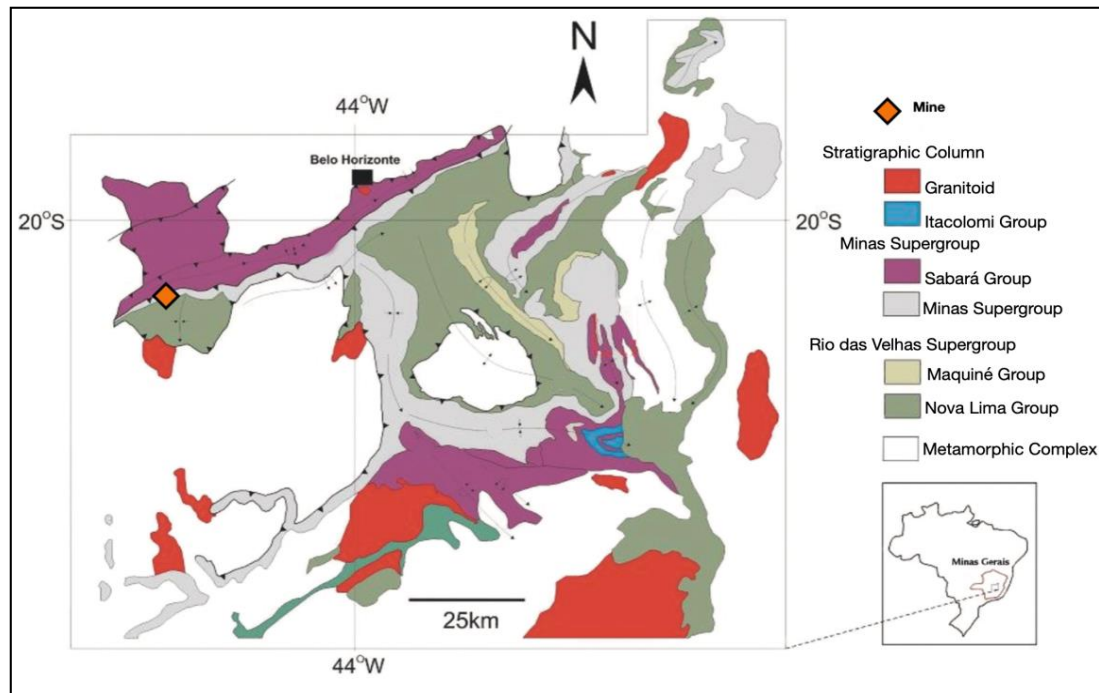


Fig.5. A summary geological map of the Quadrilátero Ferrífero and the location of the open-pit mine. Adapted from [13]

The open-pit iron mine is divided into five operational sectors and the operational slope selected for this example is located as indicated in fig. 6. The slope is highly fractured, with a Rock Quality Designation (RQD) of 37%, indicating poor rock quality. It has a height of 28 meters, an azimuth of 169°, and an inclination of 50°. The predominant lithology of the slope is compact itabirite, transitioning to friable [15].

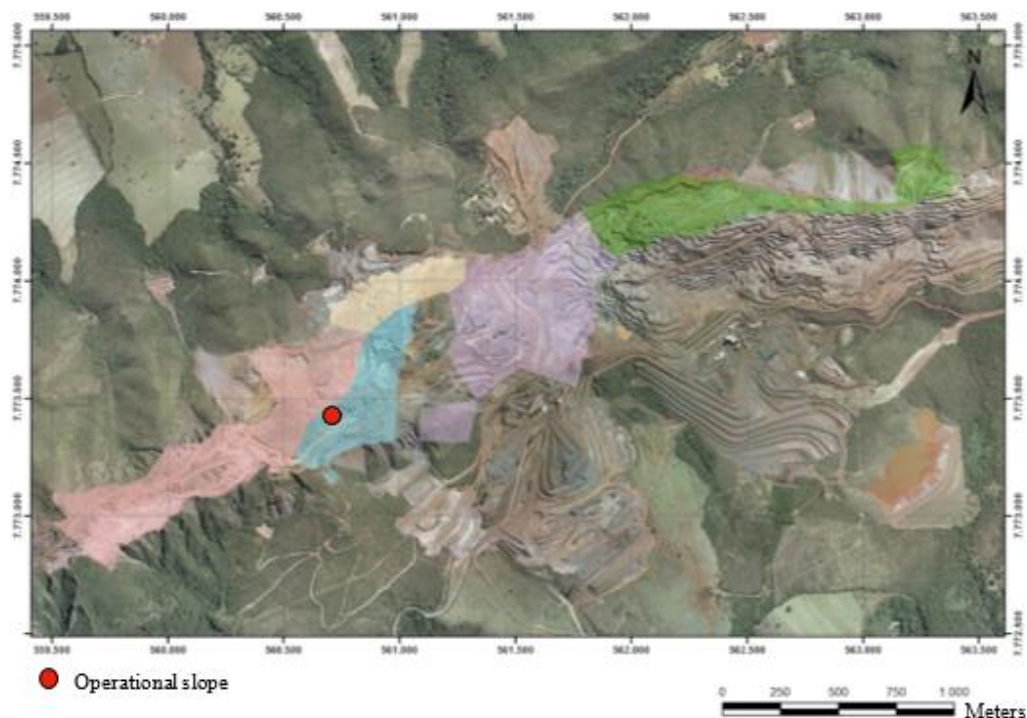


Fig.6. View of the open-pit iron mine and the location of operational slope

A total of 159 fracture orientations were collected from the operational slope using compass measurements and the scanline sampling method. During the geotechnical mapping, three predominant fracture sets were identified and labeled as **Set 1**, **Set 2**, and **Set 3**. **Set 1** is sub-vertical, **Set 3** is parallel to the banding and serves as the main conditioning structure of the slope and the open-pit mine, while **Set 2** dips in the opposite direction to **Set 3**, with a steeper inclination.

Fig. 7(a) shows typical fracture traces observed at the studied slope, along with the different itabirite lithologies. Fig. 7(b), (plotted using the Dips program from Rocscience) presents the contour plot of the fracture dataset, projected onto the upper hemisphere using an equal-area stereographic projection.

Therefore, based on information acquired during the geotechnical mapping, the FFKM must be capable of correctly identifying the number of sets and classifying the fractures. The software was run with $K_{max}=5$, and the XB index was used to determine the optimal number of sets. The XB index [16] simultaneously evaluates the compactness and separation of FFKM partitions. As such, the index achieves a small value when the clustering results are appropriate.

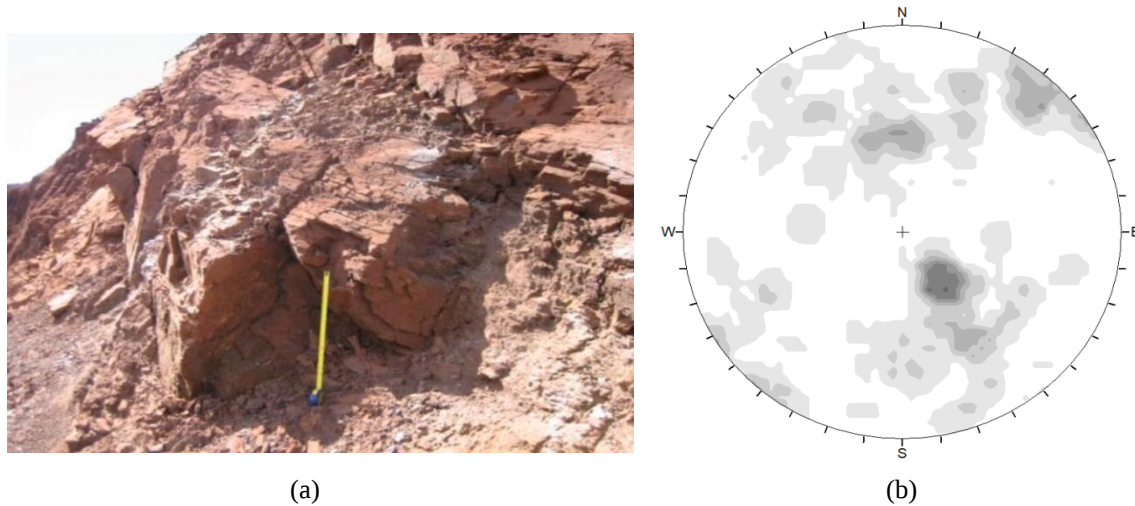


Fig.7. View of the operational slope. (a) Typical fracture traces and associated lithologies, (b) Contour plot of the fracture data set

The output report (Fig. 8) indicates that the XB indices identify three as the optimal number of fracture sets, aligning with the findings from the geotechnical mapping process. It also includes the statistical analysis of the fracture sets for $K_{best}=3$. The Fig.9 provides an example of the clustering results exported in CSV format, while the stereographic projections generated by FFKM for $K_{min}=2$ to $K_{max}=5$ are displayed in Fig.10.

"Mean orientations for K=3"	"Xie-Beni for K= 2"
"Set 1: Dip = 87, Direction = 44"	"0.2193848"
"Set 2: Dip = 45, Direction = 340"	"Xie-Beni for K= 3"
"Set 3: Dip = 39, Direction = 150"	"0.1329306"
"Set 1 count: 47"	"Xie-Beni for K= 4"
"Set 2 count: 53"	"0.1556646"
"Set 3 count: 59"	"Xie-Beni for K= 5"
"Minimum and maximum values"	"0.1678243"
Set: 1	"The optimal number of sets is: 3"
Dip: Min = 42, Max = 90	
Direction: Min = 10, Max = 265	
Set: 2	"The minimum value of XB is: 0.1329306"
Dip: Min = 12, Max = 85	
Direction: Min = 0, Max = 355	
Set: 3	
Dip: Min = 10, Max = 84	
Direction: Min = 70, Max = 225	
"The clustering results were successfully exported to a CSV file located in the Results folder."	

Fig.8. Output report showing the statistical analysis for $K_{best}=3$ and the XB indices values

```
// Example of Output Data File

// Format: Dip Angle, Dip Direction,
// K = 2, ..., K = 5

25,145,2,3,4,5
40,35,1,2,1,1
35,175,2,3,4,5
56,145,2,3,4,2
35,135,2,3,4,5
24,140,2,3,4,5
35,155,2,3,4,5
11,225,2,3,4,5
```

Fig.9. Output clustering results in CSV format

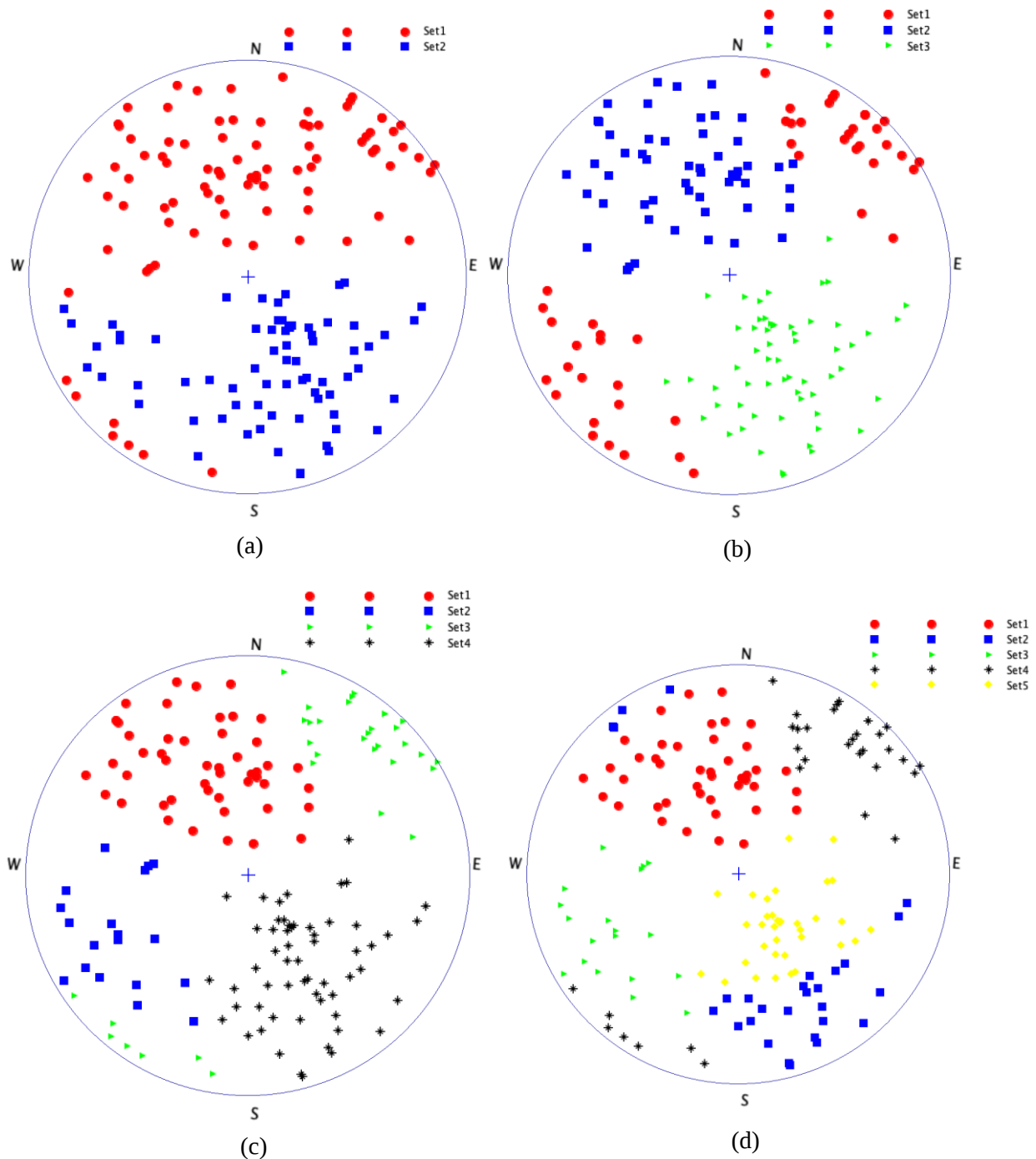


Fig.10. Stereographic projections generated by FFKM. (a) $K_{min}=2$; (b) $K_{best}=3$; (c) $K=4$; (d) $K_{max}=5$

Notably, as illustrated in Fig. 10(b), **Set 1** is sub-vertical, while **Set 2** dips in the opposite direction to **Set 3**. This demonstrates that the FFKM appropriately grouped the fracture sets, identifying the existing relationships between them. The analysis of the minimum and maximum dip angles reveals that the data set contains discontinuities that deviate significantly from the center of the sets. However, one advantage of the FFKM is its capacity to minimize the influence of these points. This is achieved through the software assigning a very low membership degree, u_{ij} , to these discontinuities, thereby reducing their impact on the calculation of mean orientations.

3. Conclusions

In this paper, the Fuzzy K-Means algorithm based on the Fisher distribution (FFKM) was presented as a tool for identifying fracture sets in an operational slope of an open-pit iron mine in the Quadrilátero Ferrífero region. Designed in Scilab, the software accurately identified the number of fracture sets and classified fractures into these sets, aligning with the results of in situ geotechnical mapping. These outputs will support future slope stability analyses and inform improved mine planning and risk management strategies.

The FFKM provides objective and reliable clustering results, reducing the subjectivity associated with conventional methods. In this version, a function was implemented to generate stereographic projections, enabling the visualization and interpretation of the spatial distribution of discontinuity sets. This feature enhances decision-making in geotechnical assessments and contributes to safer and more efficient mining operations.

The software offers flexibility for adaptation to various applications and supports integration with other tools. It is freely available for download, along with a user manual and example datasets, at <https://github.com/andrekle/FFKM-algorithm.git>.

Acknowledgements

This software has been developed through a collaboration between the Federal Institute of Education, Science, and Technology of Minas Gerais – Campus Ouro Preto, Brazil, and the Department of Civil, Chemical, Environmental, and Materials Engineering – University of Bologna, Italy.

References

- [1] Zhang K., Wu W., Liu Y., Xie T., Zhou J., Zhu H., 2024
OCFMD: An Automatic Optimal Clustering Method of Discontinuity Orientation Based on Fisher Mixed Distribution – Rock Mechanics and Rock Engineering – 57, 1735-1763, DOI: <https://doi.org/10.1007/s00603-023-03587-7>.
- [2] Klen A.M., Bonduà S., Kasmaeeyazdi S., Lana M.S., Menezes D.A., Carvalho P.G., 2024
A fuzzy K-Means algorithm based on Fisher distribution for the identification of rock discontinuity sets – International Journal of Rock Mechanics & Mining Sciences – ISSN 1365-1609, DOI: <https://doi.org/10.1016/j.ijrmms.2024.105879>
- [3] Xu L.M., Chen J.P., Wang Q., Zhou F.J., 2013
Fuzzy C-means cluster analysis based on mutative scale chaos optimization algorithm for the grouping of discontinuity sets – Rock Mechanics and Rock Engineering – 46(1), 189-198, DOI: <https://doi.org/10.1007/s00603-012-0244-z>.
- [4] Yi X., Feng W., Wu W., Zhou Y., Dong S., 2023
An Effective Approach for Determining Rock Discontinuity Sets Using a Modified Whale Optimization Algorithm – Rock Mechanics and Rock Engineering – 56, 6143-6155, DOI: <https://doi.org/10.1007/s00603-023-03364-6>.
- [5] Wu W., Feng W., Yi X., Zhao J., Zhou Y., 2024
Sparrow search algorithm-driven clustering analysis of rock mass discontinuity sets – Computational Geosciences, DOI: <https://doi.org/10.1007/s10596-024-10287-w>
- [6] Hammah R.E., Curran J.H., 1998
Fuzzy cluster algorithm for the automatic identification of joint sets – International Journal of Rock Mechanics and Mining Sciences – 35(7), 889-905, DOI: [https://doi.org/10.1016/S0148-9062\(98\)00011-4](https://doi.org/10.1016/S0148-9062(98)00011-4).
- [7] Cui X., Yan E.C., 2020
A clustering algorithm based on differential evolution for the identification of rock discontinuity sets – International Journal of Rock Mechanics and Mining Sciences – 126, 104181, DOI: <https://doi.org/10.1016/j.ijrmms.2019.104181>
- [8] Cui X., Yan E.C., 2020
Fuzzy C-Means Cluster Analysis Based on Variable Length String Genetic Algorithm for the Grouping of Rock Discontinuity Sets – KSCE Journal of Civil Engineering – 24(11), 3237-3246, DOI: <https://doi.org/10.1007/s12205-020-2188-2>

[9] **Li Y., Wang Q., Chen J., Xu L., Song S.**, 2015

K-means Algorithm Based on Particle Swarm Optimization for the Identification of Rock Discontinuity Sets – Rock Mechanics and Rock Engineering – 48(1), 375-385, DOI: <https://doi.org/10.1007/s00603-014-0569-x>

[10] **Song S., Wang Q., Chen J., Li Y., Zhang W., Ruan Y.**, 2017

Fuzzy C-means clustering analysis based on quantum particle swarm optimization algorithm for the grouping of rock discontinuity sets – KSCE Journal of Civil Engineering – 21(4), 1115-1122, DOI: <https://doi.org/10.1007/s12205-016-1223-9>.

[11] **Cardozo N., Allmendinger R.W.**, 2013

Spherical projections with OSXStereone – Computers & Geosciences – 51, 193–205, DOI: <https://doi.org/10.1016/j.cageo.2012.07.021>

[12] **Fernandes R. do C., Roncato J., Sérgio de Paula R.**, 2023

Structural model and features of the world-class Cuiabá orogenic gold deposit, Rio das Velhas greenstone belt, Quadrilátero Ferrífero region, Brazil – Journal of South American Earth Sciences – 123, 104201, DOI: <https://doi.org/10.1016/j.jsames.2023.104201>

[13] **Silva Junior A.N. da**, 2014

Definition of Mineral Resources: An Illustrative Example of an Iron Ore Deposit in the Quadrilátero Ferrífero, MG – Universidade Federal de Ouro Preto – 2022, <http://www.monografias.ufop.br/handle/35400000/4492>.

[14] **Rossi D.Q., Endo I.**, 2015

A structural model of the Fábrica Nova region, Santa Rita syncline, Quadrilátero Ferrífero: Flanking folds as a folding mechanism – REM: R. Esc. Minas, Ouro Preto – 68(2), 153-162, apr. jun., ISSN 0370-4467- <https://doi.org/10.1590/0370-44672015680058>.

[15] **Menezes D.A. de**, 2014

Geotechnical characterization and analysis of failures mode of operational slopes in compact itabirito – Universidade Federal de Ouro Preto – <http://www.repositorio.ufop.br/jspui/handle/123456789/3690>

[16] **Xie X.L., Beni G.**, 1991

A validity measure for fuzzy clustering – IEEE Transactions on Pattern Analysis and Machine Intelligence – 13(8), 841-847, DOI: <https://doi.org/10.1109/34.85677>



This article is an open access article distributed under the Creative Commons BY SA 4.0 license. Authors retain all copyrights and agree to the terms of the above-mentioned CC BY SA 4.0 license.