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

# Discontinuity data analysis from oriented boreholes

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

W.Zhou
*Department of Geological Engineering, University of Missouri-Rolla, Rolla, Missouri, USA*

ABSTRACT: This paper presents ongoing research in improvement of analytical tools for the characterization of rock mass structure from oriented borehole discontinuity data. Using automated multivariate cluster analysis and the "three dimensional stereonet" provides for fast and objective characterization of oriented core discontinuity data. Improvements to the CYL software include three new methods of cluster analysis, incorporation of roughness as well as orientation and spacing as variables, additional visualization modes, and an automated method to split the data set into different Geotechnical Mapping Units (GMU). In addition the program facilitates the selection of optimum drilling angle for boreholes.

## 1 INTRODUCTION

### *1.1 Rock mass characterization*

The characterization of the structure of rock masses is an important consideration in engineering projects in rock. Often it is the discontinuities or joints and not the intact rock that governs the mechanical and hydrological behavior of the rock mass (Figure 1). Rock characterization using oriented bore-hole core data or borehole camera data, although not as comprehensive as mapping trial excavations or outcrops, is often more useful because it is cost effective, and can target the exact location of a proposed underground structure. But because of the lack of effective tools for the interpretation of this data, it is more often than not underutilized (Maerz & Zhou 1999).

In a previous paper (Maerz & Zhou 1999) a new approach to the analysis of oriented borehole discontinuity data was introduced. Rather than considering parameters such as orientation, spacing, infilling, wall rock strength, roughness and mineralization individually, a multivariate approach was used. The new approach is designed around building a new multivariate-clustering algorithm, utilizing both spatial data, and spherical orientation data. The data is presented in terms of a "3 dimensional" stereonet (Figure 2) where joint normals are plotted on individual "stacked" stereonets, where 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 joint corresponding to that joint normal intersects the bore hole.

In this paper we report advances in the method. In addition to the nearest neighbor and K-means method of cluster analysis, we have added three more methods: an unsupervised nearest neighbor method, a fuzzy C-means method, and a vector quantization method. Unsupervised methods do not require a-priori knowledge of the number of sets, and consequently, are less subjective. Roughness has been added to the analysis as a fourth parameter in the multivariate analysis. In addition analytical methods are developed to use the results of the analysis of one borehole to determine the optimum drilling angles for subsequent boreholes in terms of optimizing the number of discontinuity intersections, as a function of the discontinuity orientations.

Figure 1. Powerhouse excavation in northern Manitoba, Canada, showing discontinuity controlled failures in granitic gneiss.

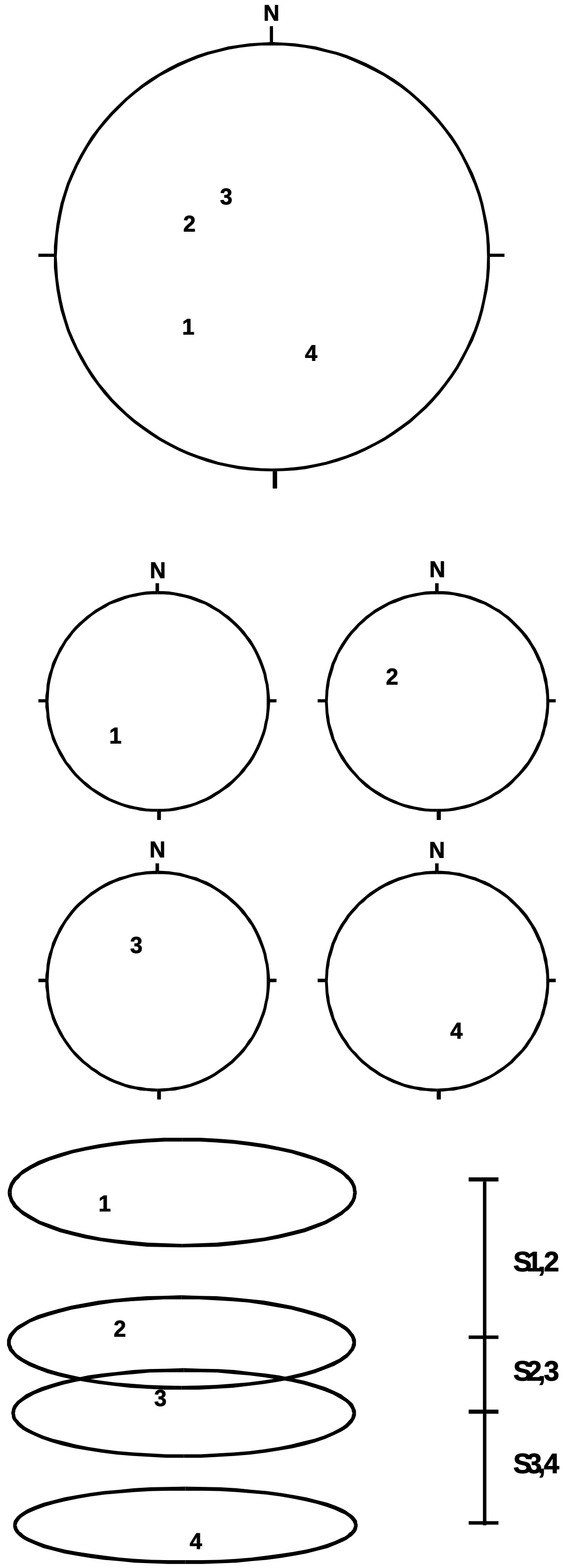



Figure 2a. Top: A lower hemisphere stereonet with four joint normals (poles), each pole ostensibly from a different depth along an imaginary vertical bore hole. Middle: Each joint normal (pole) is plotted on an individual stereonet. Bottom: The individual stereonets are stacked, with each spacing in proportion to the spacing between discontinuities in the borehole. S1,2 is the spacing distance between discontinuity 1 and 2. (Maerz & Zhou 1999).

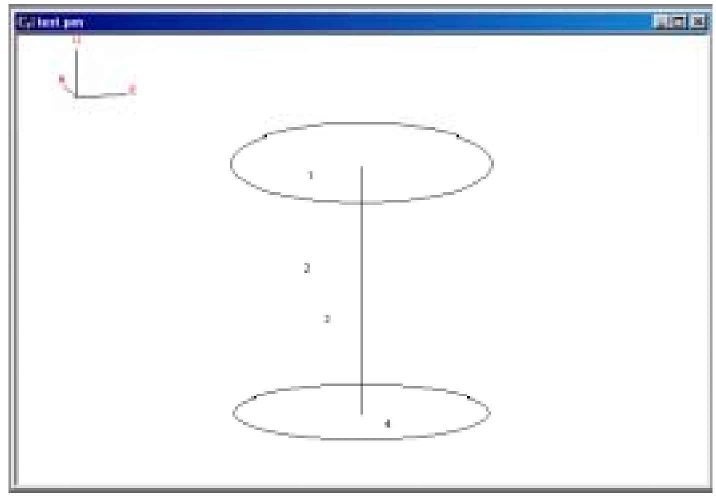


Figure 2b. “Three dimensional stereonet”, showing the example from Fig. 1a. (Maerz & Zhou 1999).

## 2 NEW CLUSTER ANALYSIS METHODS

### *2.1 Previous methods*

Two basic categories of clustering technique, the nearest neighbor or single linkage method (hierarchical) and K-means method (partitioning), were presented in a previous paper (Maerz & Zhou 1999). Three additional methods are presented here, an unsupervised nearest neighbor method, a fuzzy C-means method, and a vector quantization method.

### *2.2 Unsupervised nearest neighbor method*

For the nearest neighbor method (Dillon & Goldstein 1984), 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 borehole.

An example of a measure of distance between objects in p dimensional space is the Euclidean norm:

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

Where: $d_{ij}$ denotes the similarity distance between two objects i and j,
$|X_{ik}-X_{jk}|$ is the array of physical distances (arc length and spacing in our example).

Since the orientation measure is in terms of arc length on the unit sphere and spacing is in terms of meters or feet, the values of the individual parameters are normalized by dividing by their standard de-

viation. In addition, each of the parameters has an optional weighting factor in the algorithm.
To implement the nearest neighbor method, a matrix of Euclidean norms is created, where the element in the ith row and jth column represents the dissimilarity 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. Instead of terminating the analysis by a pre-specified number of clusters, the analysis is terminated when a threshold minimum value of Euclidean distance is reached, after which no further mergers are allowed.

### *2.3 Fuzzy C-means method*

The fuzzy c-means method (FCM) developed by Zadeh (1965) first introduced fuzzy sets as a new way to represent vagueness in everyday life. The "fuzzy" concepts have been widely adopted in the fields such as pattern recognition, neural networks, image processing, and expert system. However the geo-science literature has paid very little attention to fuzzy logic in this context, the notable exceptions being Harrison (1992) and Hammah & Curran (1998).

For clustering of discontinuity data this approach is useful because in many cases the clusters can have a wide dispersion, and the boundaries between clusters can be vague.

The FCM algorithm is based on minimization of the following objective function (Bezdek 1981):

$$J_m = \sum_{j=1}^{n} \sum_{i=1}^{k} \left(u_{ij}\right)^m d^2\left(X_j, C_i\right) \qquad k \le n \qquad (2)$$

Where: $u_{ij}$ is the fuzzy membership,
$X_j$ is the measure of individual parameter,
$C_i$ is the cluster centroid,
k is the number of clusters,
n is the number of data points,
$d^2(X_j, C_i)$ is any inner product metric (distance between $X_j$ and $C_i$), and,
m is the degree of fuzzification and is a real number greater than 1.

The variable m controls the fuzziness of the memberships. As the values of m become progressively higher, the resulting memberships become fuzzier. No theoretical optimal value for m has been determined, however, m=2 is most commonly used by researchers (Bezdek 1981, Gath & Geva 1989, Hammah & Curran 1998).

The FCM algorithm is implemented by the following procedures:

1. Selecting k number of initial cluster centers.
2. Calculating the degree of membership of all data points in all clusters.
3. Re-calculating the centroids of the clusters and updating the degree of membership from $u_{ij}$ to $\hat{u}_{ij}$. For orientation data eigenanalysis is used for finding means (Hammah & Curran 1998), for non-orientation data weighted mean is used.
4. Repeating steps 2 and 3 until the following termination criterion is met:

$$\max_{ij} \left| u_{ij} - \hat{u}_{ij} \right| < \varepsilon \qquad (3)$$

where ε is a termination criterion between 0 and 1.

### *2.4 Vector quantization method.*

The vector quantization method (VQM) is an unsupervised learning method (Pandya & Macy 1996). It begins with no clusters allocated. The first object (discontinuity) will force a cluster to be created to hold it. After that, with each input, the Euclidean distance between it and any allocated clusters is calculated. Once the distance between the current joint and all allocated clusters is known, the cluster closest to the input object may be chosen so that (Pandya & Macy 1996):

$$\left| X^{(p)} - C_k \right| < \left| X^{(p)} - C_j \right| \quad j = 1, \ldots, M \; (j \ne k) \qquad (4)$$

where: $X^{(p)}$ is the $p^{th}$ input vector,
$C_k$ is the cluster closest to the input discontinuity,
$C_j$ is the $j^{th}$ cluster center, and,
M is the number of allocated clusters.

After the closest cluster has been chosen, the Euclidean distance must be tested against a user selected distance threshold. If the Euclidean distance is less than the distance threshold, then the current member joins this cluster. If not, a new cluster is allocated. Once membership to a cluster has been established, the center of newly modified cluster must be re-calculated. The new cluster center is formed by taking the mean value of all members.

## 3 INCORPORATION OF ROUGHNESS PARAMETER

### *3.1 Introduction*

In addition to orientation and spacing as multivariate clustering parameters, the use of roughness is considered a useful parameter. Hammah & Curran (1998) have used joint roughness in clustering joint sets. In this implementation, roughness is measured as a JRC (joint roughness coefficient) as described by (Barton and Choubey 1977).

### *3.2 Implementation*

The implementation of roughness as a clustering parameter is no different than orientation and spacing. For example in the case of the nearest neighbor method (equation 1), the value of p is 3, meaning a three parameter multivariate analysis, roughness, spacing, and orientation, and $X_1$ is orientation, $X_2$ is spacing, and $X_3$ is roughness.

Since the orientation measure is in terms of arc length on the unit sphere, spacing is in terms of meters or feet, and roughness is in terms of JRC units, the values of the individual parameters are normalized by dividing by their standard deviation. In addition, each of the parameters has an optional weighting factor in the algorithm.

### *3.2 Example*

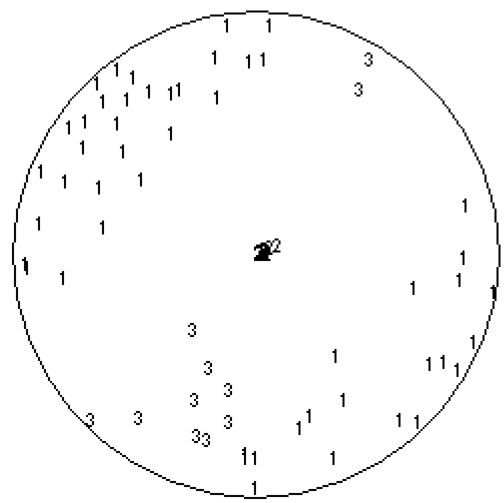
An example analysis using roughness is given in Figure 3. For this example the discontinuities along a section of outcrop of Rubidoux Sandstone were mapped along a sub-horizontal scanline. The scan line was oriented such that it crossed the near horizontal bedding planes as well as the sub-vertical fractures that were present. The data was then analyzed with CYL.

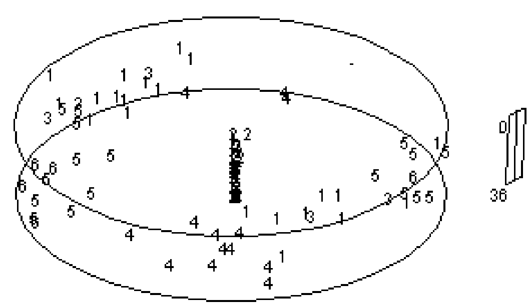
Standard cluster analysis, using orientation only, would indicate two or three specific joint sets, with a high degree of scatter in the data (Figure 3).

Clustering results using multivariate analysis show near horizontal smooth bedding with JRC values of 3.2 (set 2), and a series of five sub-vertical joint sets.

Sets 6 and 3 are distinguishable by their high mean JRC values (11.2, 12.4), and are distinguished from each other by their position along the scanline. These may represent blast-induced fractures.

Sets 1, 4, and 5 are distinguishable by their moderate values of mean roughness (JRC 5.9, 6.7, 6.1). Sets 1 and 4, 5 are distinguished from each other by their mean position along the scanline, while sets 4 and 5 are distinguished from each other by their substantially different dip azimuths.





| set # | dip dir | dip ang | position | roughness |
|---|---|---|---|---|
| 1 | 85 | 77 | 9.7 | 5.9 |
| 2 | 184 | 4 | 22.4 | 3.2 |
| 3 | 80 | 82 | 14.6 | 12.4 |
| 4 | 321 | 74 | 27.3 | 6.7 |
| 5 | 76 | 80 | 25.2 | 6.1 |
| 6 | 56 | 85 | 34.5 | 11.2 |

Figure 3. Top: Hwy outcrop of Rubidoux Sandstone, which was sampled by a sub-horizontal scanline. Upper middle: Standard cluster analysis, orientation only. Lower middle: CYL Cluster analysis of the scanline data using orientation, spacing, and roughness. Bottom: CYL Output chart showing, for each cluster, mean orientation (degrees), mean position along the scanline (feet), and mean roughness (JRC).

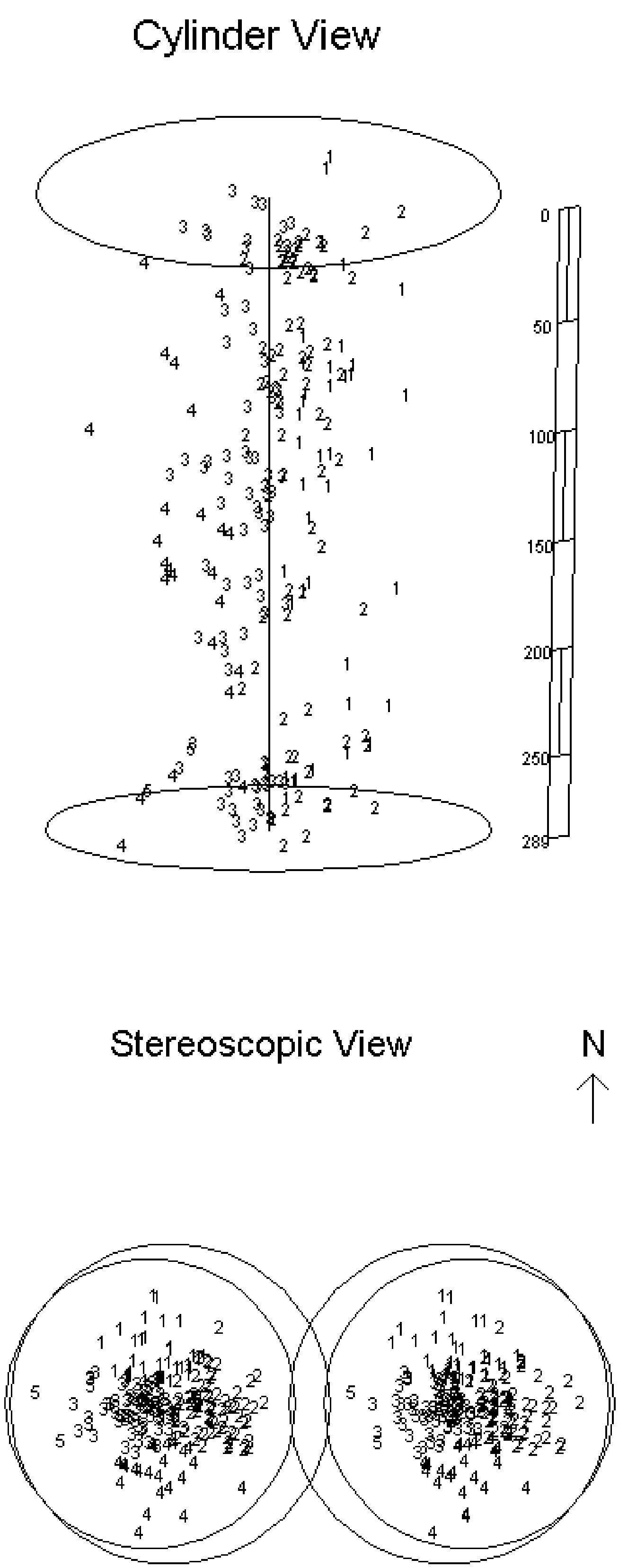



Figure 4a. Top: "3 dimensional" stereonet view showing joint sets clustered by multivariate analysis. Bottom: Stereo pair view (not to scale).

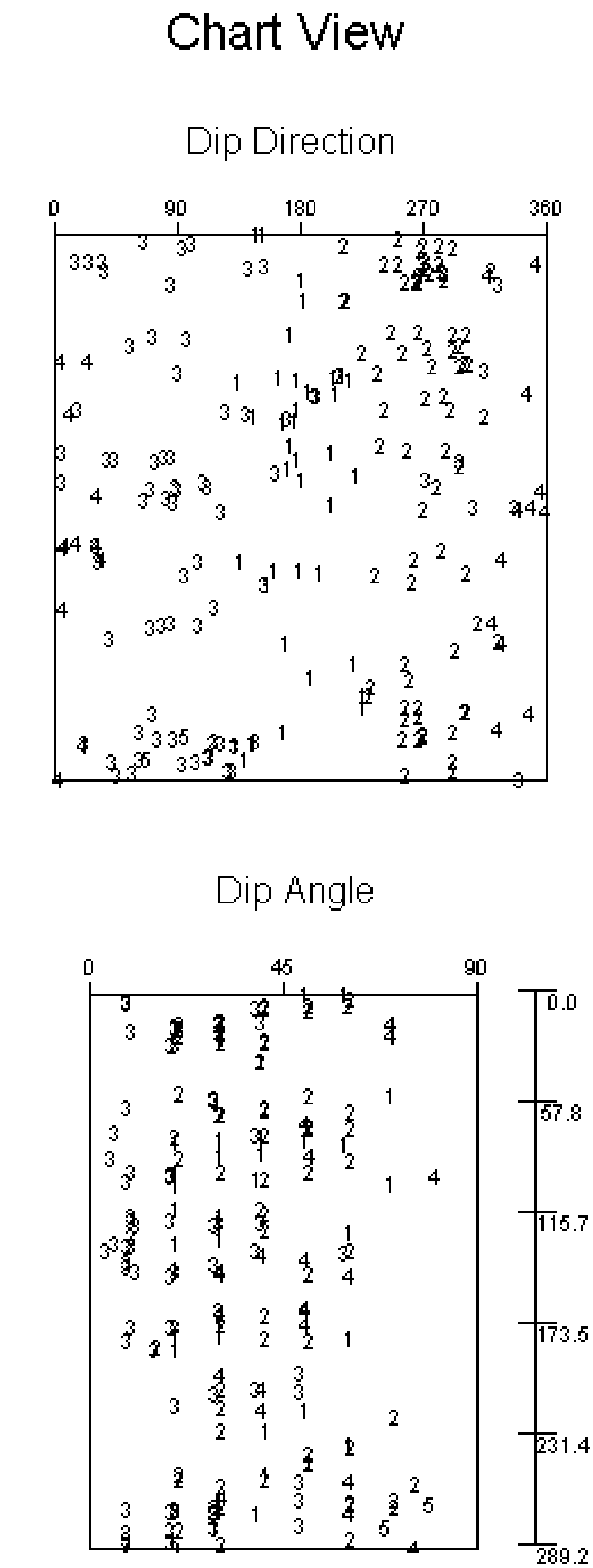



Figure 4b. Top: Dip Direction vs. depth by set. Bottom: Dip angle vs. depth, depth in feet.

## 4 PROGRAM CYL

### *4.1 Visualization improvements*

In addition to doing the clustering analysis, and reporting the results, the computer program "CYL" is extremely useful as a visualization tool. With it you can simultaneously view the projection of the discontinuity orientations, the position of the discontinuities down the borehole, and the identification of the cluster to which it has been assigned.

Figure 4a shows the cylinder view, which can be rotated and viewed from any perspective on the computer screen. The stereoscopic view is designed to be viewed with a pocket stereoscope, when printed out at the correct scale.

Figure 4b shows the chart view, where dip direction and dip angle are separately plotted against depth (position along the borehole).

### *4.2 Visual data splitting*

Long boreholes may traverse through more than one geological or structural domain. Consequently, during the analysis, it often proves useful to split the data set into different Geotechnical Mapping Units (GMU).

Figure 5 shows an example of a data set that is split into two sets by the user, by simply clicking on the point where the data is to be split. This results in two new windows opening for each data set. Each of the new windows can be saved as separate data files, and each of the new data files can further be split, as required.

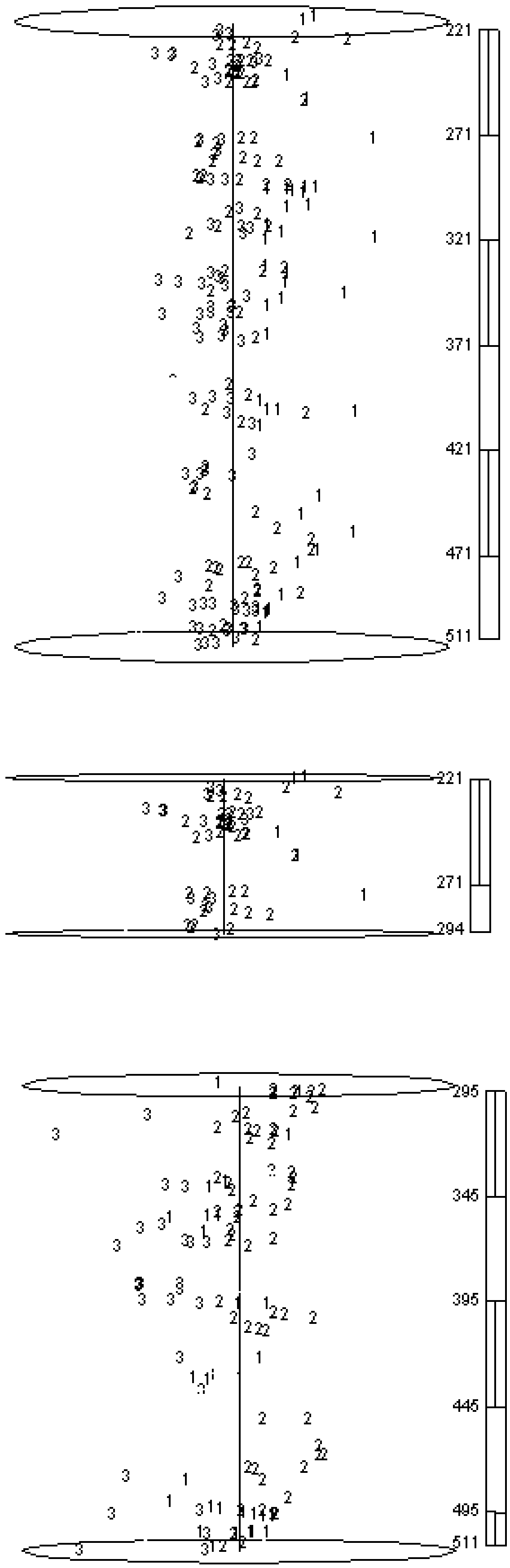



Figure 5. Top: A single data set. Manual selection of a split point results in two separate data sets.

## 5 OPTIMUM DRILLING ANGLES

### *5.1 Introduction*

Analytical tools are only as good as the data upon which they are based. In order to optimize the analysis the drilling needs to be optimized

In the case of the analysis of oriented core drilling, there is a directional bias, first documented by Terzagi (1965). Discontinuities that are near perpendicular to the borehole are much more likely to be intersected during the drilling process than discontinuities that are near parallel to the borehole. Consequently, a borehole that is optimally oriented with respect to the structure orientations will yield the most accurate data. In addition, an oriented drilling program incurs significant drilling costs, and in order to maximize efficiency, it is highly desirable to intersect as many discontinuities as possible in a given borehole.

Thus, the prediction of optimum drilling angles is of great importance.

The proposed method to find the optimum drilling direction is based on the analysis of linear sampling bias, assuming that there is some a-priori knowledge of the structure. This bias is quantified by a linear sampling bias index, which is a function of the relative angle between the orientation of the borehole and the mean orientation of the normals of each of the joint sets. The optimum drilling direction is the direction along which the minimum linear sampling bias index is obtained.

In order to quantify the bias due to linear sampling in joint surveys, the term linear sampling bias index (LSBI) is defined as:

$$F(\beta_i) = \sum_{i=1}^{n} \left( \frac{1}{\sin \beta_i} \right) \tag{5}$$

Where: $F(\beta_i)$ is the linear sampling bias index.
n is the number of joint sets.
$\beta_I$ is the angle between the borehole inclination and the dip of the $i^{th}$ joint set.

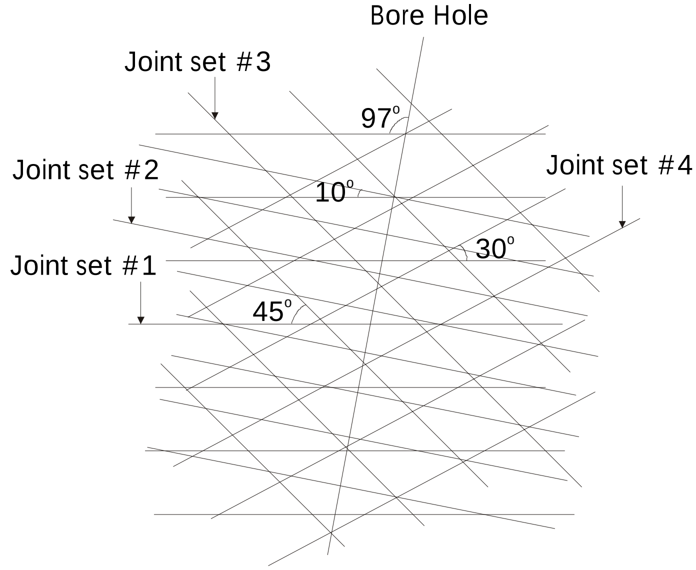
In identifying the optimum borehole inclination angle, the angle ($\beta_i$) between the borehole inclination and the joint sets inclination (dip) defines the linear sampling bias. Similarly, in identifying the optimum borehole azimuth angle, the angle ($\alpha_i$) between the azimuth angle of the borehole and joint sets strike defines the linear sampling bias index. The inclination azimuth angle ranges between 0° and 180°, borehole azimuth angle ranges from 0° to 360°.

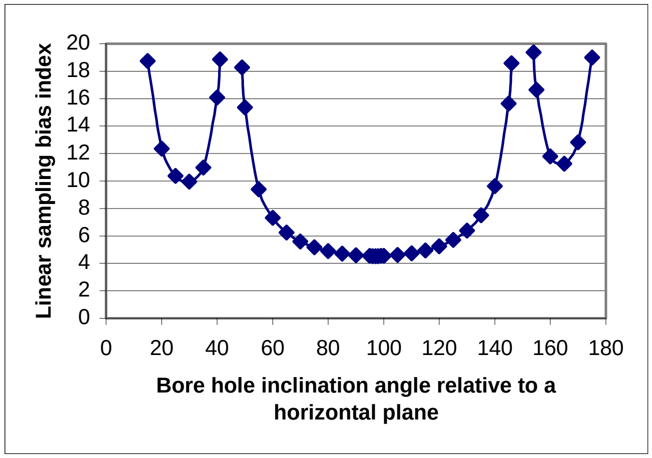
For situations with n joint sets, the optimum azimuth angle of the borehole is whenever the following value reaches the minimum:

$$F(\alpha_i) = \sum_{i=1}^{n} \left( \frac{1}{\sin \alpha_i} \right) \tag{6}$$

where $F(\alpha_i)$ is the linear sampling bias index in terms of borehole azimuth angle.
n is the number of joint sets.
$\alpha_i$ is the angle between the borehole azimuth angle and the strike of the $i^{th}$ joint set.

Consider an example of four sets of joints in Figure 6. Joint set #1 is horizontal, set #2 has a 10° dip angle and a 20° strike toward East, set #3 has a 45° dip angle and a 40° strike toward East, and set #4 has a 30° dip angle and a 150° strike toward West.



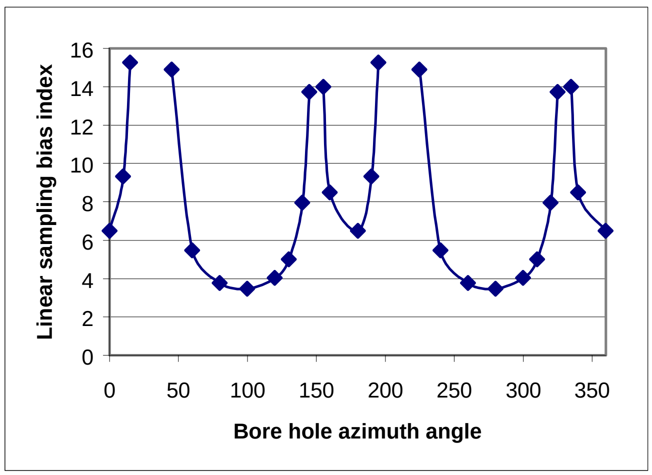




Figure 6. Top: Example of four sets of joints shown in a vertical projection. The line marked "borehole" shows the optimum drilling direction as 97° clockwise from the horizontal. Middle: For borehole inclination angle, the linear sampling bias index reaches its minimum value when the borehole inclination is about 97°. Bottom: For borehole azimuth angle, the linear sampling bias index reach cominimums at 100° or 280°. Because of the angle convention, the minimum at 100° corresponds to the inclination angle of 97°.

Figure 6 shows the LSBI as a function the angle between inclination angle of the borehole and dip angle of the joint sets. The LSBI, which is defined as:

$$LSBI = 1/\sin\beta_1 + 1/\sin\beta_2 + 1/\sin\beta_3 + 1/\sin\beta_4 \qquad (7)$$

reaches its minimum value when the borehole inclination angle is 97° (towards west). The optimum azimuth angle of the borehole is about 100°.

*5.2 Application*

Figure 7 shows the output results from program CYL. The analyzed borehole is a vertical 290-foot hole with 232 discontinuities and seven identified joint subsets.

Using this analysis, it is possible to determine that the optimum drilling direction for this situation is at an azimuth of 140°, with an inclination of 94°.

## 6 ACKNOWLEDGEMENTS

The authors would like to express our gratitude to the University of Missouri System Research Board and the National Science Foundation for funding this research. We also wish to thank Barrick Goldstrike Mine for providing discontinuity data used in this study.

## 7 REFERENCES

Barton, N., & Choubey, V., 1977. The shear strength of rock joints in theory and in practice. *Rock Mechanics* 10:1-54.

Bezdek, J. C., 1981, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press: 256 pp.

Dillon, W. R. & Goldstein, M. 1984. *Multivariate analysis methods and applications*. John Wiley & Sons, New York:487 pp.

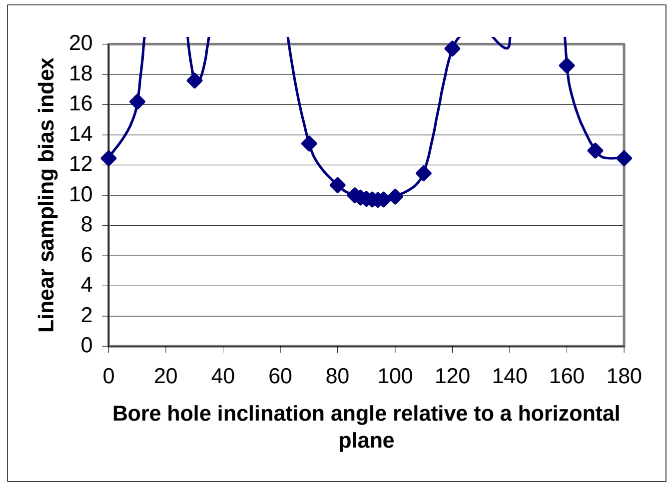
Gath, I., & Geva, A. B., 1989, Unsupervised optimal fuzzy clustering. *IEEE Trans. Pattern Anal. Machine Intell* 11(7):773-781

Hammah, R. E., & Curran, J. H., 1998, Optimal delineation of joint sets using a fuzzy clustering algorithm, *Int. J. of Rock Mech. & Min. Sci.* 35 (7):889-905.

Harrison, J. P., 1992, Fuzzy objective functions applied to the analysis of discontinuity orientation data. *Eurock'92*:25-30.

Pandya, A. S. & Macy, R. B., 1996, *Pattern Recognition with Neural Networks in C++*, CRC Press and IEEE Press: 410 pp.

| set # | dip dir | dip ang |
|---|---|---|
| 1 | 152 | 45 |
| 2 | 286 | 55 |
| 3 | 81 | 18 |
| 4 | 198 | 32 |
| 5 | 281 | 28 |
| 6 | 14 | 48 |
| 7 | 89 | 55 |



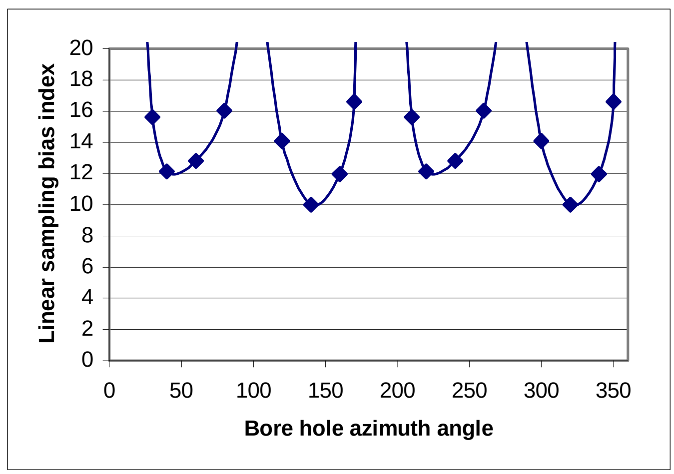




Figure 7. Top: Table showing the means of dip direction and dip angle (from the analysis of a borehole). Middle: Using this real data, the linear sampling bias index reaches its minimum value when the borehole (optimum) inclination angle is about 94°. Bottom: Graphic showing the (optimum) azimuth angle of these seven joint sets situation is either 140° or 320°. Because of the angle convention, the minimum at 140° corresponds to the inclination angle of 94°.

Terzaghi, R. D., 1965. Sources of error in joint surveys, *Geotechnique* 15:287-304.

Zadeh, L. A., 1965, *Fuzzy Sets, Information Control* 8:338-353.