ARMA/USRMS 05-877

Multivariate Clustering Analysis of the ECRB Cross Drift Discontinuities, Yucca Mountain Project

Maerz, N. H.
University of Missouri-Rolla, Rolla, MO, USA
Zhou, W.
University of Alaska Fairbanks, Fairbanks, AK, USA

This paper was prepared for presentation at Alaska Rocks 2005, The 40th U.S. Symposium on Rock Mechanics (USRMS): Rock Mechanics for Energy, Mineral and Infrastructure Development in the Northern Regions, held in Anchorage, Alaska, June 25-29, 2005.

ABSTRACT: A great deal of effort has been made to characterize the nature of the discontinuities of the Yucca Mountain proposed nuclear waste repository. Discontinuity sets were identified, using orientation only as the basis of cluster analysis. Other discontinuity characteristics such as lithology fracture frequency, continuity, roughness, aperture, and infilling have been analyzed, but separate and divorced from the clustering analysis.

The across drift data has been re-analyzed using a multivariate clustering analysis algorithm developed by the authors (CYL). This type of analysis represents a relatively recent development in characterizing the structure of rock masses. It characterizes discontinuities into subsets according to multiple parameters, such as orientation, spacing, and roughness, where rather than considering one variable at a time, a number of parameters can be treated simultaneously, so that the interactions between parameters are taken into account. The comprehensive algorithm has been developed into a software package. It enables fully automated multivariate clustering analysis and offers various visualization tools, such as a three dimensional stereonet, a stereoscopic view, a statistical table, and pie charts relating the other factors such as lithology continuity, roughness, aperture, and infilling back to each cluster.

1. INTRODUCTION

1.1. Yucca Mountain Project

The Yucca Mountain site in Nevada has been designated as United States choice for nuclear waste repository. Yucca Mountain is in a remote dry area, on federal government land. Investigative work began in 1978, construction began in 1993, and the main tunnel was completed by TBM (Tunnel Boring Machine) in 1998. The 2.7 km cross drift tunnel was completed in 1998 as part of the Enhanced Characterization of the Repository (ECRB) study. The cross drift is a linear TBM cut that crosses through the area that is proposed for nuclear waste storage.

A great deal of effort has been made to characterize the nature of the discontinuities of the Yucca Mountain proposed nuclear waste repository. Discontinuities largely determine the mechanical, hydrological, and thermal behavior the rock mass. Mongano et al. [1] detailed the structure of the 2.7 km cross drift (Fig. 1). Measurements include fracture orientation, frequency, trace length, height and width, roughness, termination type, aperture, roughness, and infilling type and thickness.

Cluster analysis, to identify the discontinuity sets was done using the old Clustran code. Cluster analysis was done on the basis of orientation only. Other attributes of the fractures were analyzed but not in the context of their clustering.

1.2. Multivariate Clustering Analysis

Multivariate clustering analysis represents a relatively recent development, characterizing discontinuities into subsets according to multiple parameters, such as orientation, spacing, and roughness, where rather than considering one variable at a time, a number of parameters can be treated simultaneously, so that the interactions between parameters are taken into account [2].

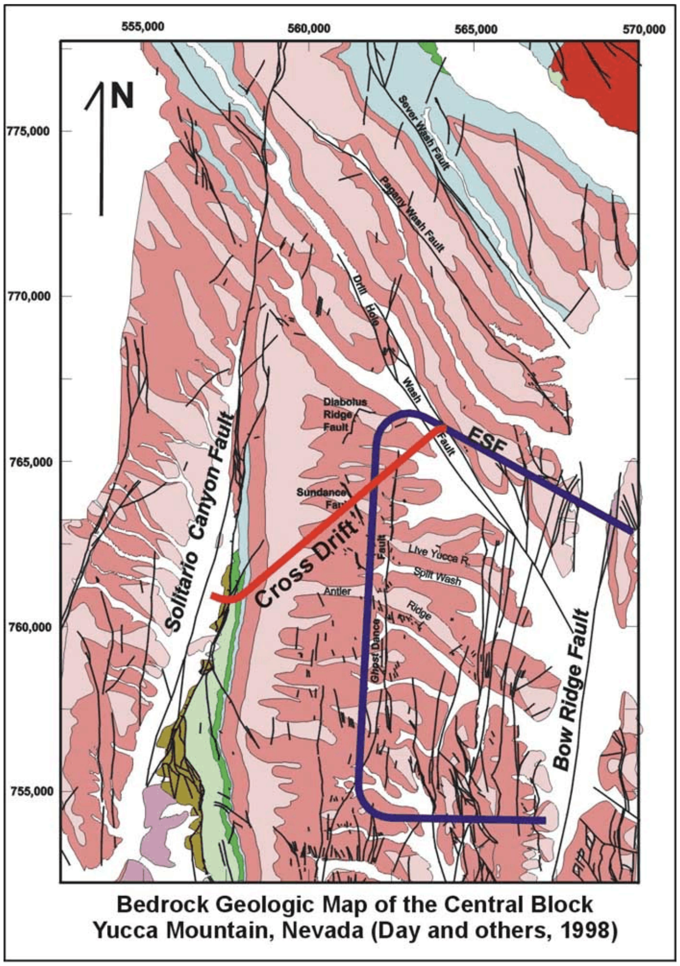

Fig. 1. Plan view of the Yucca Mountain site, showing the cross drift (Mongano et al. [1]).

Multivariate clustering has been proposed in the literature [3-9] in the last few years. In these algorithms clustering is done on basis of not just orientation, but also physical location roughness, and other quantifiable parameters. In the algorithms previously presented by the authors of this paper [3-7], a "3 dimensional" stereonet (concept shown in Figs. 2 and 3) where discontinuity normals are plotted on individual "stacked" stereonets, each normal is plotted with respect to its own stereonet, and each stereonet is plotted in a linear position that corresponds to the position where the discontinuity corresponding to that discontinuity normal intersects the bore hole or mapping scanline. This stereonet is the ideal device for visualization of clustering that is based on orientation and position, as shown in Fig. 4. The other parameter that can be used for clustering is roughness which cannot be visualized so easily. The clustering methods are best described in [4] and examples can be found in [5-7].

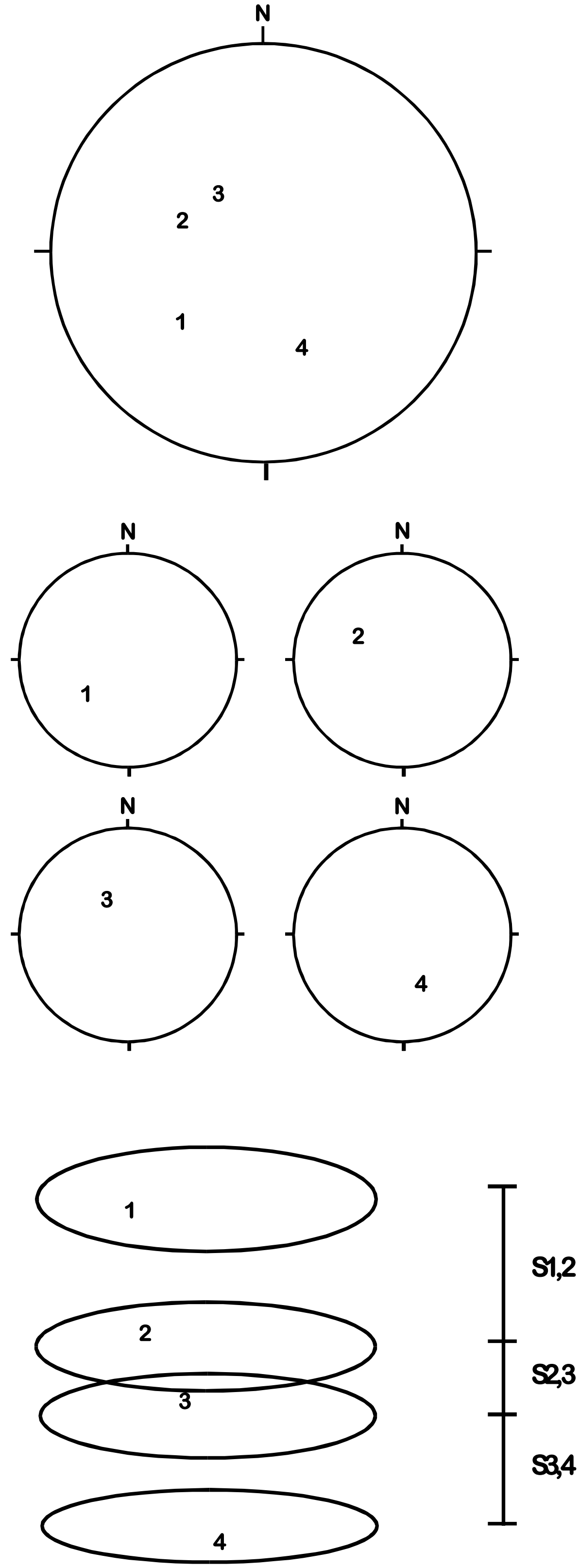

Fig. 2. Top: A lower hemisphere stereonet with four discontinuity normals (poles), each pole ostensibly from a different depth along an imaginary vertical bore hole. Middle: Each discontinuity normal (pole) is plotted on an individual stereonet. Bottom: The individual stereonets are stacked, with each spacing in proportion to the spacing between discontinuities in the borehole [7].

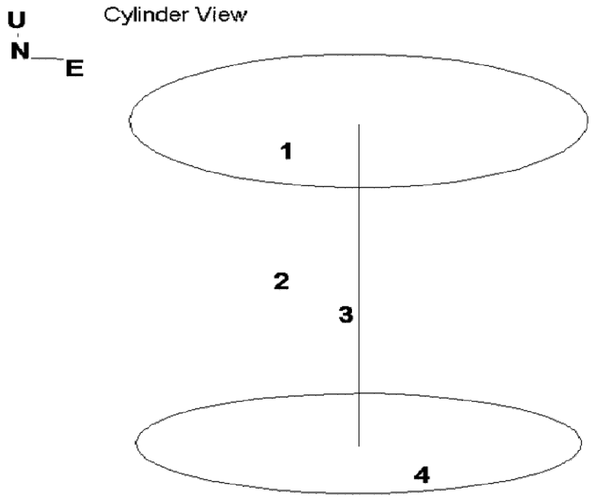

Fig. 3. “Three-dimensional stereonet” [7].

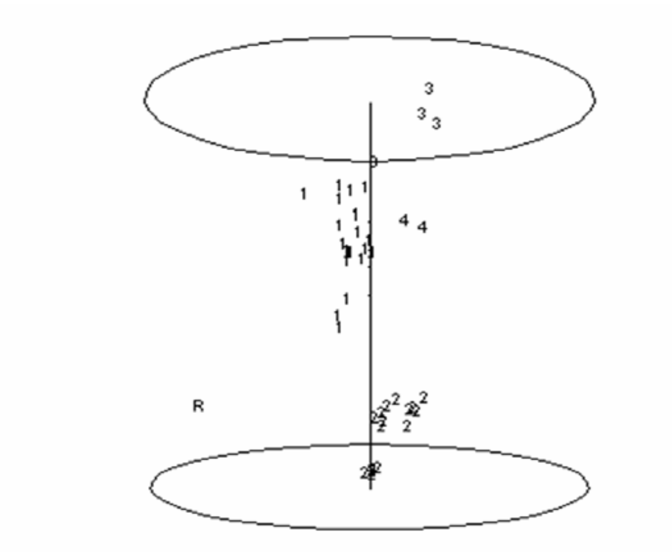

Fig. 4 Three-dimensional stereonet showing 4 sets and 1 random discontinuity clustered on the basis of orientation and position along the spatial axis [7].

1.3. *Advances in the Algorithm*

A more recent advance in the computer algorithm allows for further integrated analysis of the parameters that are not used in the clustering analysis. As an example, if the analysis identifies three clusters of discontinuities, a parameter such as trace length or infilling can be examined as a function of cluster number. It may be for example that discontinuities of cluster number 1 are predominantly filled with a particular type of material, while the other 2 clusters are not. Fig. 5 shows an example where 3 discontinuity sets are analyzed as a function of rock type.

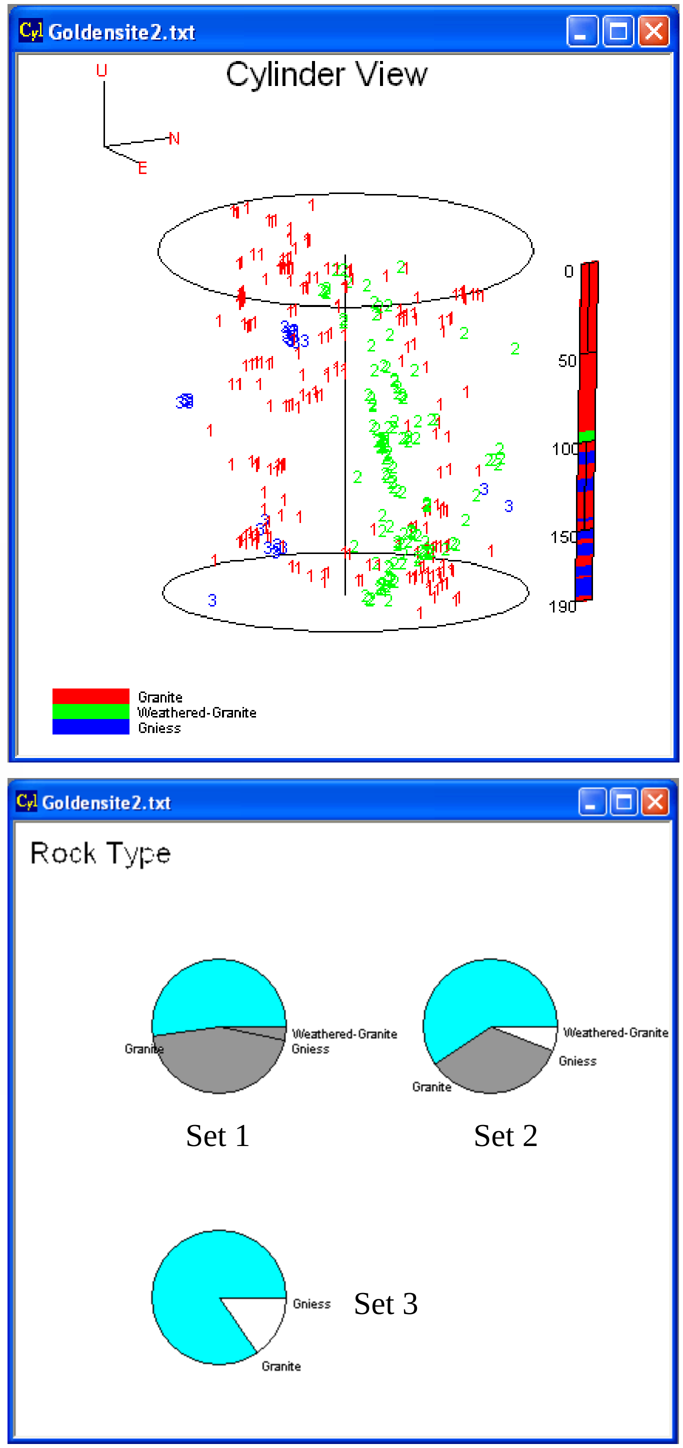

Fig. 5. Top: Analysis of a scanline into a rock mass with 3 discontinuity sets and 3 rock types. Bottom: Each pie chart represents one of the discontinuity sets, in numerical order. The weathered granite has only discontinuities of set 2, and the members of discontinuity set 3 are prominently in the granite. Outputs produced by the CYL Program.

1.4. *Re-analysis of Yucca Cross Drift Data*

This paper describes a re-analysis of the Yucca Mountain cross drift data.

2. ANALYSIS

2.1. Mongano et al. Analysis

Mongano et al. [1] produced a comprehensive report on the Geology of the ECRB crossdrift, including regional geology, litholostratigraphy, structure, and geotechnical characterization. As part of the structural characterization there is reported a set of analysis of the fractures. Analysis consisted of cluster analysis as a function of orientation, fracture frequency, continuity, aperture, and mineral or clastic infillings.

Cluster Analysis

Cluster analysis was done using the old Clustran code first developed by Shanley and Mahtab [10]. One analysis of the entire data, seven fracture sets were identified, using orientation only as the basis of cluster analysis. The cross drift traverses four main lithological units, because the drift is horizontal while the strata dips at a shallow angle. The units are called the Tptpul, Tptpmn, Tptpll, and Tptpln. The description of these units is beyond the scope of this paper but can be found in Mongano et al. [1].

For their detailed analysis, the authors determined that only four of the clusters were significant, and reanalyzed for the four major lithologies encountered along the drift. Fig. 6 shows typical results for the Tptpul section. In other lithological sections the pattern is similar, although orientation variability changes and the shallow dipping cluster is shallower in the other sections.

Other parameters

The other parameters are presented independently of the fracture sets and the clustering process, although they are presented both as a function of position (consequently lithology) and as a frequency distribution.

Fracture frequency was measured and mapped as a function of position and lithology (Fig. 7). Fracture roughness is presented in the same way and as a frequency distribution (Fig. 8). Mineral and clastic infillings are presented in terms of thickness vs. position, as a function of type of filling (Fig. 9). Roughness and apertures are presented as a frequency distribution (Figure 8).

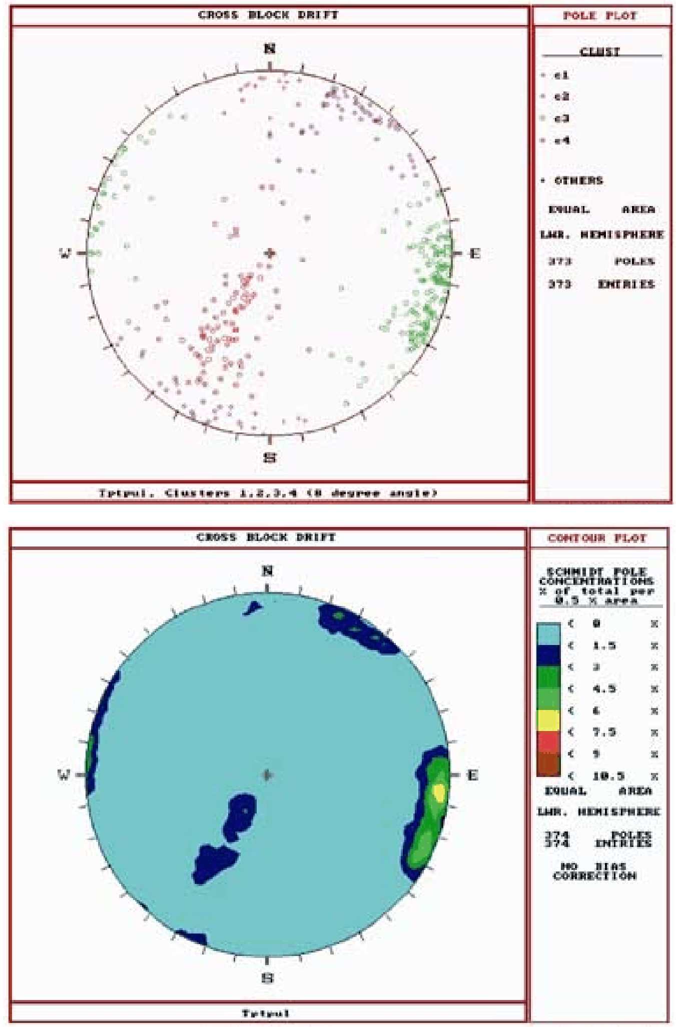

Fig. 6. Mongano et al. Analysis of the Tptpul data [1]. Top: Poles. Bottom, contouring of the data with 4 prominent fracture sets. 1. Steeply dipping to the N and S (difficult to see because of a small number of fractures); 2. Steeply dipping to the E-SE and W-NW; 3. Steeply dipping S-SW and N-NE; 4. Shallow dipping S-SW.

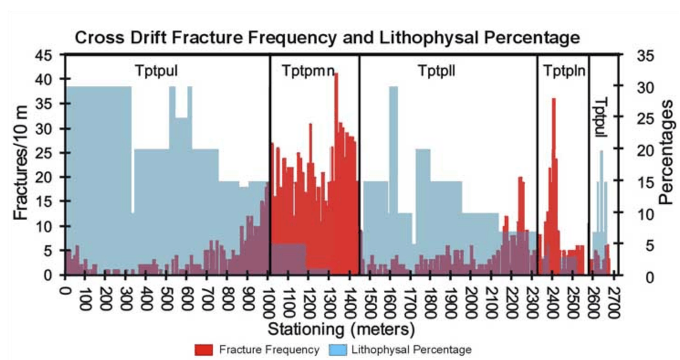

Fig. 7. Fracture frequency as a function of position and lithology. [1].

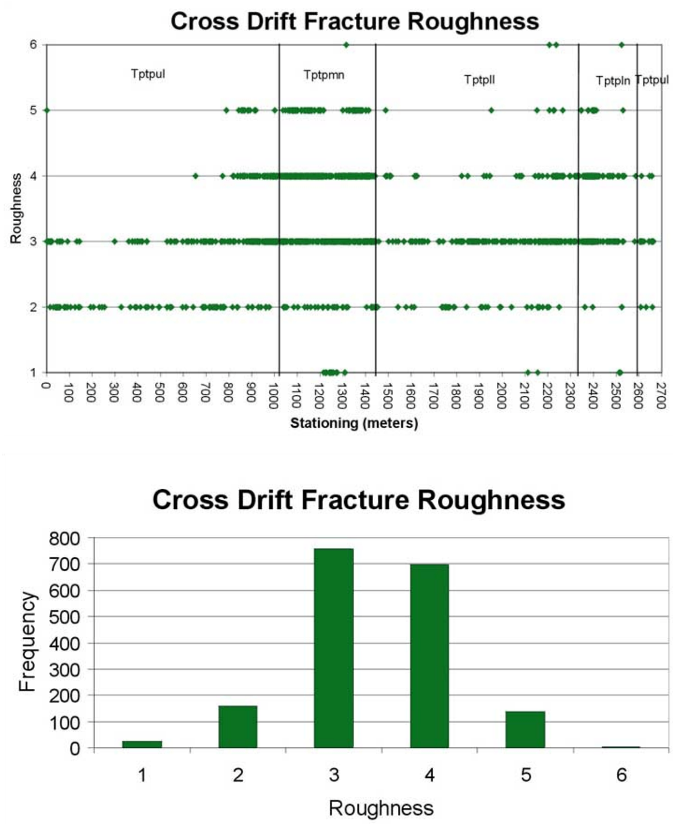

Fig. 8. Fracture roughness [1]. Top: As a function of position and lithology. Bottom: As a frequency distribution.

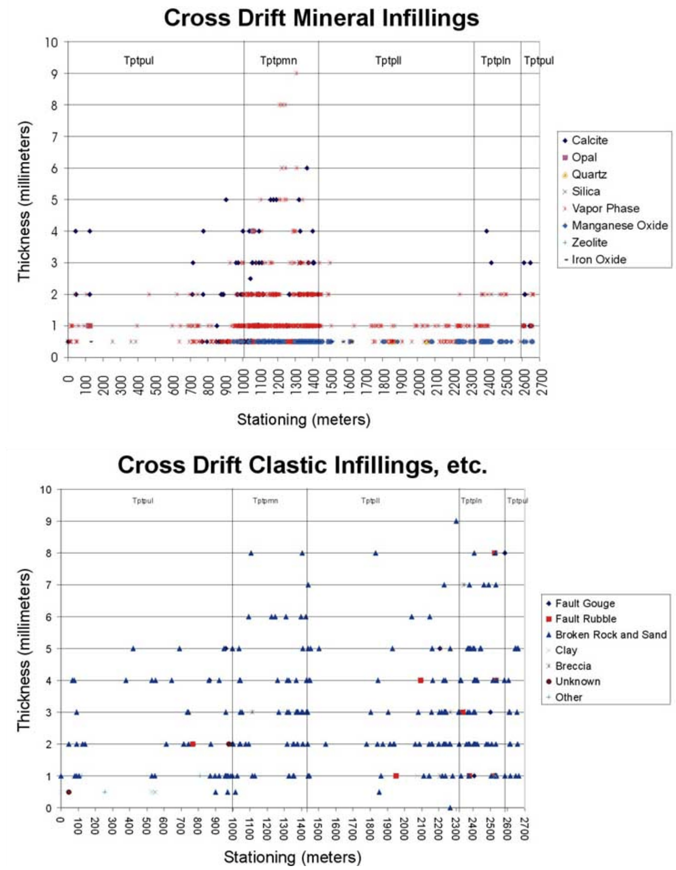

Fig. 9. Infilling as a function of position and type [1]. Top: Mineral Infilling. Bottom: Clastic infilling.

2.2. New analysis

The computer code CYL was used to reanalyze the Yucca data. Because the data set was so large, it was divided into discrete sections based on the lithological units. Orientation, position, and roughness were used in the clustering algorithm.

Cluster analysis

The cluster analysis was done using the supervised nearest neighbor method or the vector quantization method, and the number of clusters were set to match the results from the Mongano et al. [1] analysis. Position along the sampling line was included in the multivariate analysis, and given a relative weighting of 0.3. The results (Figs. 10 and 11) are very similar to the Mongano results (Fig. 12). Fig. 13 summarizes the orientation, position, and average spacing and roughness for each set. It is immediately obvious that that the spacing is different for different sets. Set 1 has a spacing of 7.0m while set 4 has a spacing of 23.2m.

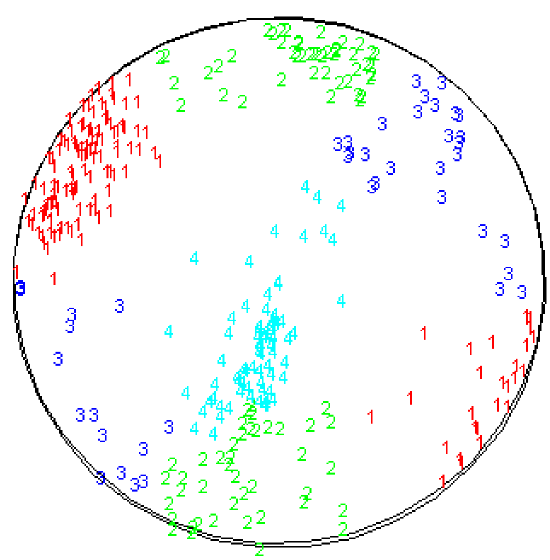

Fig. 10: Analysis of the Tptpul section, nearest neighbor clustering including position as a clustering variable.

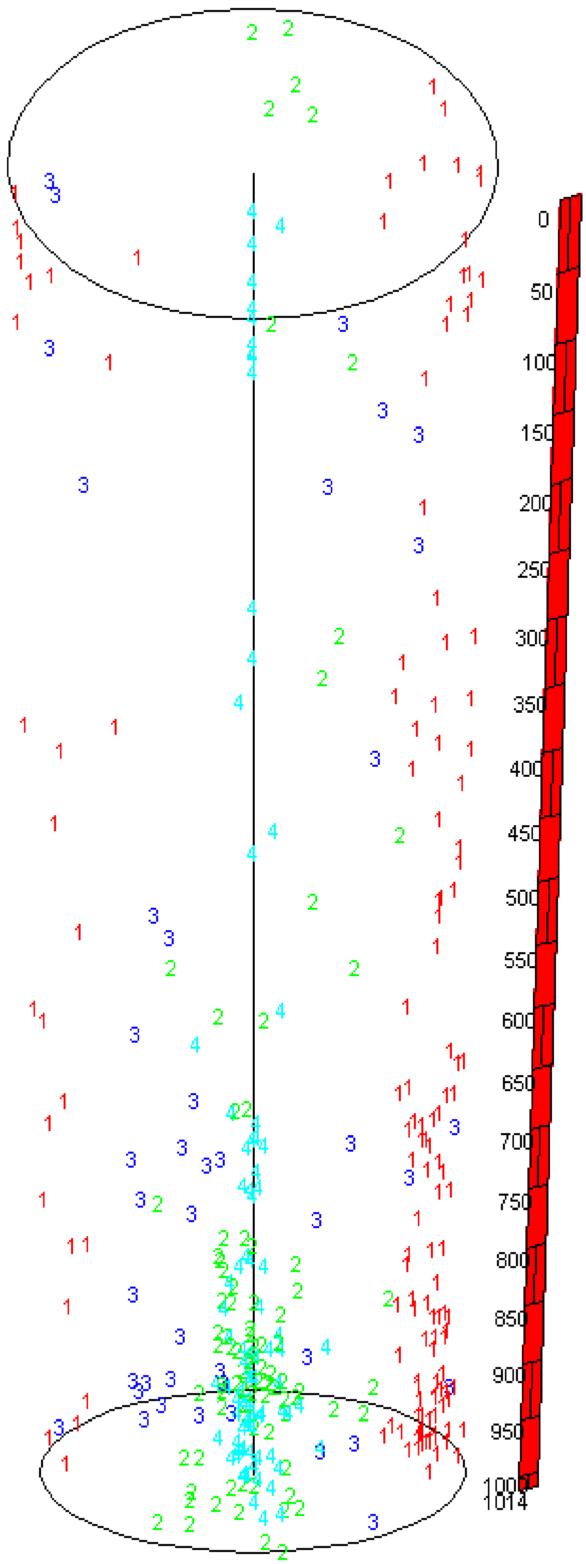

Fig. 11. Three dimensional version of Fig. 10. The increase in fracture density at about 700m (see Fig. 7) can be seen in this view

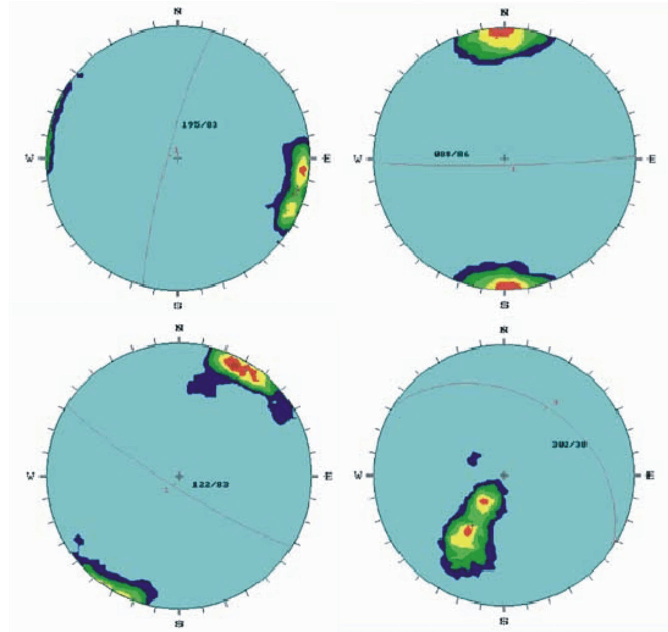

Fig. 12. Summary of the Mongano et al. clustering [1] of the Tptpul section.

Clustering Analysis

set #	joints	dip dir	dip ang	position	spacing	roughness
1	144	258	82	649.5	7.0	2.8
2	90	306	77	826.4	11.2	3.2
3	43	356	77	698.6	23.2	3.0
4	86	152	29	766.6	11.5	2.7

Fig. 13. Summary of the clustering results shown in Fig. 10. Because of the nature of the input data, the column labeled dip direction is actually strike direction (dip to right).

Clustering Analysis

set #	joints	dip dir	dip ang	position	spacing	roughness
1	143	257	82	660.0	7.0	2.8
2	158	85	68	859.2	6.4	3.3
3	16	152	82	304.7	60.4	2.7
4	46	158	15	627.6	21.2	1.9

Fig. 14. Summary of the clustering results when roughness is considered as part of the multivariate clustering. The average dip direction and dip angle parameters are not necessarily meaningful in this type of analysis.

The analysis was then modified to also include roughness as a clustering parameter, with a relative weighting of 0.25. At this point the view of the stereonet is not meaningful, as the roughness dimension does not visualize well. However looking at the clustering results (Fig. 14) shows that that there is a variability in the roughnesses of the different clusters. Set 2 in this case has an average roughness of 3.3 (on a scale of 1-6) whereas set 4 has an average roughness of only 1.6.

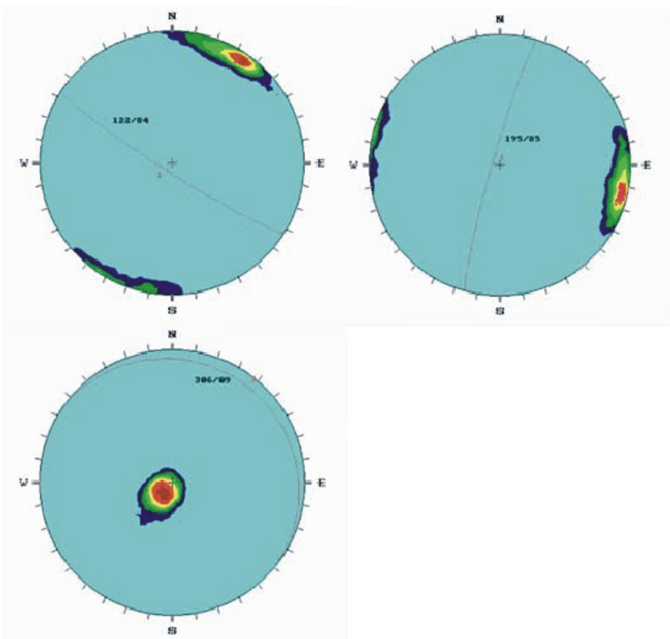

Fig. 15. Summary of the Mongano et al. clustering [1] of the Tptpmn section.

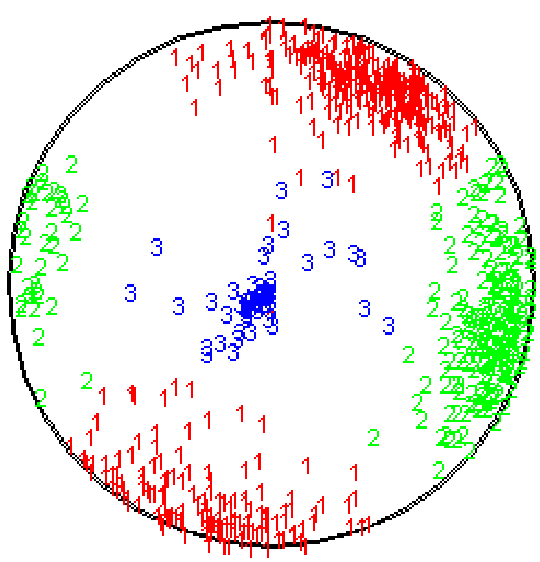

Fig. 16: Analysis of the Tptpmn section, vector quantization clustering not including position or roughness as a clustering variable.

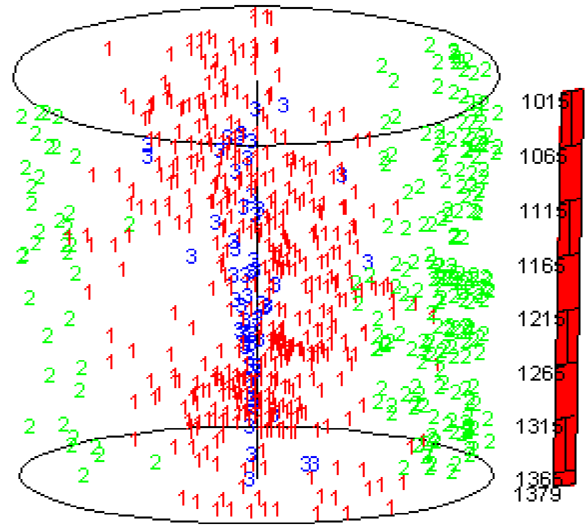

Fig. 17. Three dimensional version of Fig. 16. The high fracture density (see Fig. 7) can be seen in this view.

In Figs. 15-17 a similar analysis of the Tptpmn section can be seen, this time using the vector quantization method. In this case and in others, where discontinuity type, planarity, alteration, and aperture were referenced against set number, no obvious trends emerged. An example is shown in Fig. 18.

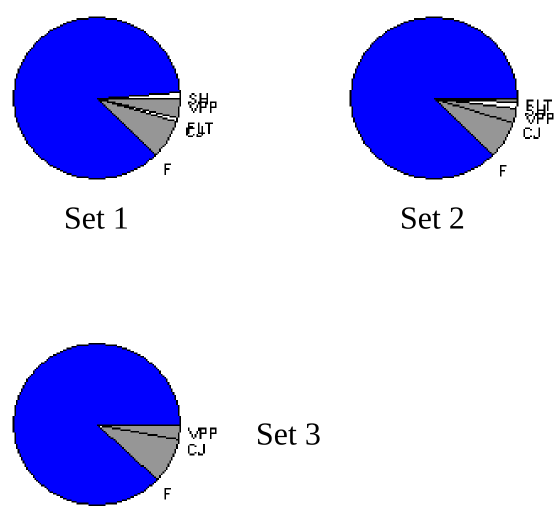

Fig. 18. Top: Each pie chart represents the discontinuity types of one of the discontinuity sets (in numerical order) from the analysis of the Tptpmn section. There is no systematic difference. The vast majority of discontinuity types are F (Fractures) with CJ (Cooling Joints) type a distance second, irrespective of cluster number. VPP (Vapor Phase Partings), SH (Shears), and FLT (Faults) comprise the rest of the data set.

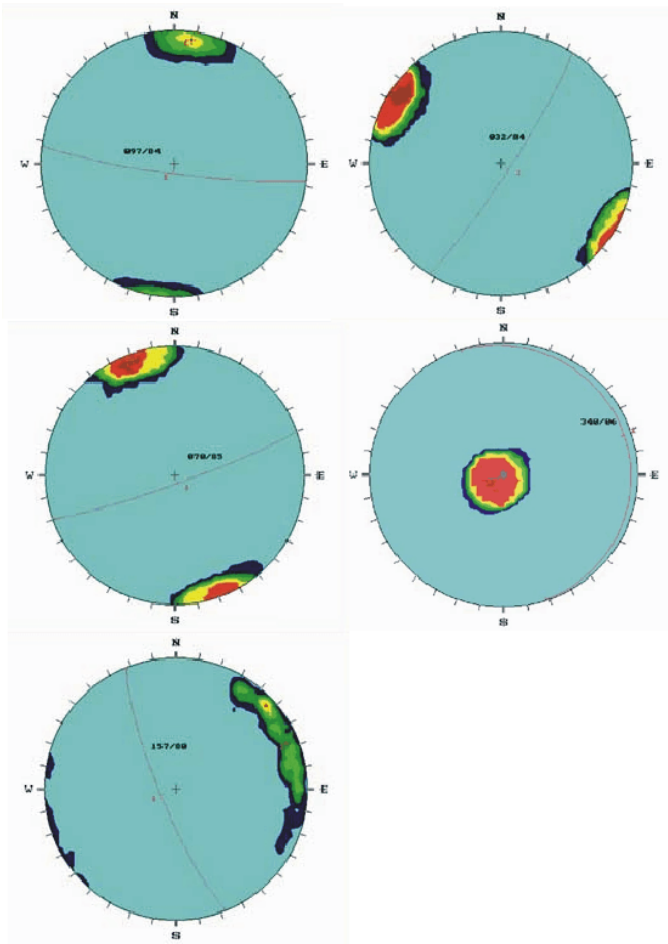

Fig. 19. Summary of the Mongano et al. clustering [1] of the Tptpll section.

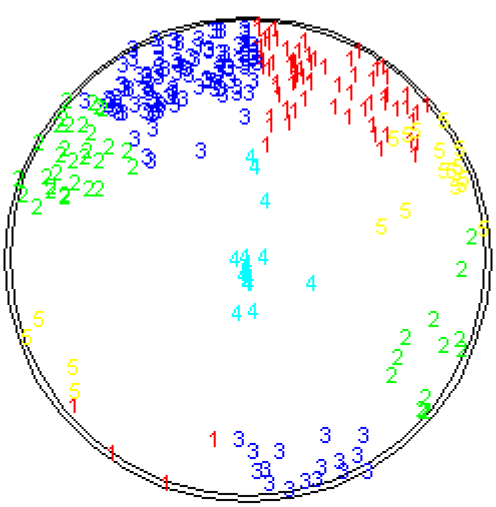

Fig. 20: Analysis of the Tptpll section, vector quantization clustering not including position or roughness as a clustering variable.

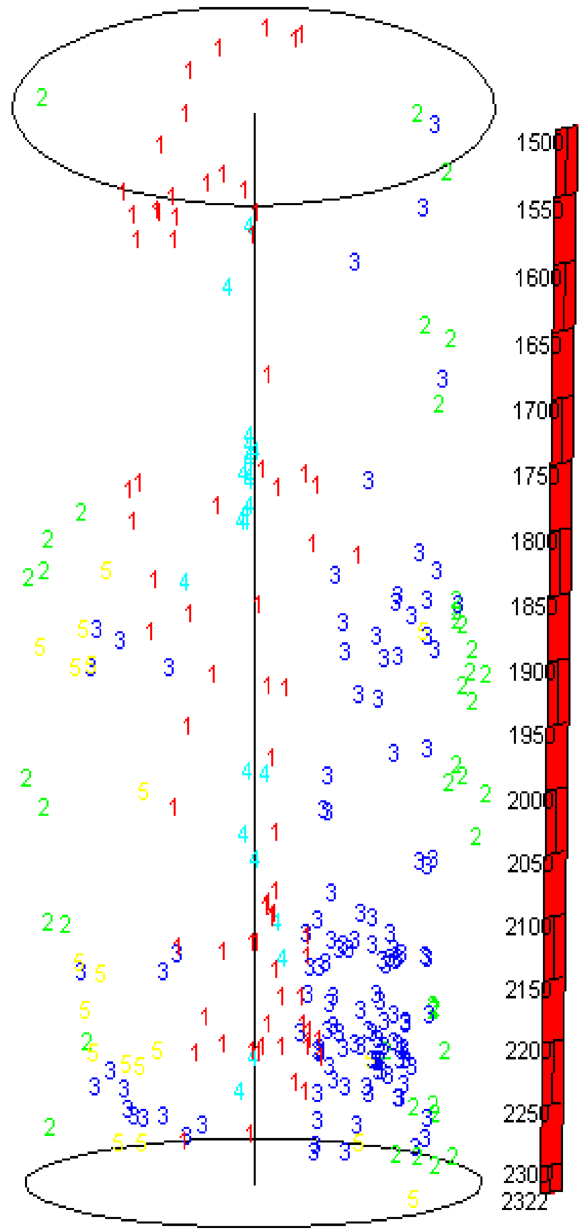

Fig. 21. Three dimensional version of Fig. 20.

In Figs. 19-21 a similar analysis of the Tptpll section can be seen, also using the vector quantization method.

In Figs. 22-24 a similar analysis of the Tptpln section can be seen, also using the vector quantization method.

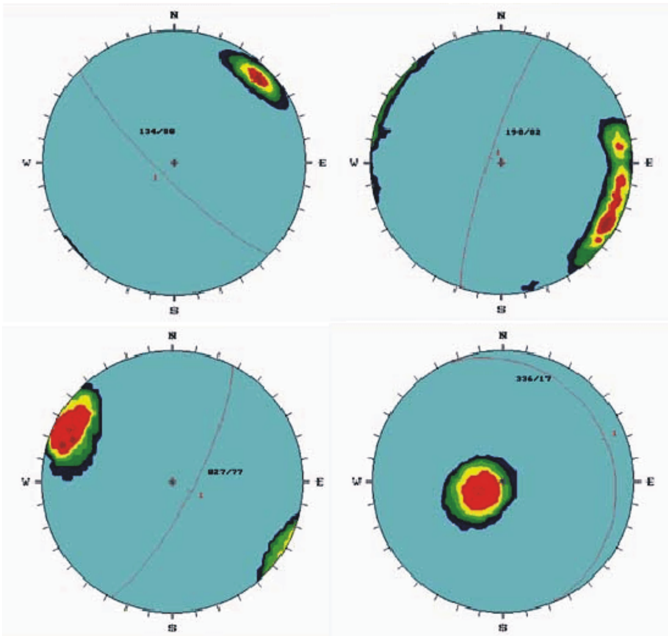

Fig. 22. Summary of the Mongano et al. clustering [1] of the Tptpln section.

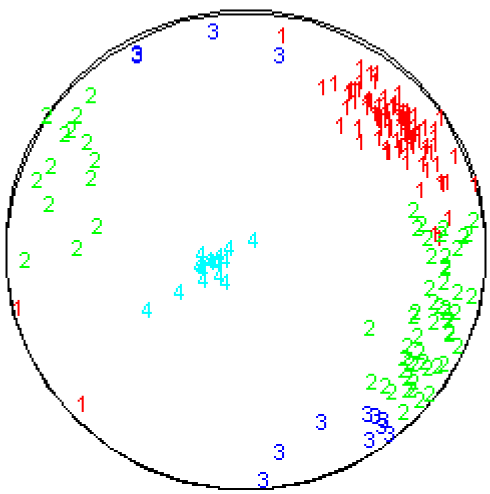

Fig. 23: Analysis of the Tptpln section, vector quantization clustering not including position or roughness as a clustering variable.

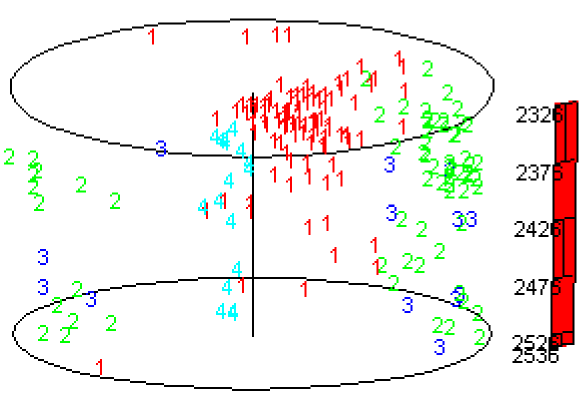

Fig. 24. Three dimensional version of Fig. 23.

3. SUMMARY AND DISCUSSIONS

The re-analysis of the discontinuity data from the Yucca Mountain Cross Drift has shown that similar analysis results to those achieved in the official report can be achieved using a single tool, the CYL multi-variate clustering and three dimensional stereonet visualization tool. The CYL analysis was fully automatic and accomplished in a time efficient manner.

At the same time the CYL algorithm could add position and roughness to the clustering algorithm. In these particular analyses, there was no advantage or insight obtained by clustering on position. This is not atypical when analyzing within a single rock type. However, when the spacing was incorporated it was clear that there were differences in space as a function of the set number that was not just related to the attitude of the sampling line with respect to the discontinuity orientation.

In the instance where roughness was incorporated into the analysis the clustering could be viewed in a completely different manner.

When secondary parameters such as discontinuity type, planarity, alteration, and aperture were referenced against set number, no obvious trends emerged. This is somewhat puzzling because one would normally expect that different genetic causes would produce different sets of discontinuities. This is certainly true in the case of bedded sedimentary and similar rocks where one discontinuity set is defined by a primary fabric like bedding, and subsequent discontinuity sets tend to orient themselves in a mutually orthogonal attitude.

What is clear however, is that with this new tool in hand more detailed analyses can be undertaken.

4. ACKNOWLEGEMENTS

The authors would like expressing our gratitude to the University of Missouri System Research Board and National Science Foundation for funding the research that developed the multivariate clustering algorithms. Special thanks to R.C. Lung of the Yucca Mountain project for making the data available.

5. REFERENCES

1. Mongano, G.S., W.L. Singleton, T.C. Moyer, S.C. Beason, G.L.W. Eatman, A.L. Albin, and R.C. Lung, 1999. Geology of the ECRB cross drift - exploratory studies facility, Yucca Mountain Project, Yucca Mountain, Nevada. *http://www.ocrwm.doe.gov/documents/spg42gm3_a/index.htm*.

2. Dershowtz, W.S., and H.H. Einstein, 1988. Characterizing rock discontinuity geometry with discontinuity system models, *Rock Mechanics and Rock Engineering*, 21, 21-51

3. Zhou, W., and N.H. Maerz, 2002. Identifying the optimum drilling direction for characterization of discontinuous rock. *Environment & Engineering Geoscience*, v. VII, no. 4, pp. 295-307

4. Zhou, W., and N.H. Maerz, 2002. Implementation of multivariate clustering methods for characterizing discontinuites from oriented boreholes. *Computers & Geosciences*, v. 28, no. 7, pp. 827-839.

5. Zhou, W., and N.H. Maerz, 2001. Multivariate clustering analysis of discontinuity data: implementation and applications. *Rock Mechanics in the National Interest; Proceedings of the 38th U.S. Rock Mechanics Symposium*, Washington, D.C., July 7-10, 2001, pp 861-868.

6. Maerz, N. H., and W. Zhou, 2000. Discontinuity data analysis from oriented boreholes. *Pacific Rocks; Proceedings of the Fourth North American Rock Mechanics Symposium*, Seattle, Washington, July 31-Aug.1, 2000, pp. 667-674.

7. Maerz, N. H., and W. Zhou, 1999. Multivariate analysis of bore hole discontinuity data. Rock Mechanics for Industry, *Proceedings of the 37th US Rock Mechanics Symposium, Vail Colorado*, June 6-9, 1999. v. 1, pp. 431-438.

8. Hammah, R. E., and J. H. Curran, 1998. Optimal delineation of discontinuity sets using a fuzzy clustering algorithm. *International Journal Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, 35: 4-5, Paper No. 136.

9. Hammah, R. E., and J. H. Curran, 2000. Standardization of variables for the fuzzy K-means clustering of discontinuity data. *Proceedings of the Fourth North American Rock Mechanics Symposium*, Seattle, Washington, July 31- Aug.1, 2000, pp. 659-666.

10. Gillett, S.L., 1987. *Online users manual: Extract clusters from axial data sets using the algorithm of Shanley and Mahtab.*