Statistical analysis for the hydrogeological evaluation of the fracture networks in hard rocks

JOHANNES HOFRICHTER [1] & GERFRIED WINKLER [2]

[1] Institute of Applied Statistics and System Analysis, Joanneum Research GmbH, Steyrergasse 25a, 8010 Graz, Austria; E-mail: Johannes.hofrichter@joanneum.at
[2] Institute of Hydrogeology and Geothermics, Joanneum Research GmbH, Elisabethstr. 16/II, 8010 Graz, Austria; E-mail: gerfried.winkler@joanneum.at

Abstract: The hydrogeological effectiveness of fracture sets is determined and evaluated by the fuzzy c-mean and hierarchical clustering. These cluster analyses combine the geological spatial attributes and the hydraulic relevant attributes of fractures. Based on the results of the clustering the fracture set volumes are estimated.

Key words: clustering, fuzzy c-mean, hierarchical, fracture network, fracture set volume

INTRODUCTION

Large areas of Austria like the Central Alps or the Bohemian Massif are built up by crystalline hard rocks. Tectonic deformations of geological units cause the development of fracture networks forming fissured/ fractured aquifers in hard rocks. The water flow in fractured hard rocks is predominantly within its fracture network, which mostly consists of several fracture sets. For the determination of favoured flow directions in fractured hard rocks it is necessary to figure out the fracture network and the hydrogeological effectiveness of the fracture sets. The hydrogeological effectiveness of fracture sets yield from two groups of fracture attributes. The first group contains the geological spatial attributes like trend, plunge and the frequency of fractures. The other group includes the hydraulic relevant attributes like aperture, trace length and linear degree of separation of the fractures. The hydrogeological effectiveness of individual fracture sets is tried to be answered with a new approach. Statistical clustering integrating both attribute types enhances the determination of the hydrogeological effectiveness of the fracture sets. Further on the volume of each fracture set can be estimated.

METHODOLOGY

Data recording

At exposures the attributes of the fractures are recorded with the scanline sampling technique based on PRIEST (1993). In advance of the data recording some assumption have to be defined. The aperture should be measured on several points along one fracture and then averaged. Fractures of the exposure face of which the aperture is immeasurable are marked separately with “n.m.” (not measurable). The total length is the complete measurable stretch of a fracture. The linear degree of separation is the sum of the trace length sections of one fracture where an aperture can be observed.

Scientific paper

Statistical data analysis – orientation related fracture volumes

Pre-processing

In advance to the statistical analyses some biases have to be corrected. Biases caused by a) the sampling technique itself (KULITULAKE ET AL., 1993) b) the immeasurable attributes (aperture, length) of fractures forming the exposure face and c) the varying orientation of scan lines to the orientation of the intersecting fractures.

The maximum trace length is standardized with 2 metres and the maximum averaged aperture with 1 centimetre deduced by existing data. The fractures are weighted by their angle between the fracture and the scan line with

$$g_i = \frac{1}{l_s} \min\left(1/\left|\langle x_s, x_i \rangle\right|,\ 5\right), \qquad (0)$$

where l_s is the length of the scanline on which the fracture is detected, x_s and x_i are the vectors of the scanline and the fracture pole on the unit sphere surface respectively. The immeasurable fractures attributes on the exposed rock face are estimated using the attributes of similar orientated fractures (HARUM ET AL., 2001).

Another goal of the fracture analysis is the determination of the fracture set volume. Therefore the individual fractures are weighted by the attributes aperture o_i and linear degree of separation d_i. We assume that the attributes recorded on the surface continue through the whole rock mass. So the area defined by the aperture and linear degree of separation can be regarded as a representative weighting value. Considering these assumptions the weights gv_i are given by

$$gv_i = o_i * d_i \qquad (1)$$

Combining the two weights g_i (0) and gv_i (1) the total weight v_i for one fracture is

$$v_i = g_i * gv_i \qquad (2)$$

This total weight describes the hydrogeological importance of a fracture based on the fracture volume.

Cluster analysis

The cluster analysis classifies objects to groups (clusters). The classification is based on the similarity and dissimilarity of the objects attributes. So we have to define the similarity respectively dissimilarity (= distance) between the objects which can be defined in many ways. The cluster analysis describes the similarity of two objects with the distance between these two objects. That means the stronger the similarity the smaller the distance. The definitions of different statistical distances is well discussed by many authors including HAMMAH ET AL. (1998 and 1999) and STEINHAUSEN ET AL. (1977). For the spherical data we consider the sine-square distance and the Mahanalobis distance of spherical data. The common distance of numeric data is the Euclidean distance. To combine the different kinds of distances, STEINHAUSEN ET AL. 1977 propose a **mixed distance** d_M. Let d_r be the distance of the direction (sine-square or Mahalanobis of spherical data) and d_e the Euclidean distance, then d_M is defined as

$$d_M = \frac{m_r d_r}{m} + \frac{m_e d_e}{m}, \qquad (3)$$

where m_r is the number of the variables representing the direction (in the discrete demand m_r =3), me is the number of variables included in the Euclidian distance (in the discrete demand m_e =2) and out of that $m = m_r + m_e$.

The **agglomerative hierarchical clustering** leads to an exact partition of objects, that means that one object is classified to exact one group. At the beginning every cluster is defined by exact one object. The two clusters with the lowest distance are combined to a new cluster. So the total number of clusters is reduced by one. This step is repeated till the demanded number of clusters is reached. For the distance between two clusters we use the "mixed distance" as described above. After the first step one cluster can include one or more than one object. So we derive the distance between two cluster with the centroid-method calculating the "mixed distance" between the mean of the clusters.

The **fuzzy c-mean clustering** differs from the hierarchical clustering in two points. The number c of cluster have to be determined before starting the cluster analysis and each object is classified to each cluster with a certain degree of membership. That means that the objects are not classified to exact one cluster but can be transformed to an exact one. So the results of both methods can be compared.
The result of the clustering is the definition of groups containing fractures with similar spatial and hydrological attributes, which can be regarded as homogenous groups (clusters).

Cluster attributes
The **center of gravity** describes the mean orientation of a fracture set (Wallbrecher, 1986). The uncertainty of the center of gravity can be described by cones of confidence (95 % and 99 %). In the structural geology the cone of confidence is parametrically estimated. Therefore the data are assumed having a certain spatial distribution. One approach, which is not bound, on a certain spatial distribution is based on the nonparametric bootstrap method (Davison & Hinkley, 1997). The advantage of that method is that the shape of the cone of confidence is a result of the empirical distribution of the data. So the shape of the cone of confidence is not bound on the shape of the assumed spatial distribution.

The orientation related **fracture volumes** define the hydrogeological effectiveness of fracture sets. Multiplying the aperture and the standardized linear degree of separation (1) leads to the standardized area (equal to the weight gv_i). Correcting the standardized area like the weights in (2) leads to a weighted area that is proportional to the fracture volume. So the sum of the weighted areas can be regarded as a first estimation of the fracture set volume.

Results and discussion

The practical application of this research work has been carried out in the area Sonnwendstein/ Semmering – Austria (S6 Semmering tunnel) (Harum et al., 2001). The new clustering methods enhance a) to calculate the cluster volumes and b) to define more precisely the orientation of hydraulic relevant clusters. The analyses of all the exposures figure out that the fuzzy c-mean clustering using the mixed distance lead to the most satisfying results considering the geological, tectonic circumstances.

CONCLUSIONS

The Clustering method enhances the description of the fracture sets in more detail combining the geological spatial and the hydraulic relevant attributes. It enhances to determine the hydrogeological relevant fracture sets and to separate clusters with equal geological but with different hydraulic attributes.
Because of the linear sampling technique (scanline) boreholes can be included as an important complementary information for the interpretation into depth. Additionally these results can be combined with results of hydraulic tests along boreholes to get a relation between fracture set volumes and their hydraulic capacity. It can help determining and quantifying the hydraulic attributes of a fracture network and their spatial distribution for a numerical realisation of fractured aquifers. So this method enhances a better three dimensional imagination of the fracture network and a better spatial characterisation of hydrogeological units.

Acknowledgements

This method was developed within a research project supported by the Federal Ministry of Transport, Innovation and Technology of Austria and Joanneum Research GmbH.

REFERENCES

DAVISON, A. C. & HINKLEY, D. V. (1997): *Bootstrap methods and their application. Cambridge Seires in Statistical and Probabilistic Mathematics*, Cambridge.

HAMMAH, R. E. & CURRAN, J. H. (1998): Fuzzy Cluster Algorithm for the Automatic Identification of Joint Sets. *Int. J. Rock Mech. Min. Sci.* Vol.35, No. 7, pp. 889-905, Oxford.

HAMMAH, R. E. & CURRAN, J. H. (1999): On Distance Measures for the Fuzzy K-means Algorithm for Joint Data. *Rock Mech. Rock Eng.* 32 (1), 1-27, Oxford.

HARUM, T., HOFRICHTER, J., KLEB, U., KUPFERSBERGER, H., POLTNIG, W., REICHL, P., REINSDORFF, S., RUCH, C., SCHÖN, J., STROBL, E., WINKLER, G. (2001): *Erfassung von Fließprozessen zur hydrogeologischen Bewertung von klüftigen Festgesteinen*: unpublished report, Joanneum Research GmbH, S. 124, Graz.

KULITULAKE, P. H. W., WATHGALA, D. N., STEPHANSSON, O. (1993): Joint network modelling with a valiation exercise in Stripa Mine Sweden. Int. J. of Rock Mech. a. M. Sc. 30: 503-526, Oxford.

PRIEST, S. D. (1993): *Discontinuity analysis for rock engineering*, 1st ed. Chapman & Hall, 473 pp, London.

STEINHAUSEN, D. & LANGER K. (1977): Clusteranalyse Einführung in Methoden und Verfahren der automatischen Klassifkation. Walter de Gruyter, Berlin New York.

WALLBRECHER, E. (1986): *Tektonische und gefügeanalytische Arbeitsweisen*. S.244, Stuttgart.