Maerz, N. H., and Zhou, W., 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.

MULTIVARIATE ANALYSIS OF BORE HOLE DISCONTINUITY DATA

Norbert H. Maerz,
Rock Mechanics and Explosives Research Center, University of Missouri-Rolla, Rolla, Missouri, USA

Wei Zhou,
Department of Geological Engineering, University of Missouri-Rolla, Rolla, Missouri, USA

ABSTRACT: This paper describes a new and cost effective method of analyzing hard rock discontinuities from oriented core bore hole data. This algorithm uses multivariate cluster analysis to group discontinuities (joints) into sets based on orientation and spatial position (spacing) along the bore hole, and to display the data in a three dimensional stereonet. Although drift or surface exposure mapping data allows better characterization of discontinuities, bore hole data is often more readily available, because of lower costs. In addition, bore hole data may be more useful because bore holes can be drilled to the exact location where the ground needs to be characterized and bore hole data is usually available earlier in the life cycle of an engineering project.

1 INTRODUCTION

1.1 Discontinuities in Rock Mechanics

The characterization of the structure of rock masses is an important consideration in engineering projects in rock. Often it is the nature of the discontinuities (joints, fractures, bedding planes, faults, and other breaks in the continuity of the rock) and not of the intact rock that governs the mechanical and hydrological behavior of the rock mass (Fig. 1). With a few exceptions, most of the rock masses that engineers deal with are influenced to some extent or another by discontinuities.

In rock engineering analysis it is necessary to understand the mechanical and hydrological behavior of rock masses in order to predict such aspects of design as:

1. The stability of a rock mass (how likely is it that the rock will fail, and how catastrophic will the failure be?);
2. The degree of remediation and/or ground support required (how do we make it safe?);
3. The expected amount of deformation as a result of applied structural loads (how much movement do we have to design for?);
4. The amount of effort needed to excavate the rock (do we need to use explosives, and if so how much?);
5. The degree and effect of water infiltration (how do we keep it dry?).

While we understand much about the mechanical properties of intact, solid rock, our understanding of discontinuous rock is significantly less developed.

Fig. 1. The discontinuous nature of rock masses.

Fig. 2. Scribed oriented core.

1.2 Analysis of Discontinuities

State of the Art

The current state of the art in ground (discontinuity) characterization in engineering projects consists either of drilling and logging the discontinuities in oriented core (Fig. 2) or from bore hole video, or trial excavations and mapping of the discontinuities on the walls and backs of the excavation. In analyzing the discontinuities, the attributes of the individual discontinuities must be identified, measured or characterized, and input into a predictive model. Methodologies for this have been proposed and or are being used by Kulatilake and Wu (1984; 1986), Baecher (1983), Dershowitz and Einstein (1988), Hudson and Priest (1979; 1983), Priest (1994), etc. Models in the past have tended to be physical or electrical analogs, and now are predominantly numerical, but also utilize empirical rock mass classification and prediction systems. The attributes for discontinuity classification are described in ISRM 1981, and shown figuratively by Hudson (1989) (Fig. 3). These include:

1. The discontinuity attitude (orientation),
2. The distance between adjacent discontinuities (spacing),
3. The physical extent of discontinuities (persistence),
4. The surface characteristics of the discontinuities (roughness, strength, mineralization and alteration)
5. The filling material (filling or infilling).

These parameters all in someway affect the mechanical and hydrogeological behavior of the rock.

The advantage of mapping a two dimensional exposure is of course that more and better data can be obtained, which can ultimately result in better characterization. The advantage of oriented bore hole core logging or bore hole video logging is that it is much more cost effective, and is capable of exploring much deeper, and much larger volumes of ground, an important consideration when the ground conditions are not homogeneous, or far below the ground surface.

In analysis, the joints are typically clustered or contoured into groups or sets based on their orientation only. At that point the other attributes may be summarized for each joint set.

In reality it may not be just orientation that is common among members of a set. Spacing or persistence or roughness may also play a role, and over the length of a long bore hole, orientations may not cluster very well. Do discontinuities group conveniently into sets based on more than orientation alone? Can we consider clustering of sets using orientation and other attributes in a clustering algorithm?

Actual Practice

These above basic discontinuity attributes are more or less routinely measured during drilling and drift mapping programs in significant engineering projects.

In more routine situations, such as mining development or projects with limited budgets, drift mapping is not done because of the extra costs and time involved. Drilling is however routinely carried out, and it takes relatively little extra effort to measure a wide range of characteristic properties from recovered oriented drill core. Such measurement and analysis is, however, not routine, because of the lack of useful tools which can make use of the measured data and incorporate it into a model with practical value.

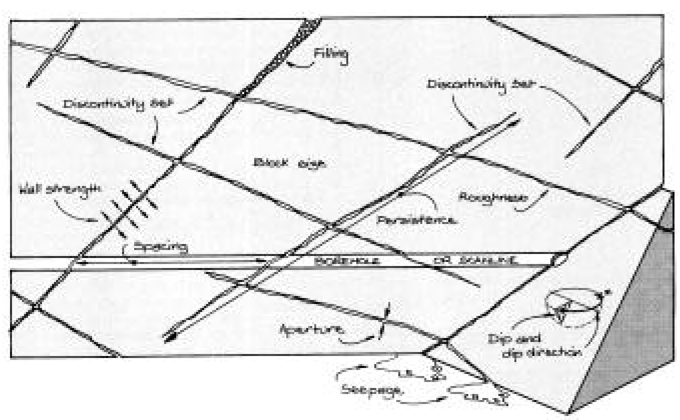


Fig. 3. Schematic of a bore hole intersecting a rock mass, (Hudson, 1989).

2 MULTIVARIATE ANALYSIS

2.1 Concept

This paper describes a new approach for the analysis of bore hole data. Rather than considering parameters such as orientation, spacing, infilling, wall rock strength, roughness and mineralization individually, a multivariate approach is used. The conventional analysis is done by grouping the discontinuities into families or sets based on orientation only, and then trying to generalize the other attributes to these groupings. While this approach works well in situations where the clustering of joints into orientation families is very obvious, this approach breaks down under typical conditions where the geologic structures vary, or the in situ stress field rotates, especially in long bore holes. Under these conditions, the joints often do not cluster well into families, and the geological engineer must start making arbitrary decisions on how to interpret this data. Consequently, the available data is typically underutilized, or not even collected in the first place.

The new approach is designed around building a new multivariate-clustering algorithm. Hammah and Curran (1998) have used joint roughness in clustering joint sets. Dershowitz, et al. (1998) have used discontinuity orientation and spacing to define structural domains.

The current approach uses both spatial data (position along the length of the bore hole), and spherical orientation data to identify joint sets. Future development will consider secondary attributions such as roughness, wall strength, fillings and moisture condition, etc. The eventual goal is to automatically identify and characterize geological domains through which the bore hole passes.

Consider the lower hemisphere stereo plot of Fig. 4a, showing four poles to discontinuity planes. If we plot each joint normal on a separate stereonet Fig 4b, and stack them on top of each other (Fig. 4c), we get a “three dimensional stereonet” (Fig. 5)
Fig. 7 shows a lower hemisphere stereo plot using a contrived data set, using multivariate clustering.

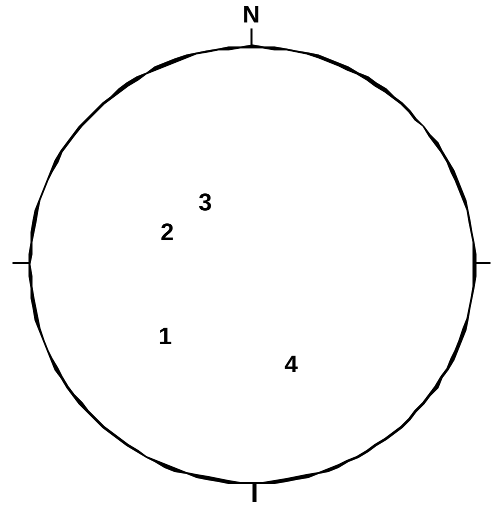
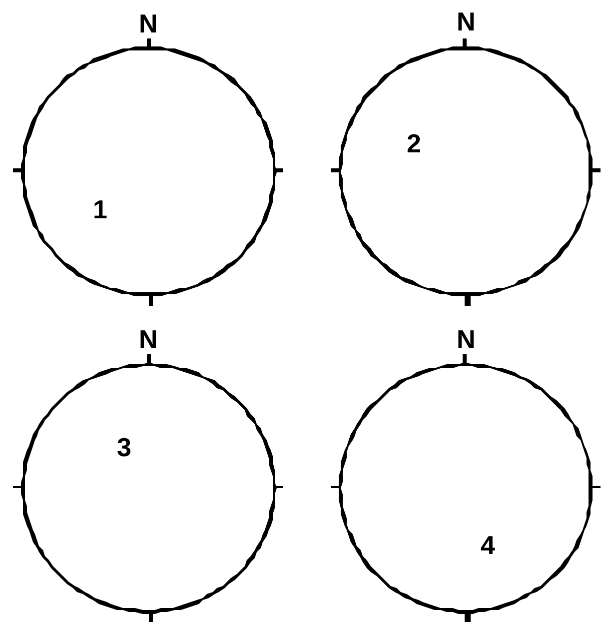


Fig. 4a (top right) A lower hemisphere stereonet with four joint normals (poles), each pole ostensibly from a different depth along an imaginary vertical bore hole.



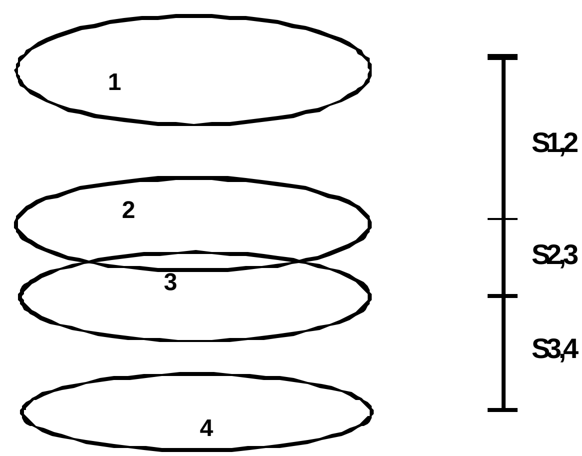


Fig. 4b (middle) Each joint normal is plotted on an individual stereonet. 4c (bottom) The individual stereonets are stacked, with each spacing proportional to the spacings between discontinuities in the bore hole. S1,2 is the spacing between discontinuity 1 and 2.

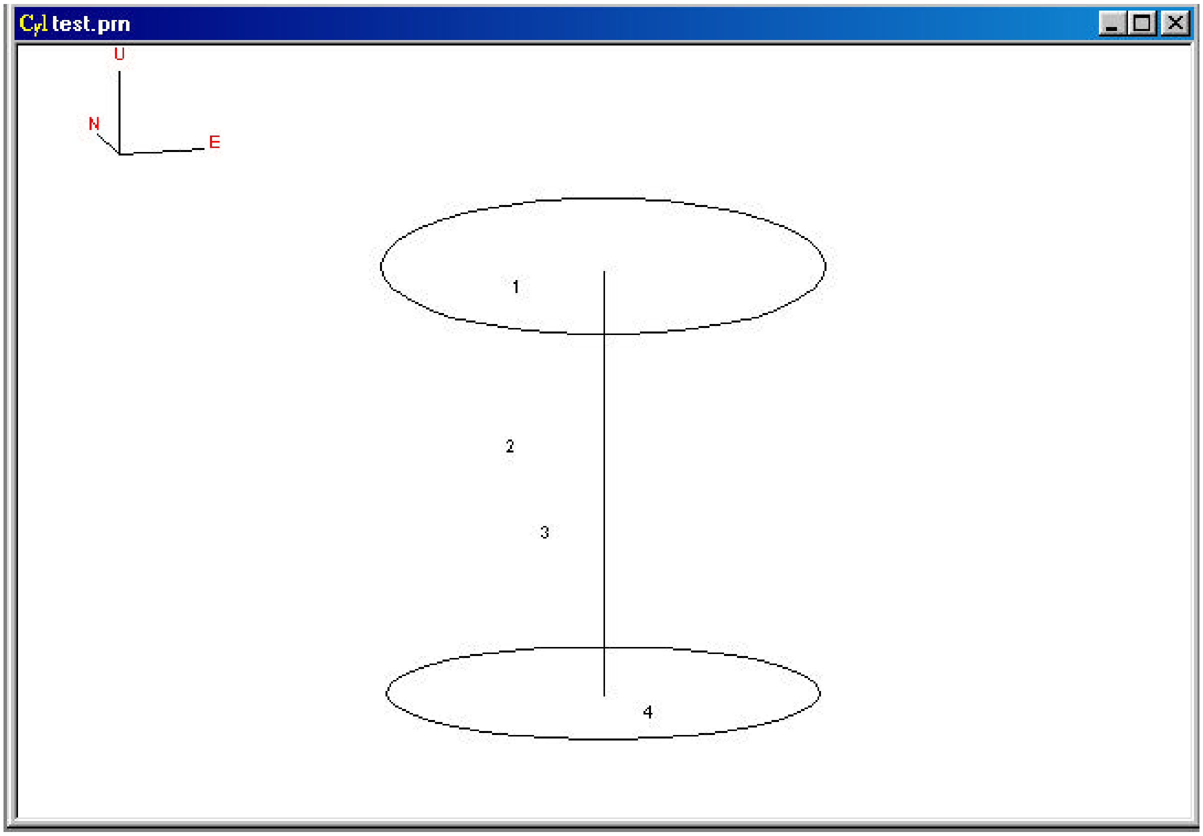


Fig. 5. "Three dimensional stereonet", showing the example from Fig. 4.

2.2 Cluster Analysis Methods

The basis of the grouping of discontinuities in Fig. 5 is cluster analysis. The two basic categories of clustering techniques are hierarchical and partitioning. One of the primary features distinguishing hierarchical techniques from partitioning techniques is that the allocation of an object (discontinuity or joint) to a cluster is irrevocable in the hierarchical technique. That means once an object joins a cluster, it is never removed and fused with other objects belonging to some other cluster.

The hierarchical scheme can be supervised, meaning the number of clusters obtained are determined by the user, or unsupervised, meaning the number of clusters are a function of a distance threshold value.

Partitioning clustering techniques do not require that the allocation of an object to a cluster to be irrevocable. Thus objects may be reallocated if their initial assignments were indeed inaccurate.

Two algorithms, the nearest neighbor or single linkage method (hierarchical) and K-mean method (partitioning), were used in this analysis. Equations for these methods were taken from Dillon and Goldstein (1984).

Nearest neighbor method

For the nearest neighbor method, the similarity between joints is based on distance measurements. Joint orientation can be expressed as a vector in spherical coordinates. The arc length between the two vectors in a spherical coordinate is adopted as the first variable. The second variable is the distance between each pair of joints in the direction down the bore hole.

An example of a measure of distance between objects in p dimensional space is defined by the Minkowsi metric:

$$d_{ij} = \left[\sum_{k=1}^{p} \left| X_{ik} - X_{jk} \right|^{r} \right]^{1/r} \qquad (1)$$

Where d_{ij} denotes the similarity distance between two objects i and j, X is the array of physical distances (arc length and spacing in our example). If we set r=2, then we get the familiar Euclidean distance between object i and j:

$$Ed_{ij} = \left[\sum_{k=1}^{p} \left| X_{ik} - X_{jk} \right|^{2} \right]^{1/2} \qquad (2)$$

Table 1 Clustering results of 6 objects, based on the number of clusters desired.

Clusters desired	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
1 cluster	1,2,3,4,5,6					
2 clusters	1,2,3,4,5	6				
3 clusters	1,2	3,4,5	6			
4 clusters	1,2	3	4,5	6		
5 clusters	1,2	3	4	5	6	
6 clusters	1	2	3	4	5	6

To implement this, a matrix of Euclidean distances is created, where the element in the ith row and jth column represents the similarity distance between objects i and j induced by the above Euclidean formula.

During the first stage, the two objects, for example i and j, with the smallest Euclidean distance are merged to form a cluster, since they are closest. The smaller Euclidean distance between i or j and all other objects are retained in the matrix, and the matrix is reduced by one row and one column. The array is now one whose elements are inter-individual and inter-group (object cluster) distances.

The above procedures are repeated to form new clusters and further reduce the matrix. In the end the last two clusters would be fused to form a single cluster containing all the individual objects. The analysis can be terminated when a pre-specified number of clusters have been identified, or by a threshold value of Euclidean distance between clusters, above which no further mergers are allowed.

Fig. 6 shows a dendrogram illustrating an example of nearest neighbor method summarizing the various stages at which merging are made. The numbers in horizontal axis represent the individual objects and the vertical axis represents the element values in an Euclidean distance matrix.

First individuals 1 and 2 are merged to form a cluster, group-individual 1-2. Then individuals 4 and 5 are merged to form a cluster, group-individual 4-5. Next inter-individual 3 and group-individual 4-5 are merged to form a cluster, group-individual 3-4-5. Eventually, individual 6 and group individual 1-2-3-4-5 fusion to form a single cluster containing all six individuals.

The individuals joined together first have the most similarity and individuals joined together last have least similarity. In the above example individual 2 has most similarity with 1, while 6 has least similarity with 1. Depending on the number of clusters selected or determined by the threshold value of the Euclidean distance, we will get the grouping as shown in Table 1.

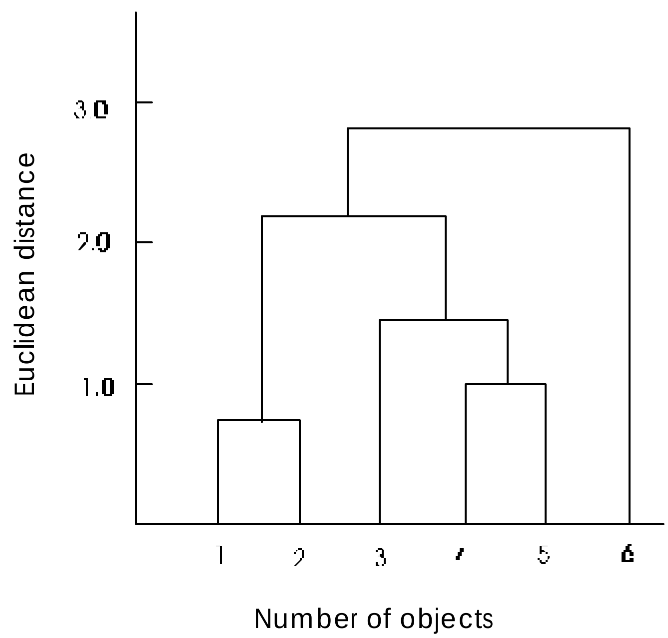


Fig. 6. Nearest neighbor dendrogram of the six objects of table 1.

K-mean method

The K-mean method uses a supervised classification for which the number of the final clusters is specified in advance.

For this method, three variables (discontinuity dip, dip direction and position down the bore hole) are considered in the analysis. Each joint is expressed in a cartesian coordinates in a cylindrical space. The relationships between the cartesian coordinates (x,y,z) and the dip direction (β), dip angle (α), and position are as following:

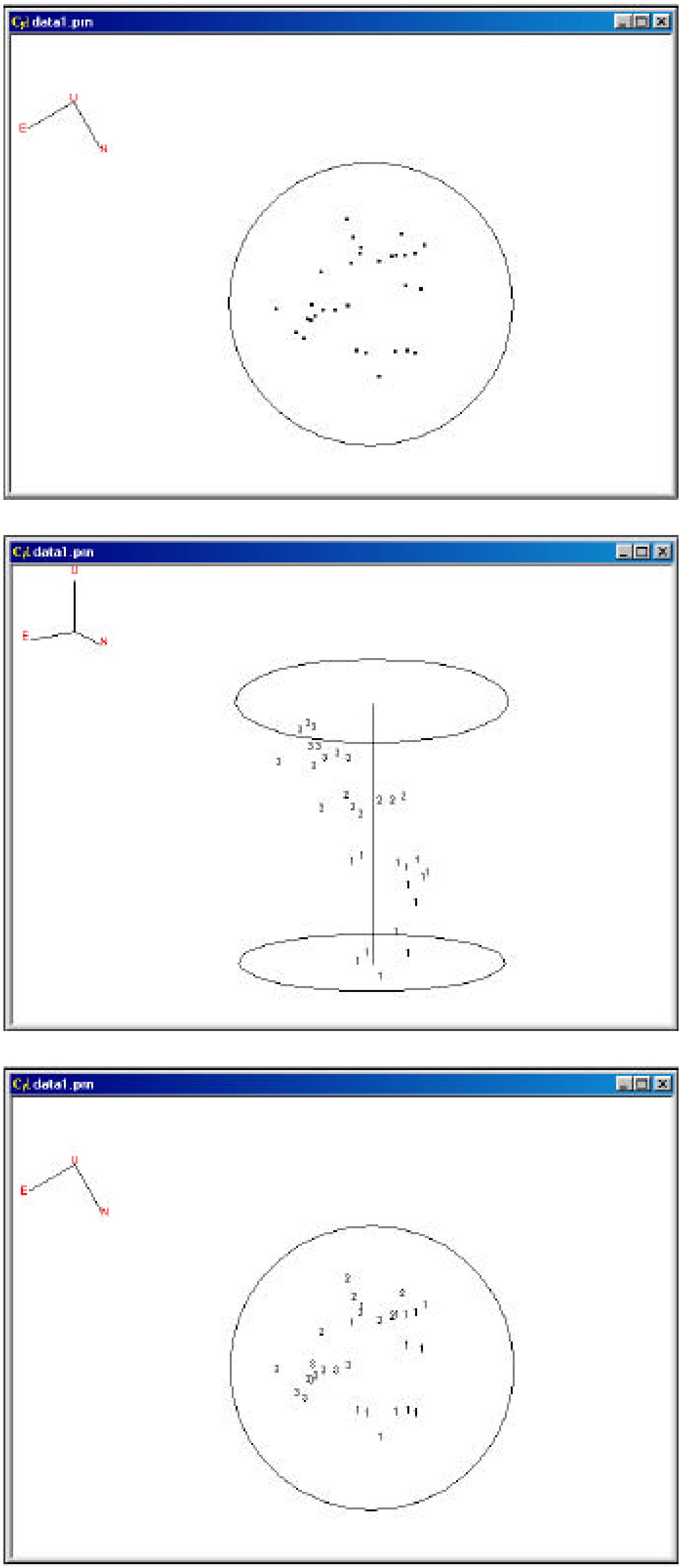


Fig. 7a (top): Traditional stereonet of contrived joint normal (pole) data. Fig. 7b (middle): Three dimensional stereonet with three joint sets clustered by multivariate analysis. Fig. 7c (bottom): Two dimensional view of Fig. 7b, also showing multivariate clustering.

$$x = \tan(\boldsymbol{b}/2) \times \sin \boldsymbol{a} \quad (3)$$

$$y = \tan(\boldsymbol{b}/2) \times \cos \boldsymbol{a} \quad (4)$$

$$z = -position \quad (5)$$

Assume we have n discontinuities and have the above three attributes (variables) of each discontinuity. We denote by x(i, j) the value of the ith joint on the j^{th} attribute; i=1, 2, …, n and j= 1, 2, 3. The mean of the j^{th} variable in the l^{th} cluster will be denoted by x(l, j), and the number of joints belonging to the lth cluster by n(l). Based upon the above notation the Euclidean distance between the ith joint and the center of the mean of l^{th} cluster can be expressed as:

$$D(i,l) = \left[\sum_{j=1}^{p} [x(i,j) - \bar{x}(l,j)]^2 \right]^{1/2} \quad (6)$$

where p is the number of dimensions, in this case three.

The specific procedures for implementation are summarized as follows:

1. To form the initial clusters, we denote the sum of all the variables for joint i as Sum(i) and denote the maximum and minimum values of Sum(i) by Max and Min respectively. The initial clusters are formed by considering object (joint) i as part of the l^{th} cluster, where l is the cluster number:

$$l = (\text{int}) \left[\frac{k \times (Sum(i) - Min)}{(Max - Min)} + 1 \right] \quad (7)$$

where k is the number of clusters desired.

2. Calculate the arithmetic mean of the three attributes of each initial cluster.
3. Calculate the Euclidean distance between the i^{th} joint and lth cluster as given by equation 6, in which p=3.
4. Check to see if any joint should be re-allocated from one cluster to another by considering the error reduction function:

$$R_{l(i),l} = \frac{n(l) D(i,l)^2}{n(l)+1} - \frac{n(l(i)) D(i,l(i))^2}{n(l(i))-1} \quad (8)$$

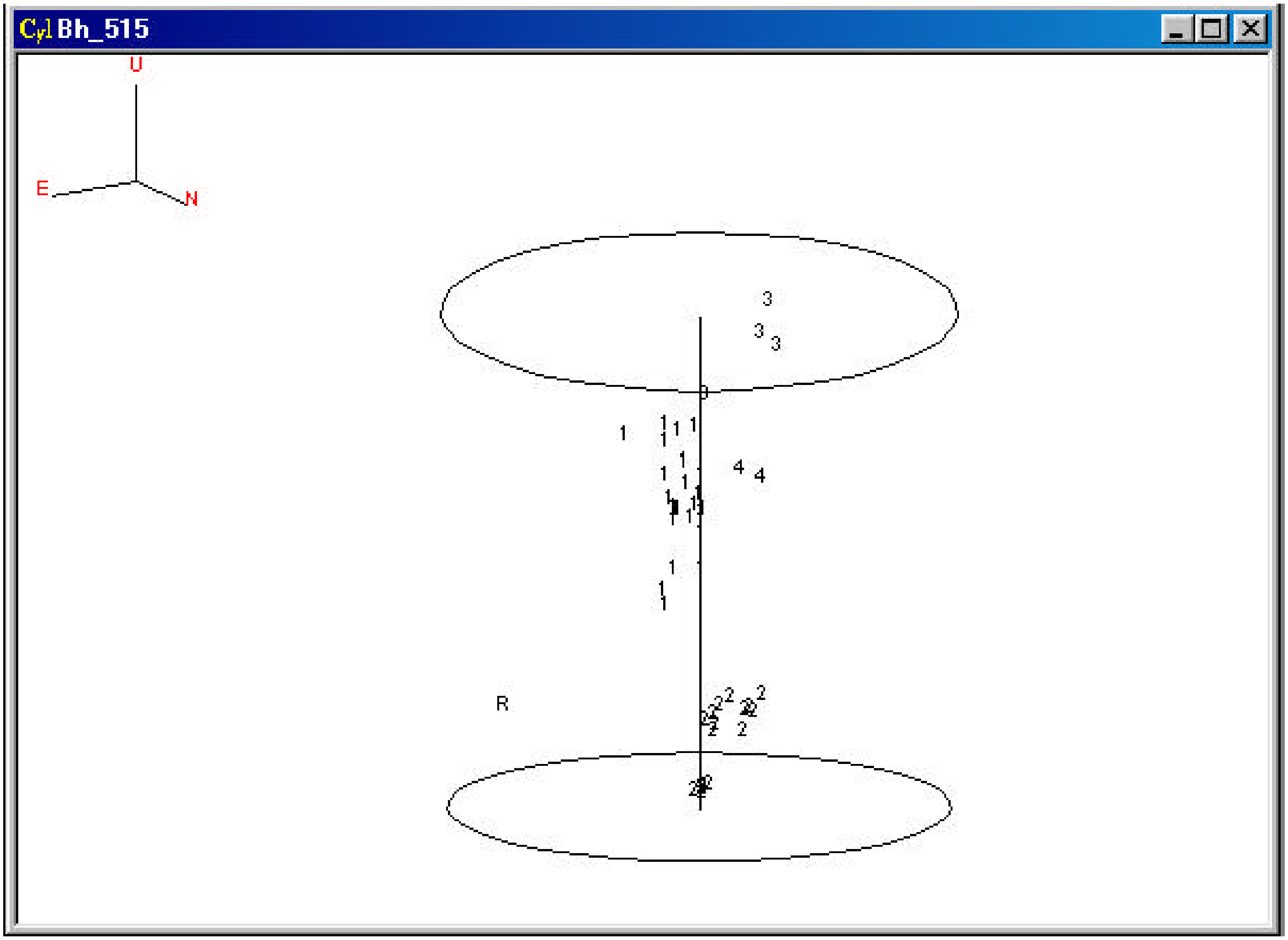


Fig. 8. Example of data from a 27 m hole in a granodiorite gneiss. Numbers indicate joint clusters; R is a random joint.

2.3 Example

Fig. 8 illustrates an example from a 27m vertical bore hole in a weakly foliated gneiss with well developed jointing. This illustrates two dominant jointing clusters (1 and 2) found at different depths. The argument could be made that the two represent different geological domains. Two other minor clusters appear, as well as a random joint.

3 DISCUSSION

The concept of multivariate analysis of bore hole data is a much overdue innovation. This paper presents a new method to cluster joints using both orientation and spacing data that promises to be automatic, fast and cost efficient.

Further research will be done on other clustering methods. It will also need to address the use of other attributes such as joint roughness into the analysis, and the automatic determination of geological domains. The issues of bore hole orientation bias, and analyzing data from multiple bore holes will also need to be addressed.

4 ACKNOWLEDGEMENTS

The authors would like to acknowledge the University of Missouri System Research Board for funding of this research.

5 REFERENCES

Baecher, G. B. 1983. *Statistical analysis of rock mass fracturing*. Journal of Mathematical Geology 15 (2): 329-347.

Dershowitz, W. & Einstein, H., 1988. *Characterizing rock joint geometry with joint* *system models*. Rock Mechanics and Rock Engineering 21:21-51.

Dershowitz, W., LaPoint, P., & Cladouhos, T., 1998. *Derivation of fracture spatial pattern parameters from bore hole data*. Int. J. Rock Mech. & Min. SCI. 35:4-5, Paper No. 134.

Dillon, W. R. & Goldstein, M. 1984. *Multivariate analysis methods and applications*. John Wiley & Sons, New York.

Hammah, R. E., & Curran, J. H. *Fuzzy Cluster algorithm for the automatic identification of joint sets.* Int. J. Rock Mech. & Min. SCI. 35:889-905.

Hudson, J. A. 1989. *Rock mechanics principles in engineering practice.* Butterworths, London.

Hudson, J. A. & Priest, S. D. 1983. *Discontinuity frequency in rock masses*. Int. J. Rock Mech. & Min. SCI. 20:73-89.

Hudson, J. A. & Priest, S. D. 1979. *Discontinuity and rock mass geometry*. Int. J. Rock Mech. & Min. SCI. 16:339-362.

ISRM Commission on Standardization of Laboratory and Field Tests 1978. *Suggested methods for the quantitative description of discontinuities in rock masses*. Int. J. Rock Mech. & Min. Sci.15:319-368.

Kulatilake, P. H. S. W. & Wu, T. H. 1986. *Relation between discontinuity size and trace length.* 27th U. S. Symposium on Rock Mechanics:130-133.

Kulatilake, P. S. H. W. & Wu T. H. 1984. *Estimation of mean trace length of discontinuities.* Rock Mechanics and Rock Engineering 17:215-232.

Priest, S. D. 1994. *Discontinuity Analysis for Rock Engineering.* Chapman & Hall, London.