Rock Mechanics
and Rock Engineering 16, 39—72 (1983)

Rock Mechanics
and Rock Engineering

Probabilistic and Statistical Methods in Engineering Geology

Specific Methods and Examples

Part I: Exploration

By

H. H. Einstein and G. B. Baecher

Massachusetts Institute of Technology, Cambridge, Mass., U. S. A.

Summary

After having pointed out weaknesses in traditional approaches to exploration in a preceding paper, it will now be shown what can be gained by using statistical methods in practical application to exploration. In particular, statistical data collection for joint surveys will be presented including methods for error reduction. Extensive treatment will be given to subjective assessment of uncertainty, a methodology that is well suited to engineering geology; its use in describing tunnel geology is given as a practical example. Finally, it will be shown how these methods of rational uncertainty description are employed to plan exploration in an optimum manner, with an example in exploration planning for underground gas storage caverns.

I. Introduction

Uncertainty about geologic conditions and geotechnical parameters is perhaps the most distinctive characteristic of engineering geology compared to other engineering fields. This is evidenced by the central role of "engineering judgement", adaptable design approaches, and other procedures for dealing with uncertainty or hedging against it. The profession has developed many qualitative strategies, and the intent of this paper is to show that most can be improved by rational analysis. Rational analysis of uncertainty usually involves probability theory and statistics. These analyses are not meant to replace present approaches — particularly engineering judgement — but to add systematic consideration which is essential to engineering decisions.

In a preceding paper (Einstein et al., 1982) sources of uncertainty and their consequences were described, followed by a description of traditional approaches for analyzing uncertainty and of their deficiences. A short summary of these deficiencies will be given below. This will be followed

Received July 9, 1982
Revised October 28, 1982

0723-2632/83/0016/0039/$ 06.80

by the main body of the text, in which improved methods in application to exploration are summarized and illustrated with practical examples. In part II, to be published later, methods used in design and construction will be discussed.

II. Traditional Approaches for Analyzing Uncertainty in Engineering Geology and Their Deficiencies

Since this topic has been discussed extensively in the preceding paper (Einstein et al., 1982), it shall only be recalled to the extent necessary to serve as an introduction to the following section.

Exploration is on the one hand affected by uncertainty, on the other hand it includes specific procedures to deal with uncertainty. Traditional approaches are however often unsatisfactory:

— Current exploration practice is seldom founded on a systematic approach. Particularly quantitative analysis for exploration planning is exceedingly rare.

— Exploration reliability is quantified only to the extent of manufacturers' literature and limited experience with some methods. How factors affecting reliability of a particular method can be tied together, and how they can be incorporated in reliability expressions for an exploration approach is often an open question. Hypotheses on site geology or conditions are seldom compared and tested rationally.

— While the precision of individual experiments can often be established, the precision of estimates of in situ properties — and more importantly, the range of properties — cannot be assessed with deterministic, often qualitative approaches.

— Finally, economics of exploration is a continuing matter of dispute. Current programs are typically based on a percent of total project cost (e. g., 3% for large embankment dams). These numbers do however not indicate whether a program was overdesigned, components should have been planned differently, or if adding or omitting components would have been economically advantageous.

The following section will give the reader some idea on statistical methods that can reduce if not eliminate the aforementioned limitations.

III. Statistical Methods in Exploration

The use of statistical methods in geological exploration involves difficult philosophical issues about the primary products of exploration, how observations are organized and explained, and how predictions or interpolations are made. These have been discussed previously in detail (Baecher, 1977), and are not further considered here. Instead, this section concentrates on three specific examples of data collection, subjective assessment, and exploration planning.

III.1 Statistical Data Collection: Joint Surveys

For many decades, the collection of geometric data on rock mass jointing has been recognized as a problem of statistical sampling, and beginning in the mid-1960's, a number of workers have devoted effort to developing sound survey procedures and to interpreting the voluminous empirical data available.

In common joint surveys, three geometric properties are of interest. These might be recorded in a number of equivalent measures, but involve:

— *Density* — Spacings, numbers per rock volume or outcrop area.

— *Size* — Trace lengths, areas, radii.

— *Planear Orientation* — Strike and dip, direction cosines of pole, azimuth and dip.

It is important to recognize that these properties manifest in observed data in interdependent ways. The measures commonly used to describe joint survey data are, in fact, only facets of more fundamental description.

Empirical Results

Typical results *for spacing distributions* are shown in Fig. 1, plotted against best fitting exponential functions, $F(s)=1-\exp\{-\lambda s\}$, in which s is joint spacing along a sampling line and λ is a parameter. The cumulative

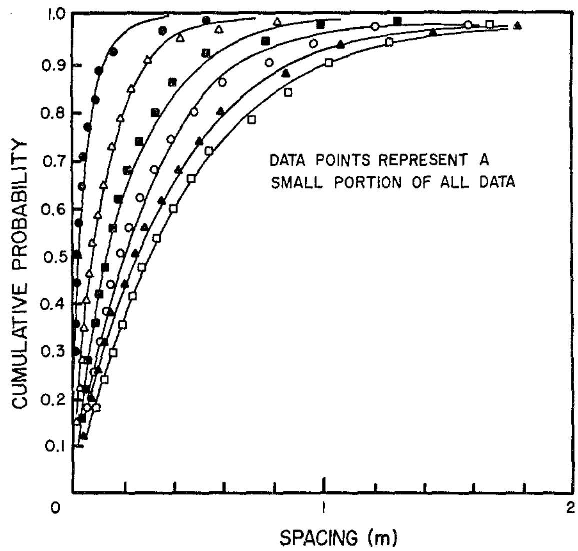


Fig. 1. Joint spacing distributions plotted against best fitting exponential functions

density functions (*cdf*) were fit using the estimator $\hat{\lambda}=\bar{s}^{-1}$, in which $\bar{s}=$ sample average spacing. Common statistical tests were used to test goodness-of-fit. In only 8% of the spacing distributions analyzed did the Exponential Model fail to satisfy goodness-of-fit criteria at the 5% confidence level. In all, some 25 data sets were evaluated for spacing, each having a sample size between 500 and 1000. The Exponentiality observed is consistent with other published results (e. g., Priest and Hudson, 1976) and is a strong verification for the Exponential Model of joint spacing.

An interesting observation is that, while average spacing varies with orientation of the sampling line, Exponentiality does not. This can be seen in Fig. 2, in which mean spacing and coefficient of variation (Cov = standard deviation/mean) along six non-coplanar directions are shown. For Exponentiality Cov = 1.0. Relationships among average spacings can be accurately calculated simply from trigonometric considerations.

Trace length distributions do not exhibit the exceedingly consistent behavior that spacings do; however, in 82% of the samples, trace lengths satisfied 5% goodness-of-fit tests for logNormality. Decreasing to 1% error allowed certain of these to pass (Table 1).

Table 1. *Results of Goodness-of-Fit Tests for Trace Length Distributions*

Site	Exponential	Gamma	Lognormal
Site A, top	fail	fail	fail
Site A, bottom	fail	fail	fail
Site A, sides	fail	fail	pass*
Site A, sides	fail	fail	pass*
Greene Co., trench A	fail	fail	pass
Greene Co., trench B	fail	fail	pass
Greene Co., trench C	fail	fail	pass
Greene Co., trench T	fail	pass	pass
Site B	fail	fail	pass
Blue Hills	fail	fail	pass
Pine Hills	fail	fail	pass

* All significance levels set at 5%, except where indicated by (*) which were set at 1%.

It is of interest to note that visual classification of trace length distributions at times can be misleading, and that clustering procedures for grouping data into histograms can mask important distributional information in a sample. Trace length data from Site B are shown in Fig. 3 against logNormal and Gamma *pdf*'s, fit by Maximum likelihood estimations. By visual inspection both the logNormal and Gamma *pdf*'s provide reasonably good fits to the data. However, χ^2 and *K-S* tests show only the logNormal to provide an acceptable fit at 5%.

If the data of Fig. 3 were clustered in five foot intervals, the histogram of Fig. 4 would result. Presuming that short trace lengths are under represented in the sample, either by design (i. e., sample truncation) or unintentionally (i. e., through sample bias), the conclusion might be adopted erroneously that the density function (*pdf*) is Exponential.

Whereas, considerable success was enjoyed in fitting analytical *pdf*'s to spacing and trace length data, the opposite was true of *orientation data*. The conclusion is interesting, in part because more statistical work has been

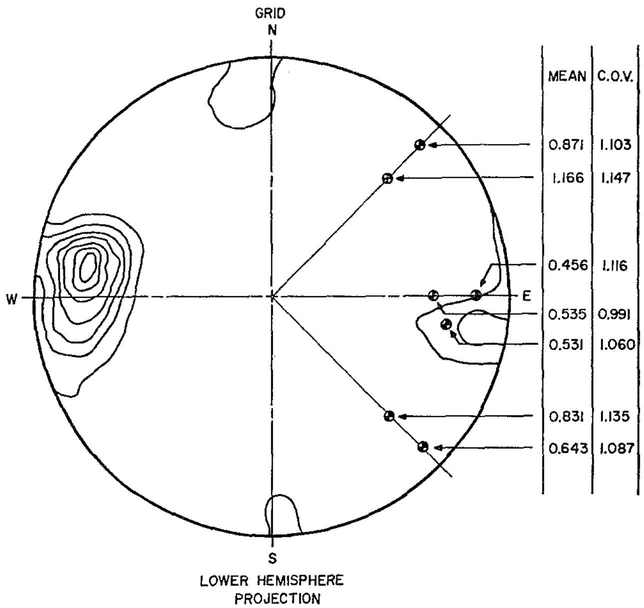


Fig. 2. Relation between spacing distribution and orientation of sampling line. For exponential distribution cov = 1.0

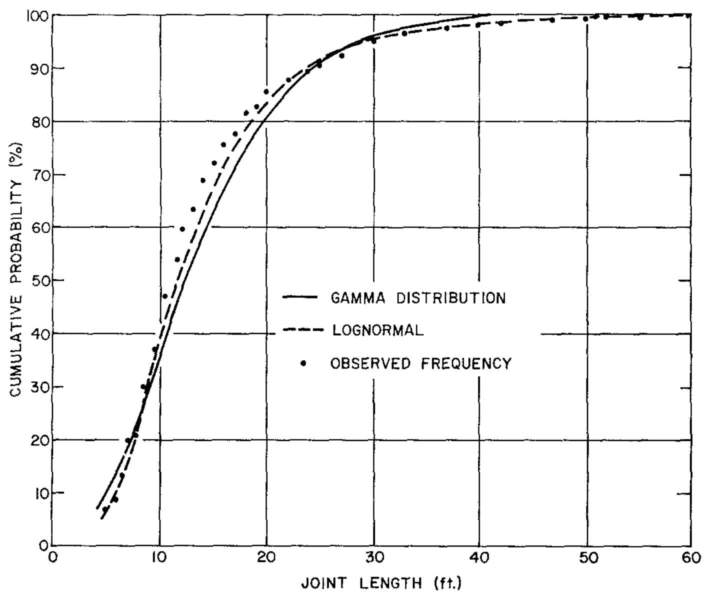


Fig. 3. Comparison of lognormal and gamma distributions vs. observed frequencies of trace lengths

Table 2. *Distributional Forms for Orientation Data*

Name	Form
Uniform	$f(\theta, \phi) \propto \sin$
Fisher	$f(J) \propto \exp\{\underline{\underline{\kappa}} J \cdot \mu\}$
Elliptical	$f(J) \propto \exp\{\mathrm{tr}\, \underline{\underline{\kappa}} J^t \underline{\underline{\kappa}}\}$
Bingham	$F(J) \propto \exp\{\mathrm{tr}\, \underline{\underline{\kappa}} \mu^t J J^t \mu\}$
Normal	$f(\theta. \phi) \propto \exp\{-1/2\, (x-\mu)^t \underline{\underline{\Sigma}}\, (x-\mu)$

in which: $J = \{l, m, n\}$, the vector of direction cosines,
$x = \{\theta, \phi\}$,
θ = spherical coordinate,
$\underline{\underline{\kappa}}$ = dispersion matrix or constant,
μ = mean vector of direction cosines or coordinate,
$\underline{\underline{\Sigma}}$ = covariance matrix,
ϕ = spherical coordinate.

Table 3. *Goodness-of-Fit for Orientation Data*

Joint set	Fisher	Bi-variate	Bingham	Pass χ^2-test
1A	−558.7	−420.3	−480.0	NONE
1B	−80.0		−8.0	NONE
1C	−294.8	−144.0	−107.0	NONE
2A	−127.1	−31.5	−47.0	NONE
2B	−71.5	−60.1	−282.0	NONE
2C	−442.2	−322.9	−283.0	NONE
2D	−131.0	−90.7	−326.0	NONE
2E			−29.0	NONE
2F	−89.4	−68.5	−51.0	NONE
3A	−125.0	−122.0	−114.0	ALL
3B			−20.0	NONE
3C	−20.0	−20.0	−91.0	NONE
4	−567.6	−554.8	−517.0	NONE
5A	−444.5	−287.7	−272.0	NONE
5B	−236.1	−141.7	−149.0	BINGHAM
6A	−79.4	−57.0	−37.0	NONE
6B	−62.3	−27.0	−27.0	ALL
7A	−383.6	−298.9	−280.0	NONE
7B	−251.8	−91.4	−140.0	NONE
7C	−574.7	−574.8	−555.0	NONE
7D	−119.9	−45.4	−41.0	BINGHAM, BIV., F.
8	−294.8	−144.0	−107.0	NONE
Total best Fits:		7.0	13.0	

Note: Best Fit is highest log-likelihood.
Normal Distribution was fit on direction cosines and results are therefore not comparable. Magnitude of log-likelihood is related to number of observations, so it is not possible to compare values between different sets. Arnold Distribution results are indistinguishable from Fisher results.

performed on the description of joint plane orientation than perhaps on all other rock mass properties taken together. Attempts to correct field data for implicit biases in sampling plans make the situation, if anything, worse.

The primary distributional forms used in the study are shown in Table 2. The Fisher, Bingham, elliptical and uniform are defined on the unit sphere; the bivariate Normal and bivariate logNormal are defined on the plane. Maximum likelihood estimators were used to fit the distributions. Results of goodness-of-fit tests are shown in Table 3. Only the most ideal pole

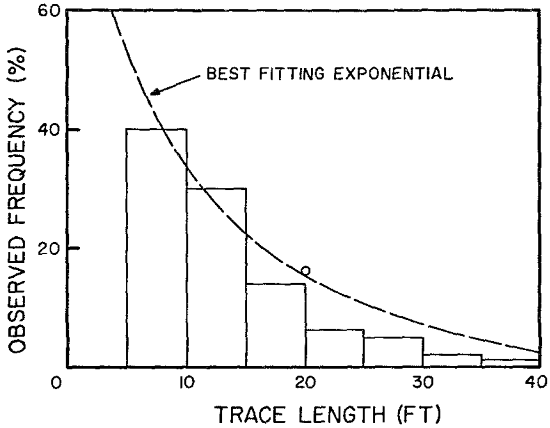


Fig. 4. Histogram of data of Fig. 8 showing effect of contour interval on inference of distributional form

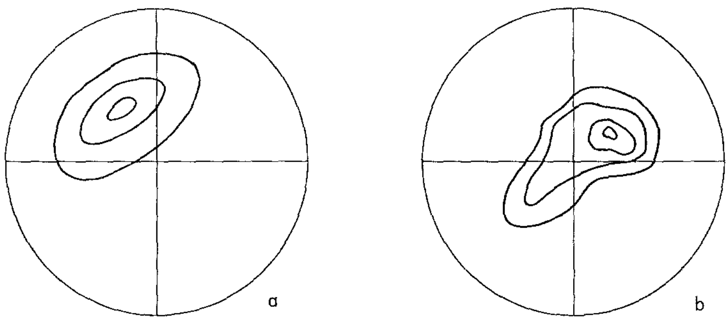


Fig. 5. Joint orientation distributions
a) Distribution can be approximated by parametric distributional form
b) Distribution cannot be approximated parametrically

distributions, like that of Fig. 5a, can be well approximated by analytical forms. Typical distributions, like that of Fig. 5b, are more erratic than allowed by the limited flexibility of standard analytical forms.

A problem in fitting distributional forms to orientation data is how to separate subparallel sets of joints from one another. A number of algorithms are available for numerically clustering orientation data, but most suffer drawbacks and experience suggests that visual clustering may lead to better results (Einstein et al., 1980).

These procedures, whether visual or numerical, partition the projective hemisphere into non-overlapping regions and treat the poles within each region as corresponding to an individual set of joints. This truncates the tails of most analytical forms that could be fit, severely complicating parameter estimation and limiting the conclusions drawn from goodness-of-fit testing. Using a mixed distribution procedure, in which the overall *pdf* is set equal to a sum of *pdf*'s, is in principle a good approach, but again leads to difficult estimation problems. Non-parametric techniques can be used as a last resort, but allows few general conclusions on distributional form to be drawn.

Sampling Errors

Errors in sampling arise from three sources: sampling error, estimation ("statistical") error, and measurement error. Sampling error is caused by plans that are not representative, estimation error is caused by statistical fluctuations from one sample to another, measurement error is caused by inaccuracies in the way individual elements are measured.

Errors and Biases in Sampling for Joint Size (Length)

— Proportional Length Bias

Presuming that outcrops or excavations are statistically independent of the joint populations to be sampled, the probability of joints intersecting a sampled surface is proportional to their size. The sampled population therefore contains traces of a disproportionate number of large joints, and does not accurately represent the population of joints within the rock mass. Outcrop geometry and location do depend on jointing, of course, but are influenced primarily by joints parallel to the outcrop surface which appear with low frequency in the sampled population.

Using intersections with an arbitrary scan-line as the sampling procedure, a second geometric bias is produced. Longer trace lengths have proportionally larger probability of interesting the line and therefore of being sampled (Baecher, 1978, Cruden, 1977, Priest and Hudson, 1981). Thus, for inferring the size of trace lengths on the outcrop, the sample is linearly biased. For inferring the size of joints within the rock mass, the sample is quadratically biased.

The effect of a linear bias is shown schematically in Fig. 6. The probability of a trace length l appearing in the sample is the product of the probability of it appearing on the outcrop, $f(l)\,dl$, and the conditional probability of it intersecting the sampling line if it does appear on the outcrop, kl,

$$f_s(l)\,d = kl\,f(l)\,dl, \tag{3.1}$$

in which $f(l)$ = the *pdf* of trace lengths in outcrop, and k = a normalizing constant, which can be shown to equal the reciprocal of the mean of l on the outcrop (Priest and Hudson, 1981). Any higher order bias introduces the conditional probability kl^n, in which k equals the reciprocal of the n-th central moment of $f(l)$.

An interesting property of the linear (or higher order) bias is that it serves as a filter that transforms many common distributions $f(l)$ into approximately logNormal forms. In the sense of common goodness-of-fit tests

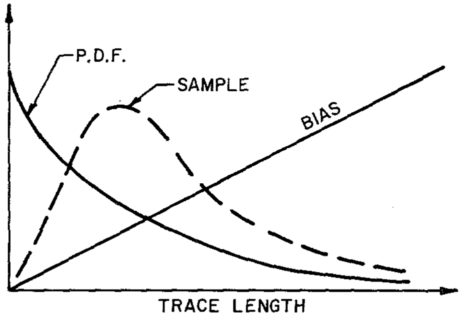


Fig. 6. Simple length bias in sampling trace lengths

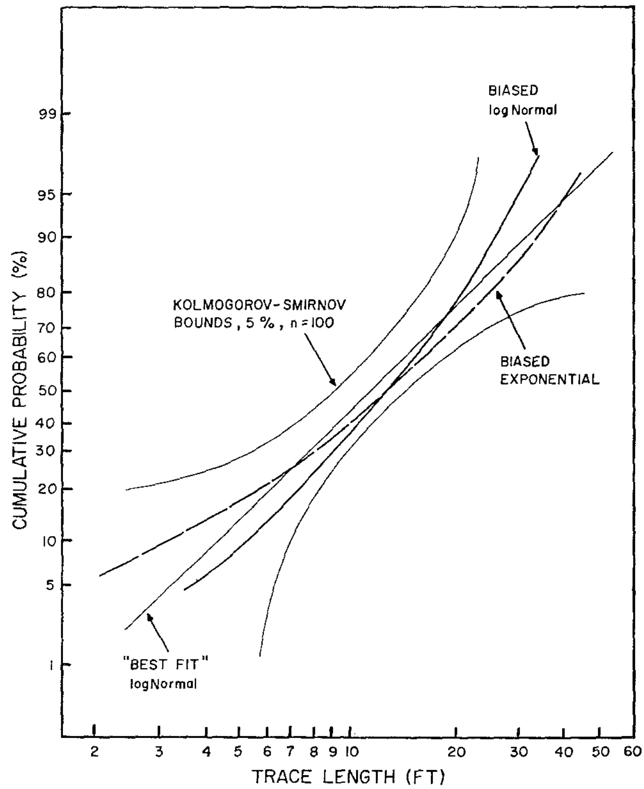


Fig. 7. Sample distributions of exponential and logNormal distributions filtered through a linear bias note, both pass standard goodness-of-fit distributions for logNormality (After Baecher and Lanney, 1978)

these transformed *pdf*'s are indistinguishable from logNormal *pdf*'s at realistic sample sizes. This is demonstrated in Fig. 7 in which linearly biased

Exponential and logNormal $f(l)$'s are tested against best fit logNormals and shown to satisfy *K-S* criteria at the 5% level. Since size biases are common in geological sampling, it is interesting to speculate that the common observation of logNormal *pdf*'s for geometric properties is primarily an artifact of sampling procedures.

— Censoring Bias

The trace length data of Fig. 8 were collected as area samples (i. e., every joint within a very large sampling field was measured) at the ground surface ("top of rock") and on the floor of a 20 m deep excavation ("bottom").

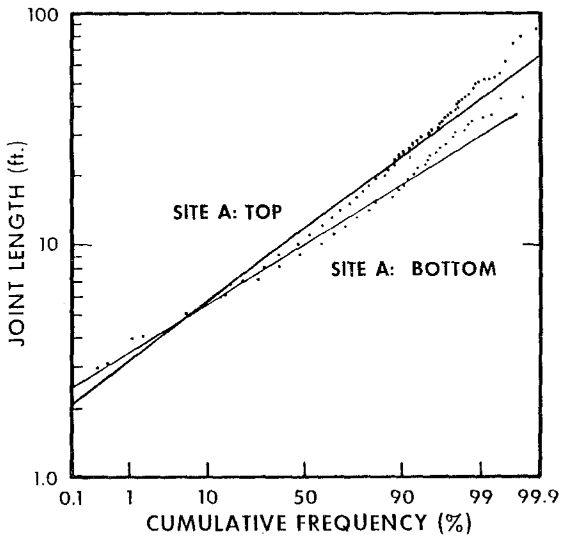


Fig. 8. Effect of censoring bias on joint length distribution

The joint populations are for present purposes essentially identical, and yet the bottom of rock sample has a lower mode and a much thinner upper tail. The reason is that many of the traces observed in the excavation run off into the rock walls, and cannot be observed in their entirety. Since this censoring occurs with proportionally higher probability to longer traces, the sample is biased toward shorter lengths and the extreme upper tail disappears completely.

Censoring is a well known sampling problem in life testing and other fields of statistics. For the traditional problem in which the point of censoring is constant (i. e., all traces longer than l_c are censored and shorter than l_c are observed completely) a large literature of both frequentist and Bayesian methods has been developed. Primarily, this literature deals with Exponential distributions (Epstein, 1959, Kendall et al., 1967), but results also exist for other forms (Fisher, 1931, Hald, 1949). The question of fixed-point censoring for joint surveys has been considered by Cruden (1977), Baecher and Lanney (1978) and Priest and Hudson (1981).

Unless the sampling program for joint surveys is constrained such that joints longer than a fixed length l_c are not measured even if they in fact could be, the problem of censoring becomes more difficult. In particular, the point of censoring is itself a random variable. The observations recorded are (1) a set of completely observable traces, $\underline{l}_x = \{l_{x,1}, \ldots, l_{x,r}\}$; and (2) a set of traces for which only one or neither end is observable, $\underline{l}_z = \{l_{z,1}, \ldots, l_{z,t}\}$.

The likelihood of $(\underline{l}_x, \underline{l}_z)$ is,

$$L(\underline{l}_x, \underline{l}_z | \underline{\Theta}) = \prod_{i=1}^{r} f(l_{x,i} | \Theta) \cdot \prod_{j=1}^{t} \int_{l_{z,j}}^{\infty} f(l_{z,j} | \underline{\Theta})\, dl \tag{3.2}$$

in which $\underline{\Theta}$ = the parameters of the trace length *pdf* (corrected for other biases). The second term in the right hand side is the probability that a censored trace would be longer than that observed. Clearly, closed form maximization of Eq. (3.2) with respect to Θ is only possible for *pdf*'s having analytical cumulative distributions (*cdf*). Therefore, while analytical results for censored Exponential sampling are available, only numerical solutions are available for Normal and logNormal sampling. For Exponential sampling the *ML* estimator of λ is

$$\hat{\lambda}_{\mathrm{ML}} = \frac{r}{\Sigma l_{x,i} + \Sigma l_{z,j}} \tag{3.3}$$

and the posterior *pdf* on λ in a Bayesian sense and starting from a non-informative prior is Gamma (Baecher, 1980). The sampling variance of $\hat{\lambda}_{LM}$ in the Exponential case is,

$$V[\hat{\lambda}_{\mathrm{ML}}] \cong \frac{\hat{\lambda}^2}{n} \tag{3.4}$$

and for other parent *pdf*'s $V[\hat{\Theta}_{ML}]$ can be found by numerical approximation.

— Truncation Bias

In collecting joint data a decision is usually made not to record traces shorter than some cut-off length. This decision is made either out of expediency or because short traces are difficult to distinguish, as for example in photographs. Several workers have noted that this form of truncation introduces bias into the sampling plan, increasing the sample mean (Baecher et al., 1978, Cruden, 1977, Priest and Hudson, 1981).

Fig. 9 shows the bias in the sample mean resulting from truncating at a given fraction of the mean trace length, for an Exponential *pdf* of trace length. This bias is smaller for distributions, like the logNormal, with zero density at the origin. The figure clearly shows that the effect of truncation bias on estimates of central tendency of the trace length *pdf* is small, unless the chosen truncation level is large (e. g., >10% of the mean). For most purposes this bias can be safely ignored.

Errors and Biases in Sampling for Joint Orientation

— Sampling Error: Weighted Sampling Plans

In joint surveys, differences in the probability of being sampled are caused by geometric relationships (e. g., relative orientations of joints and outcrops), and non-geometric relationships (e. g., difference in the degree of joint weathering). Only geometric biases are considered here. Non-geometric differences tend more to be questions of geology alone, and not statistics.

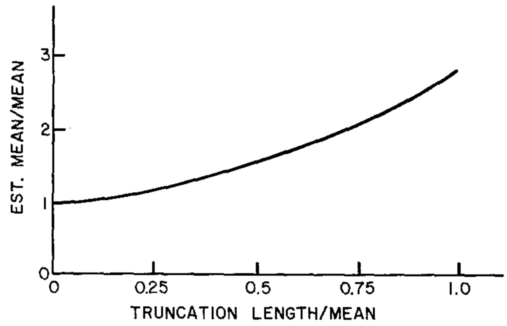


Fig. 9. Bias in sample mean due to truncation of short trace lengths

Geometric relations which cause joints to have low probabilities of being sampled were brought to the attention of the literature by R. Terzaghi (1965), although Sander et al. (1954) and others had earlier considered related problems with thin sections.

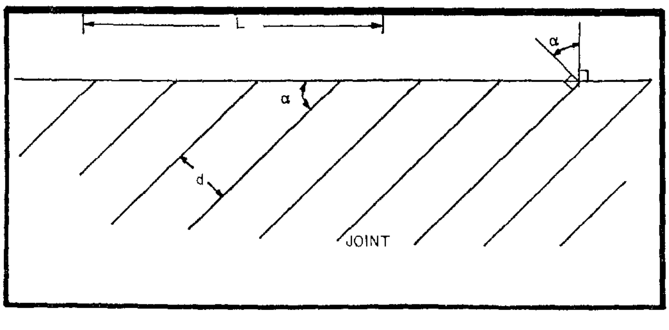


Fig. 10. Probability of joints intersecting an outcrop

To be sampled a joint must be a member of the sampled population. It must intersect an outcrop, a boring, or an excavation from which samples might be drawn. Joints which do not, cannot be measured. Assuming that the sampling plan by which the joints are sampled from outcrops and borings is self-weighting (in the sense that bias on the outcrop is corrected), the probability of an orientation entering the sample is proportional to the probability of it intersecting an outcrop or boring. Consider the two-dimensional situation of Fig. 10. The probability of a joint of a given orientation intersecting the ground surface in an interval ΔL is

$$P_{\Delta L} = \frac{\Delta L \sin \alpha}{d}. \tag{3.5}$$

In words, given d, joints that are flatter with respect to the surface appear less frequently in the sample than joints that are steeper. In order to be a probability sample this difference must be accounted for by weighting. Since the ratio of probability must be constant, if w_α is the weighting factor for joints at angle α,

$$w_\alpha \propto 1/\sin \alpha. \tag{3.6}$$

The use of computers allows consideration of more precise methods: methods in which the sample is considered as a whole and in which closer approximations can be made. This can be done by evaluating the probability of a given orientation appearing in the total sample (i. e., in any of the outcrops or borings) and applying weights accordingly. The probability of a joint of a given orientation occurring in outcrops i is proportional to $B_i \sin \alpha_i$, where B_i is some dimension of the outcrop and α_i is the angle the orientation makes with the normal to the plane that best models the outcrop. Since a single joint may be sampled in more than one outcrop or boring (i. e., sampling "with replacement") the probability of the orientation being measured in the entire set of outcrops sampled is proportional to

$$\Pr \begin{pmatrix}\text{given orientation being} \\ \text{seen in the entire sample} \\ \text{of outcrops}\end{pmatrix} \propto \Sigma B_i \sin \alpha_i. \tag{3.7}$$

Similar consideration for boreholes leads to

$$\Pr \begin{pmatrix}\text{given orientation being} \\ \text{seen in the entire sample} \\ \text{of borings}\end{pmatrix} \propto \Sigma L_j \cos \beta_j \tag{3.8}$$

where L_j is the length of the j^{th} boring and β_j is the angle the joint makes with the j^{th} bore hole axis. The probability of a given orientation appearing in the entire sample is

$$\Pr \begin{pmatrix}\text{given orientation} \\ \text{in entire sample}\end{pmatrix} \propto \Sigma B_i \sin \alpha_i + \Sigma L_j \cos \beta_j \tag{3.9}$$

and the weighting factor, being proportional to the reciprocal of the probability, is,

$$W(\text{given orientation } Z) \propto \frac{1}{\Sigma B_i \sin \alpha_i + \Sigma L_j \cos \beta_j}. \tag{3.10}$$

— Estimation Error

Two approaches to estimation errors in joint surveys can be taken: analytical and empirical. The analytical approach is based simply on the sampling variance of statistical estimators for the various analytical forms or non-parametric descriptors. For example, for a spherical root mean square variation of about 10^0, a sample size of $n=100$ leads to a standard error on the spherical mean of about 1^0. $n=200$ reduced this to 45′.

The purely empirical approach is based on observed changes in pole diagrams as sample sizes increase (e. g., Larsson, 1952). Some results of simulations in which data sets of size $n=25$, 50, 75, 100, 125 and 150 were

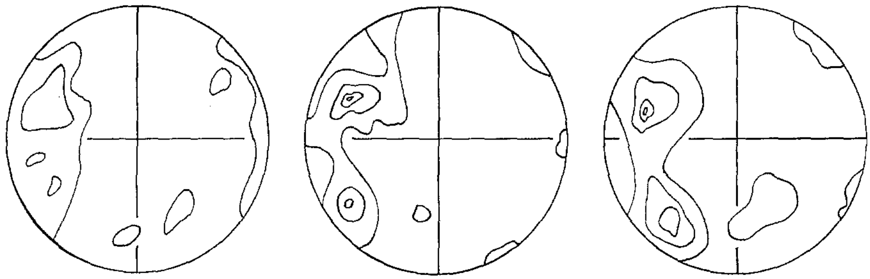


Fig. 11. Typical changes in pole diagram as number of poles sampled increases, poles randomly sampled from population of size 725

randomly sampled from surveys of much larger size ($N=725$) are shown in Fig. 11. Apparently, sample sizes of about $n=100$ yield acceptably precise orientation diagrams.

— Measurement Error

Measurement errors are caused by inaccuracies in either the instruments of measurement or the reading of instruments. They are of two types, random and systematic. Random errors are unpredictable both in magnitude and direction. The treatment of random errors usually assumes magnitude to be Normally distributed with mean zero. This allows confidence limits to be placed on the measurements themselves. Random errors are usually much smaller than sampling or estimation errors, and can often be neglected,

Although random errors in joint orientation measurements arise from a host of sources, several general comments can be made about them. Random error in the strike direction is greater for "flat" dipping than "steep" dipping joints. The sensitivity of the direction of the line of intersection of two planes to error in the orientation of one or both is a function of the angle. The smaller the angle, the more sensitive. Since flatter joints form a smaller angle with the horizontal, they are more sensitive to errors in leveling the geologic compass.

Random error in the dip is greater for steeper joints than for flatter joints. This comes primarily from two sources other than reading and round-off errors: inaccuracy in leveling the pendulum inclinometer, and inaccuracy in aligning the geologic compass parallel to the dip direction. While the first is independent of the dip, the second is not. The scale of roughness to compass size also contributes to random measurement error.

While attempts could be made to determine measurement error analytically, the simplest and most reliable way is to simply perform several measurements on a single joint, and empirically determine the dispersion of values. This was done in the laboratory, using two fixed planes and having several people measure the strikes and dips with a geologic compass. The

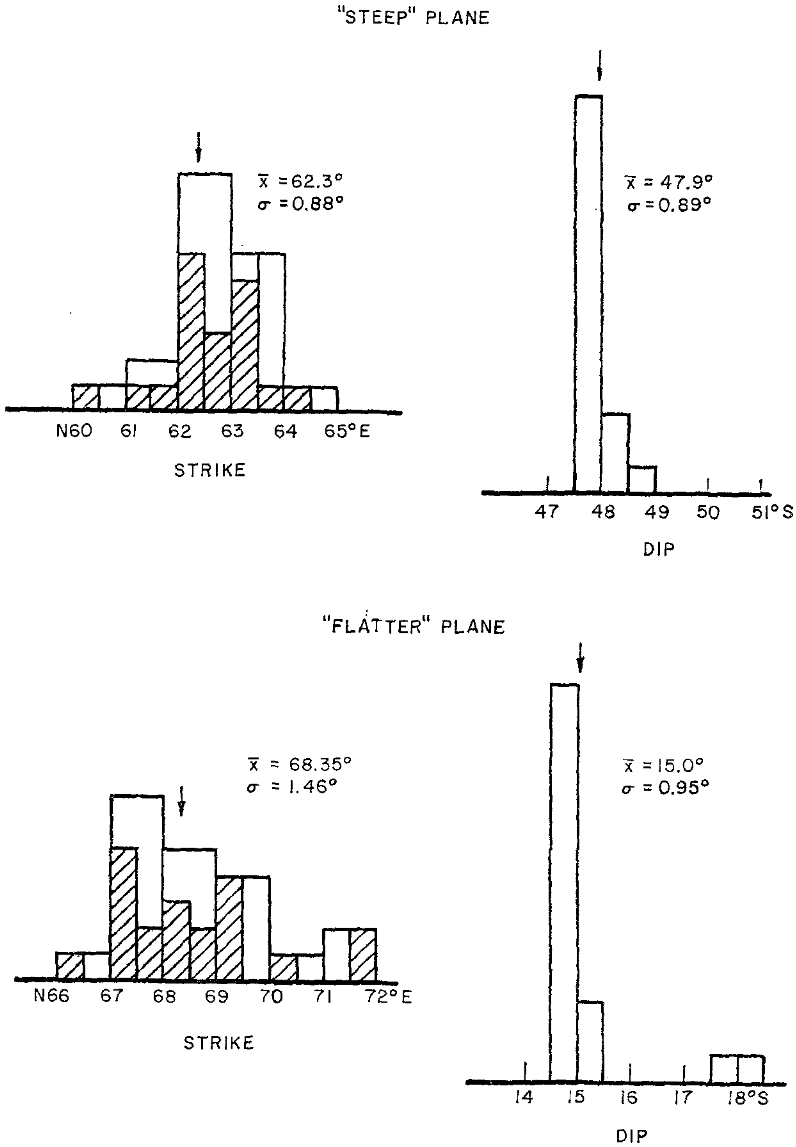


Fig. 12. Effect of joint inclination on measurement error

results are only intended to be illustrative. The variance in the dip measurements was about the same for both planes but the variance in the strike measurements was almost twice as great for the flat plane as for the steep plane (Fig. 12).

Systematic errors are errors whose mean value is different from zero. Measurement values of a quantity will, therefore, be almost consistently high or consistently low. Systematic errors, like random errors, are inevitable; however, unlike random errors, they are not reduced by large sample sizes. The only strategy against systematic errors is to hold them to a "reasonable" level. One can never know whether systematic errors have been sufficiently reduced. One can only carefully consider possible sources of error, and search for inconsistencies in sample data.

Sampling Plans

The purpose of statistical sampling is that it allows estimates of rock mass properties to be made that are optimal in some agreed upon sense, and for which estimate precisions can be determined. In order to do this the sampling plan must be representative in the sense that (1) every element of the sampled population have a non-zero probability of appearing in the sample, (2) the relative probability of each element appearing is known, and (3) the importance given to observing a particular element be in inverse proportion to its probability of appearing in the sample.

In sampling, three populations are of interest. The *target population* is that collection of elements about which information is desired. For joint surveys this might be the population of joints at some depth in a rock mass. The *sampled population* is that collection of elements that are available for sampling. For joints surveys this might be the population of joints intersecting outcrops, borings or excavations. The *sample* is that collection of elements whose properties are actually measured. This might be the joints whose traces intersect sampling lines or fields.

Statistical procedures allow quantitative inferences to be made about properties of the sampled population from observations on the sample. They do not, however, allow formal inferences to be made about properties of the target population. Such inferences are based on geology; they have little to do with statistics.

Typical Sampling Plans for Joint Surveys

Sampling plans for joint surveys must meet two criteria:

1. They must allow valid statistical inferences to be drawn, whose precision can be evaluated (i. e., probability sampling).
2. They must be economical and easy to implement.

In most cases the cost of analysis is much less than the cost of field data collection, so plans which minimize the sampling effort are to be favored.

Simple random sampling of joints is almost always infeasible. These plans require randomly selecting individual joints around the site and measuring their orientations. For the same reason, stratified random sampling is infeasible unless strata are small (e. g., the size of outcrops). Stratification is not an innate property of populations in general; if a population is stratified into internally homogeneous subpopulations, the strata sizes are a property of the population and not just the sampling plan. Joint populations are naturally stratified into joint sets, and joint populations are frequently stratified into geographical or lithological subpopulations. Prior stratification by these properties will improve the performance of any sampling plan.

Systematic plans for sampling joints are easier to use than simple or stratified random ones because joints to be measured are easily located. A plan that specifies every 100th joint, say, is infeasible, but a plan that specifies "joints within a 6′ circle every 50 feet", say, is not. Problems of periodicity in the sampled population might be encountered if stratification by lithology does not precede systematic sampling.

Cluster sampling plans have long been favored for joint surveys because the time required to sample several joints at one outcrop is less than the travel time between outcrops. In cluster plans several outcrops are selected by some random process, and from each selected outcrop a sample is taken. Many sampling plans based on clustering could be suggested. The following is outlined only as an example.

Different geological formations at a site frequently have different patterns of jointing due to differences in rigidity, friction angle, age of jointing, and the like. Therefore, the initial step is to stratify the site by major formations. Since one does not know *a priori* whether the populations of joints are homogeneous from formation to formation, these data sets are maintained separately.

Next, each formation is arbitrarily stratified by superimposing a large regular grid, and the data from each quadrate kept separately. The dimensions of this grid might be on the order of 1000 feet depending on site dimensions and available effort. This stratification allows a "nested analysis" of variance and variances in joint population properties to be obtained as a function of spatial dimensions. This information is desired because the variance of joint properties generally increases as the volume of rock considered increases. For a structure only affecting a small volume of rock, estimates made from the total joint population overestimate true local variance, and perhaps either overestimate or underestimate the local mean. The formation is not stratified to improve the overall estimate of population parameters (the usual reasons for stratification) since each stratum is treated identically, but simply to maintain separation of the data sets.

Within each stratum clusters of joints are selected for sampling. If few outcrops exist, all of them are sampled; if not, a random process must be used for selection. The orientation and size of all selected outcrops must be recorded.

Several joints intersect each outcrop to be sampled. If their number is large, not all of them may be measured and a second-stage sampling plan is required.

Two second-stage sampling plans which should be avoided, even though they are frequently used, are systematic plans (e. g., sampling every tenth joint along a line) and plans randomly locating points on the outcrop and measuring the closest joint. Systematic plans should be avoided because periodicities are likely to exist in the way joints intersect an outcrop, while sampling by measuring the closest joint to a random point should be avoided because joints whose individual spacing is large have a higher probability of being sampled than ones whose individual spacing is small.

Snow (1966) has suggested sampling a single outcrop by randomly locating a line segment and measuring every joint which intersects it. This is a satisfactory method in that it is random and does not allow personal bias in selection, no matter how tight, small, or hard a joint is to measure. However, this procedure leads to large weighting factors and a "blind zone" for those joints whose pole is perpendicular to the sampling line. These large weighting factors and the blind zone can be reduced by using two perpendicular line segments. A better alternative is to place a sampling line on the outcrop and measure every joint intersecting a rectangular "window" of fixed width centered on the sampling line.

III.2 Subjective Assessment of Uncertainty

Much of the uncertainty in geological exploration can only be expressed subjectively. Because of this, fairly well-developed methods of subjective probability theory (see, e. g., Barnett, 1973, for a discussion of philosophy)

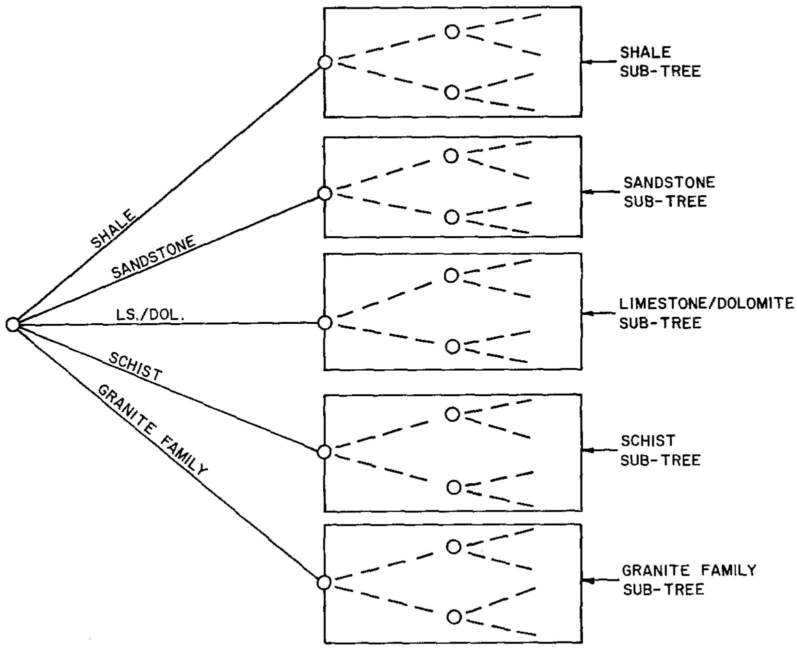


Fig. 13. Parameter tree and "subtrees"

can sometimes be used to help quantify these types of uncertainties and to help rationalize the way they are treated.

Subjective assessment of uncertainty in engineering geology is best illustrated with the geologic submodel of the so-called Tunnel Cost Model (TCM)

(Moavenzadeh, 1974, Einstein et al., 1977). The TCM will later be used in discussing uncertainty in design and construction; several successful practical applications have been made.

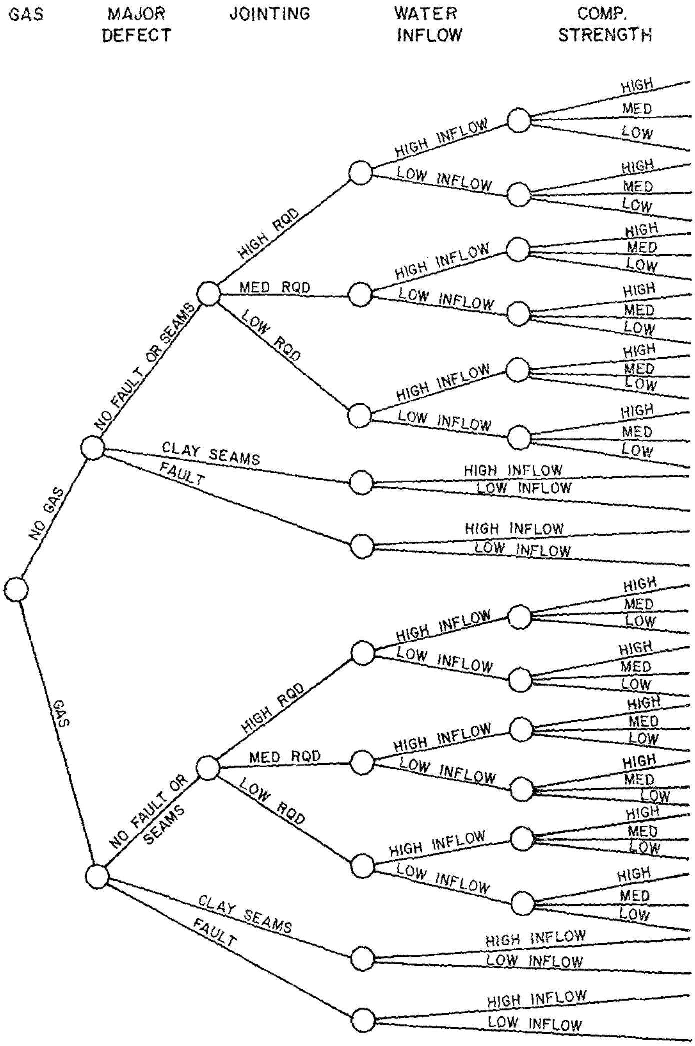


Fig. 14. The sandstone parameter tree

Within the TCM-geologic submodel:

— Geology is summarized by geotechnical parameters which can be related to design and construction. Table 4 lists one possible set of parameters and states.

— Geology is described by a combination of parameter states, all possible combinations of which can be developed in a parameter tree (Figs. 13, 14).

— The parameter tree is used to assign uncertainties. Beginning with Fig. 14, subjective estimates are made at each node of the parameter states which may occur. Such subjective estimation is usual in geology, although typically summarized in verbal description (e. g., "likely condition", "improbable", "predicted with great confidence", etc.). The only difference here is to express these probabilities quantitatively. Fig. 15 shows an example where one estimates at a particular location that shale is less likely to occur than granite, where high *RQD* is expected with a 50% chance, and so on. These numbers can be directly estimated by intuition or through formal procedures (e. g., Staël von Holstein, 1974). Using the parameter

Table 4. *Geotechnical Parameters and Parameter States*

Parameter	Parameter states	Major construction consequences
Rock type	Shale Sandstone Limestone/Dolomite Schist Granite, Basalt Diabase, Intrusive Basalt, Gneis, Quartzite	Wear of cutters or drill bits (Indication of other rock types)
Jointing RQD	High Medium RQD Low	Support requirements, rate of advance, overbreak
Major defects	Fault of Shear Zone Clay Seams	Support requirements, rate of advance, overbreak
Foliation	Highly foliated Non-foliated	Support requirements, overbreak, boring machine rate of advance
Gas	Gas exists No gas exists	Delays, ventilation requirements
Water inflow	High water inflow Low water inflow	Remedial measures like grouting
Compressive strength	Very high High Medium Low	Boring machine rates of advance, supports

tree one obtains probabilities of combinations of parameters. In Fig. 20 the probabilities are assumed independent, but the procedure can accommodate dependence (Einstein et al., 1974).

— The combinations of parameter states and their probabilities are not the same throughout the tunnel. Therefore, the profile is segmented (Fig. 16) whenever parameter combinations or associated probabilities change. It is, of course, possible to include dependence between segments if the geology is related. Segmenting a tunnel by geologic conditions and estimating their possible occurrence is the standard procedure for tunnel geologists. Except for increased quantification, the procedure here is identical.

The Roberts Tunnel in Colorado provides a good example of this procedure. Specifically, consider the sections where the tunnel intersects the William's Range Thrust Fault (Fig. 17). The procedure starts with the con-

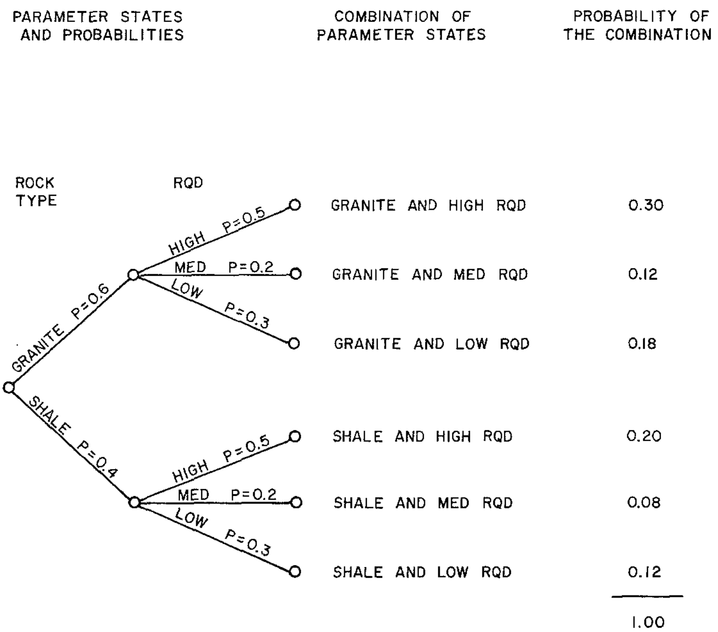


Fig. 15. Parameter tree and probabilities for independent parameters

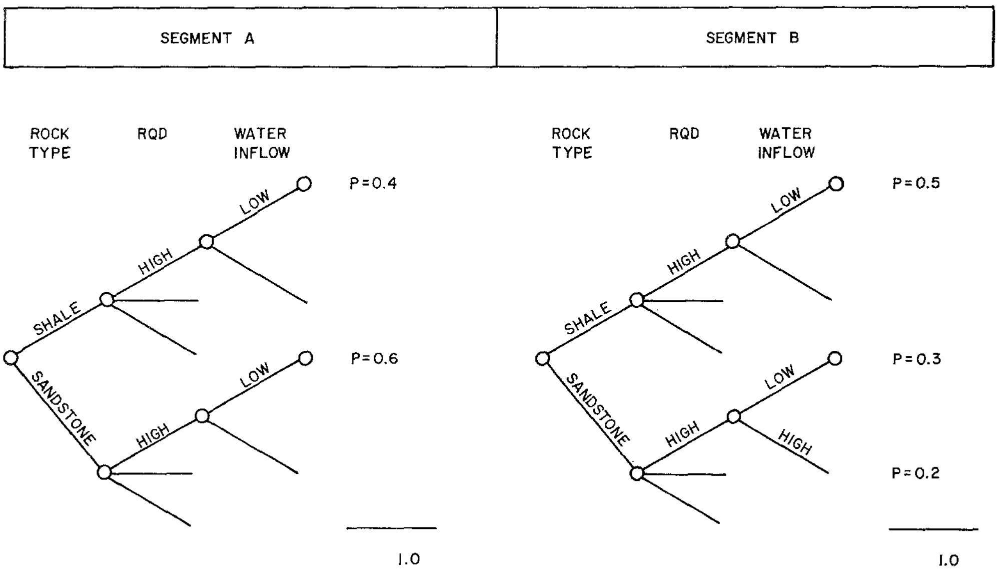


Fig. 16. Segments with different combinations of parameter states and associated probabilities

struction of a normal geologic profile using information from outcrops and borings. The William's Range Thrust Fault is a bowl shaped feature; on the outside is either Pierre shale or a baked shale, inside are metamorphic or igneous rocks. There is also evidence of a heavily sheared zone and of sound rock (boring DDH6). The profile is next scrutinized and marked where the geologist is uncertain. For example, markers 7 and 10 indicate uncertainty on the intersection of the fault with the tunnel axis (and thus of the fault width), 8 refers to the width of the sheared zone, and 9 represents the uncertainty about the width of the sound rock zone. The next step is segmen-

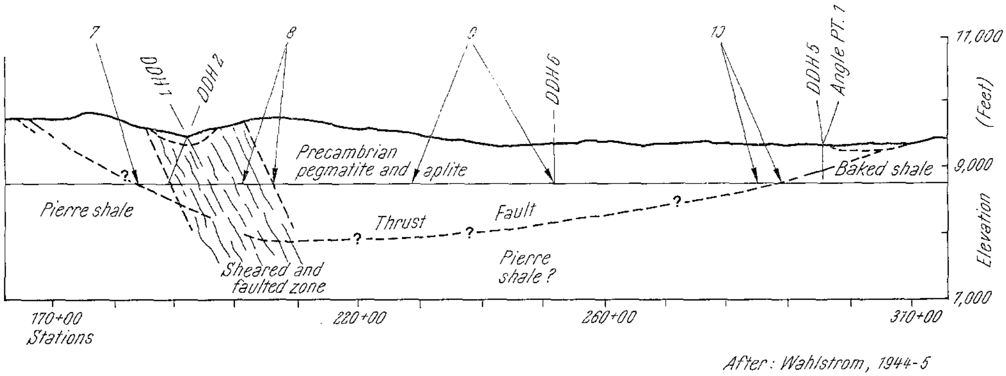


Fig. 17. Geologic profile and uncertainties Williams range thrust fault, Roberts tunnel Colorado

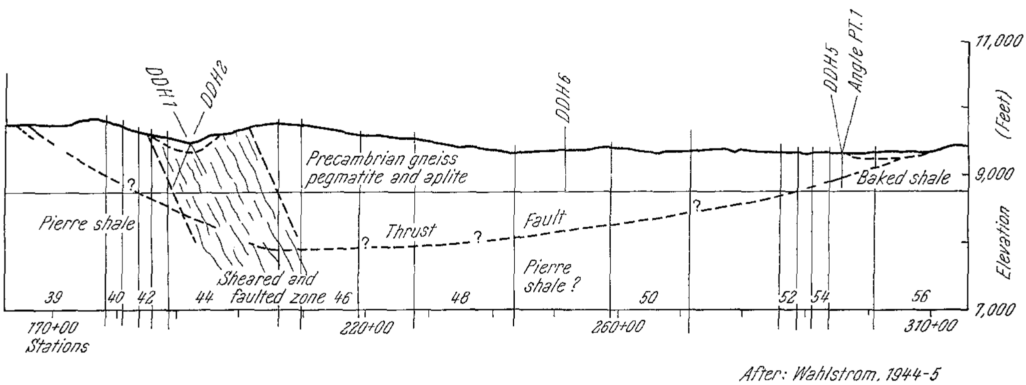


Fig. 18. Segmentation Williams range thrust fault, Roberts tunnel Colorado

tation (Fig. 18) and creation of parameter trees (Figs. 19 and 20). A geologic unit as used in this context is a particular rock type, in combination with specific other parameter states and their associated probabilities. Figs. 19 and 20 show examples for the relatively sound units, Granite 1 and 2, and the heavily sheared units, Granite 5 and 6. In the segmented profile of Fig. 18, worst conditions (i. e., Granite 6) are assigned to segment 44 with a probability of 100%. The adjacent segment is evidently a transition to better granite. Dependencies among segments are used at fault intersections. These descriptions of geology are the input of the geologic submodel pro-

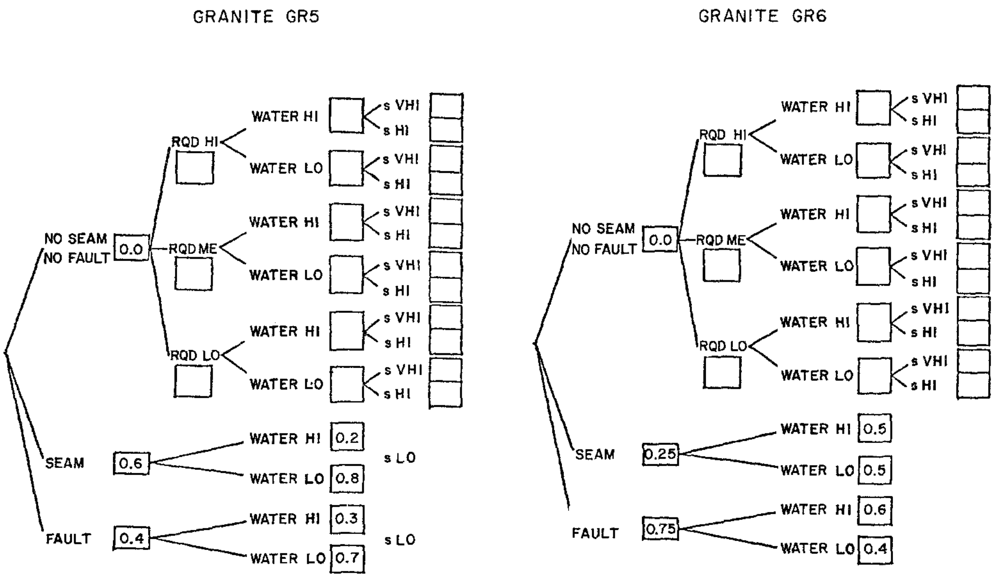


Fig. 19. Granite parameters trees

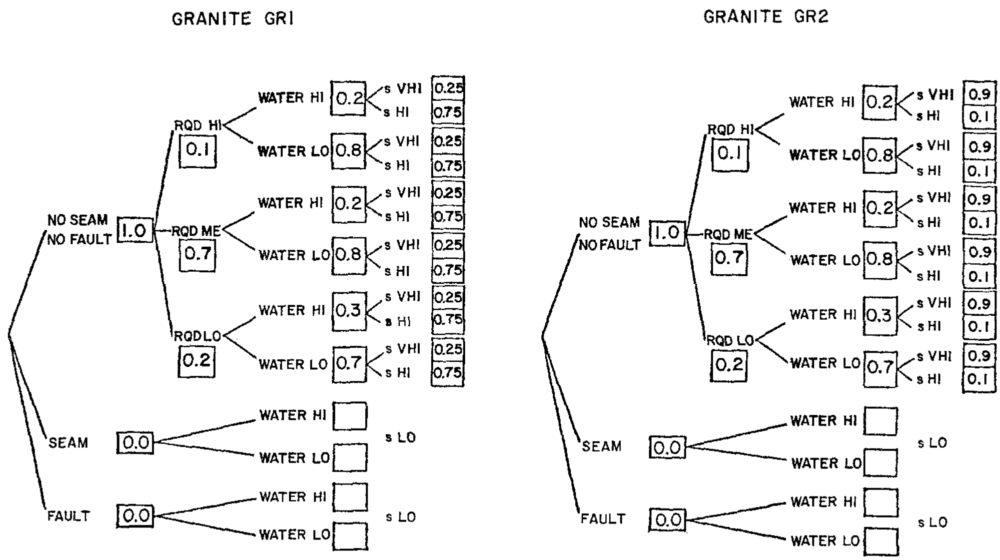


Fig. 20. Granite parameters trees

gram, which stores this information and supplies it to the tunnel simulator program. The latter then simulates geologic profiles like those in Fig. 21.

The three example profiles shown in this figure represent favorable, medium and unfavorable conditions. The frequency with which a certain unit is simulated in a particular segment corresponds to the input probabilities.

Later, the Tunnel Cost Model simulates tunnel construction through each profile resulting in a time-cost pair for each profile and time-cost distributions for a complete run.

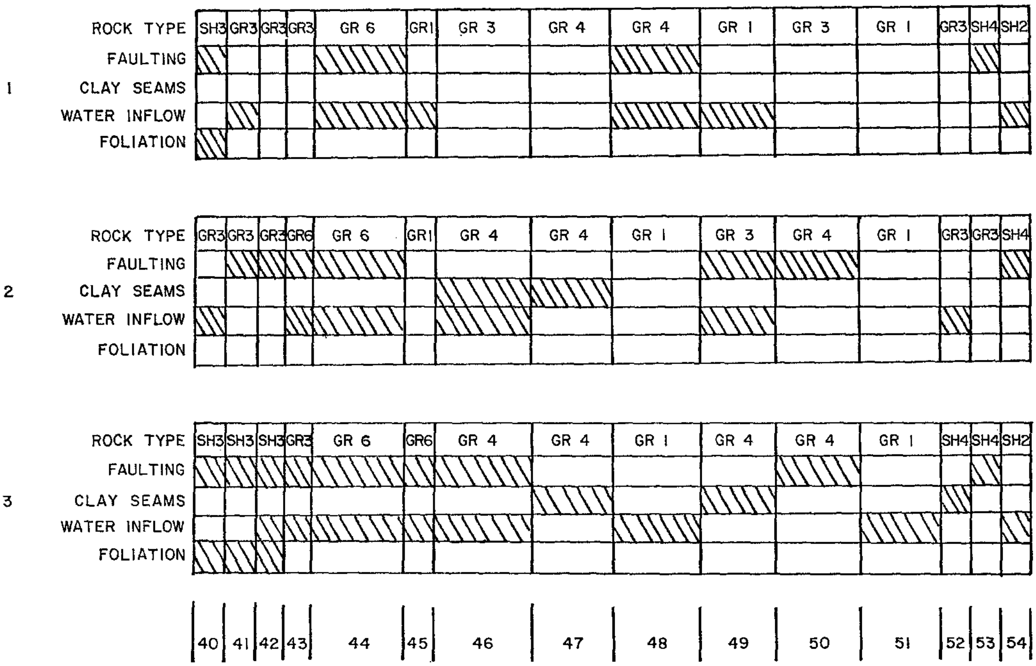


Fig. 21. Three simulated geologic profiles for Williams range thrust fault (dark bands: indicate unfavorable parameter states)

The TCM has been successfully used on several tunnel projects. However, detailing geologic descriptions, incorporating objective information (e. g., from boreholes) and estimating location specific probabilities are often difficult. Thus an improved model for estimation, the Geologic Prediction and Updating Model, has also been developed. This model is easier to use, but makes somewhat stronger stochastic assumptions (see, Ashley et al., 1981).

III.3 Exploration Planning

Sections III.1 and III.2 have shown that it is possible to collect information statistically and to express uncertainties quantitatively. This is the basis for formal exploration planning in which technical and economic aspects are balanced.

Exploration reliability and the marginal benefit of detailed information are major considerations in Exploration Planning. Here, an application to underground construction is used to illustrate a formal planning procedure. The method has also been applied to other types of geotechnical projects (Baecher, 1978). Fig. 22 shows a generic decision-analysis cycle leading to a decision on collecting more information; Table 5 describes the steps in detail.

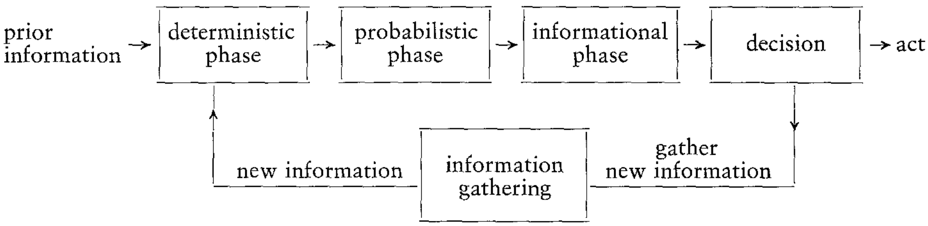
Table 5. *Decision Analysis for Tunnel Exploration*

Steps in the analysis of the exploration decision	
Deterministic	In exploration for tunnels
1) Define the decision problem ..	Determine if exploration is beneficial; if it is, what is the optimal exploration program.
2) Identify the alternatives	Exploration No exploration
3) Outcomes	Value of information; Construction costs and exploration sites are intermediate results
4) Decision and state variables ...	Exploration cost; Exploration reliability; Construction method costs geology
5) Relationships between variables and outcomes	Effect of geology and exploration on expected construction cost; Decision tree
6) Value (of outcomes)	Expected cost
Probabilistic	
1) Encode uncertainty in state variables	Prior probabilities of geologic states
2) Probabilistic model	Effect of geology and exploration on expected construction cost; Decision tree
3) Choose among distributions ..	e. g., mean of cost-time scattergrams
4) Probabilistic sensitivity analysis	Critical ranges of probabilities, exploration reliability, and construction costs established
Information	
1) Value of perfect information ..	Probabilistic sensitivity analysis
2) Best information gathering scheme	Exploration method and configuration (geometry along tunnel)

In the *deterministic phase* one enumerates a course of action and outcomes they may lead to (1—3). Outcomes depend both on decision variables that can be controlled and state variables that cannot (4). Variables and outcomes are related to each other in a model (5). Outcomes are compared with one another through an objective function usually involving cost or time criteria, but which can also include technical performance criteria (6, 7). Sensitivity studies (9) make it possible to identify most important variables.

The *probabilistic phase* is basically a "revision" of the deterministic phase introducing uncertainty of the state variables. Outcomes in the probabilistic phase are in the form of probability distributions.

A decision could be made at this point, but generally one wants to establish whether further information might be beneficial. This is analyzed in the *information phase*. If further information is expected to be beneficial, this also results in the optimal level of information to be gathered. *Preposterior analysis* is a key element in the information phase. The consequences



Deterministic Phase

1. Define problem and limits of investigation
2. Alternative course of action
3. Outcomes of each alternative
4. Select decision and state variables
5. Relate outcomes and variables
6. Method of comparing relative values of each outcome
7. Time preference
8. Dominated alternatives eliminated
9. Sensitivity of outcome to variables

Probabilistic Phase

1. Express uncertainty in variables by means of probabilities
2. Probabilistic model
3. Establish relative value of probabilistic outcomes
4. Probabilistic sensitivity analysis

Information Phase

1. Value of perfect information
2. Evaluate various information collection schemes

Fig. 22. The decision analysis cycle
(after Staël von Holstein, 1973)

of potential future actions (collecting new information) are assessed before the action is taken. Preposterior analysis makes use of Bayes' theorem to perform updating,

$$P\,[B_j/A] = \frac{P\,[A|B_j]\,P\,[B_j]}{\sum\limits_{j=1}^{n} P\,[A|B_j]\,P\,[B_j]} \qquad (3.11)$$

Or in words:

$$\begin{bmatrix}\text{Posterior probability}\\ \text{of } B_j \text{ given the new}\\ \text{information, } A\end{bmatrix} = \begin{bmatrix}\text{Likelihood of the}\\ \text{new information,}\\ A\text{, given } B_j\end{bmatrix} \times \begin{bmatrix}\text{Prior}\\ \text{probability}\\ \text{of } B_j\end{bmatrix} \times \begin{bmatrix}\text{Normalizing}\\ \text{factor}\\ \end{bmatrix}$$

The likelihood function $P\,[A|B_j]$ is the probability that the particular observation A (or "data A") would be made in exploration, given that the true state of nature was B_j. In applying the procedure of Fig. 22 to tunnel exploration, the phases and steps of Table 5 result.

The *alternatives* "exploration" or "no exploration" (step 2) can be defined for any stage of a geotechnical study. The "exploration" alternative includes many exploration strategies which are combinations of methods, locations and number of explorations. An exploration method is characterized by its cost and its reliability. The reliability is the probability that the exploration results indicate the true conditions and is represented by the likelihood function in Bayes' theorem.

The *value of information* (step 3) is the difference between expected construction cost without exploration and expected construction cost with the particular exploration alternative. The goal of exploration usually is to reduce the expected construction cost. However, exploration involves some cost also. The objective of the decision analysis is thus to minimize exploration cost plus expected construction cost, or in other words, to establish the maximum value that one is willing to pay for exploration.

The *decision variables* are exploration methods which can be described by exploration cost and reliability, and construction methods which are described by their costs.

The *state variables* are the geologic conditions affecting tunnel construction such as jointing, water inflow, major defects. In this paper a simplified description of geologic conditions with the three states, good, fair and poor, is used.

The establishment of *relations* between variables and outcomes is the major problem that needs to be solved in any analysis, and is discussed later. At this point, it may suffice to say that geologic conditions, exploration and construction costs have to be related to each other. The value of an outcome in the present case is measured only by expected cost of construction.

In the *probabilistic phase*, degree of belief probabilities are assigned to the geologic states (Einstein and Vick, 1974). The probabilistic model makes use of the deterministic relations and, by introducing uncertainties in the form of subjective probabilities, produces distributions of outcomes rather than point estimates. One means for relating geology to construction cost in a probabilistic manner is the Tunnel Cost Model.

A *sensitivity analysis* is performed to evaluate changes in the "best" decisions and changes in predicted costs that result from variation in the input parameters. In this way, the "sensitivity" of optimal exploration strategies to probabilities of geologic conditions and estimates of construction sequences can be determined. If the optimal strategies are insensitive to minor fluctuations in the variables, one can say that the decisions are "robust" and one has more confidence in them.

In the *information phase*, the expected value of perfect information (EVPI) is calculated to eliminate those tunnel sections for which even perfect information (e. g., knowing the true geologic conditions precisely) would not be cost effective. Then, in the remaining sections, the expected value of alternative imperfect exploration schemes is evaluated.

This then is the general procedure for exploration decision analysis. More specific details are described below.

Details of the Exploration Decision Analysis

Three facts of the decision analysis approach need further attention, the relationships among input variables and outcome, the comparison among outcomes, and the sensitivity analyses. These are discussed through an example.

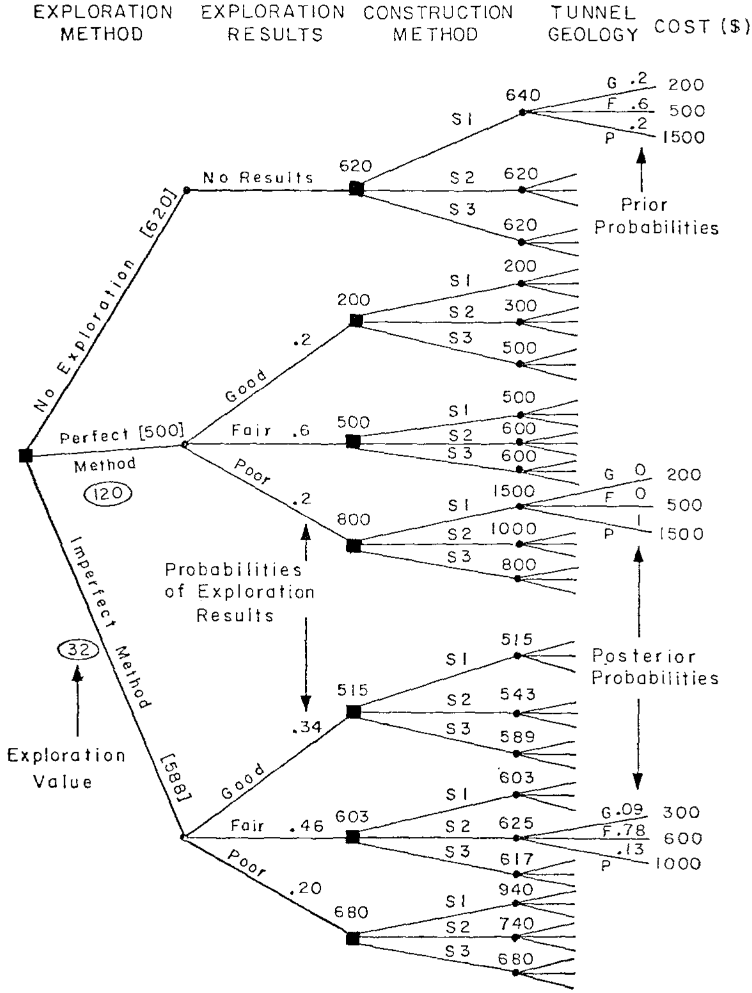


Fig. 23. Exploration decision tree

The tool for relating variables and outcomes and for comparing different outcomes is the decision tree (Fig. 23). Using this tree, the expected cost of the "no exploration" case is computed by: (1) multiplying the cost of any of the construction strategies in a particular geology by the originally estimated subjective probability of that geology; (2) summing these "expected costs" for each construction strategy (e. g., S1 = 640); and (3) selecting the construction strategy with minimum expected cost (S3 or S2 = 620).

The expected cost for the "exploration" cases are computed similarly by: (1) calculating the posterior probability of each state conditioned on each possible result of exploration (e. g., if the exploration program indicates "fair" conditions, the probabilities of "poor", "fair", and "good" conditions, respectively, might be 0.09, 0.78 and 0.13); (2) determining the expected cost (i. e., probability times cost of any of the construction strategies in the particular geology) for each exploration result (analogous to step 1 above: 300×0.09, and so on, and these summed); (3) selecting the minimum-expected-cost construction strategy for each exploration result; (4) finally, weighing each minimum cost strategy by the probability of the corresponding exploration result and summing ($515 \times 0.34 + 603 \times 0.46 + 680 \times 0.20 = 588$). Adding the cost of the particular type of exploration to this sum yields the expected total cost of that exploration strategy (588 plus cost of exploration). The expected value of exploration (expected value of sample information, EVSI) is the difference between the expected cost of the best strategy (step 4) and the expected cost of the "no-exploration" case (i. e., 32 minus cost of exploration). (The expected value of perfect information is obtained similarly as shown in Fig. 23). The analysis expressing these relations must be performed for each location of possible exploration.

Table 6. *Reliability Matrix*

		Exploration result		
		E_G	E_F	E_P
Geologic states	G	0.6	0.2	0.2
	F	0.3	0.6	0.1
	P	0.2	0.3	0.5

To calculate EVSI, both exploration reliability and construction costs must be known. *Exploration reliability* is expressed in the form of a reliability matrix, a matrix of likelihood functions as shown in Table 6. The likelihoods or reliabilities are the result of subjective assessment of the performance of an exploration method in a certain geologic condition (e. g., the method in Table 6 has an 0.5 reliability of indicating "poor" conditions if the real conditions are "poor"). Exploration reliability includes "mechanical" uncertainties of the exploration method (e. g., breaking of intact rock due to drilling) and uncertainties in inferences. Reliabilities apply to entire segments. Therefore the reliability of a single boring may decrease as the segment increases. Reliabilities vary between completely reliable (identity matrix) and completely ambiguous (all probabilities equal $1/n$, where n = number of geologic states).

The *construction costs* of a certain construction strategy in certain geologic conditions can also be represented in matrix form (Table 7). Normalizing such that the strategies are ideal for the geologies along the diagonal, one can form a so-called penalty matrix with 0-penalties along the diagonal

(Table 8). The numerical example in Table 8 is corrected to avoid negative numbers. If a strategy is the best for all geologic conditions, a row of 0-penalties would occur; for the numerical example, strategy S1 is the best for "good" and "fair" geology, strategy S3 for "poor" geology. Reliability and penalty numbers are then used to compute EVSI as shown in the decision tree.

Table 7. *Construction Cost Matrix*

		Geology G	F	P	e. g.	G	F	P
Construction strategy	S1	C_{1G}	C_{1F}	C_{1P}	S1	200	500	1500
	S2	C_{2G}	C_{2F}	C_{2P}	S2	300	600	1000
	S3	C_{3G}	C_{3F}	C_{3P}	S3	500	600	800

Table 8. *Penalty Matrix*

	G	F	P		G	F	P	e. g., from Table 7	G	F	P
S1	$C_{1G}-C_{1G}$	$C_{1F}-C_{2F}$	$C_{1P}-C_{3P}$		0	P_{1F}	P_{1P}	S1	0	0	700
S2	$C_{2G}-C_{1G}$	$C_{2F}-C_{2F}$	$C_{2P}-C_{3P}$	⇒	P_{2G}	0	P_{2P}	S2	100	100	200
S3	$C_{3G}-C_{1G}$	$C_{3F}-C_{2F}$	$C_{3P}-C_{3P}$		P_{3G}	P_{3F}	0	S3	300	100	0

With this procedure it is thus possible to eliminate sections for which further exploration would not be cost effective and then to rank the segments in which exploration is beneficial (Einstein et al., 1978).

An interesting extension and application of the exploration analysis for underground structures is the exploration analysis method for underground

Table 9. *Estimated Costs of Construction*

	Caverns size	Capacity	Cavern volume	Geology: Excellent MM$	Good MM$	Fair MM$	Poor MM$	Very poor MM$
build 25′ a_1 @ 2000′	42 caverns 2420′ × 2370′	2000 MMCF	66.806 MMCF	87.4	131.1	174.8	218.5	262.2
build 55 a_2 @ 2000′	15 caverns 1530′ × 1500′	2000 MMCF	66.806 MMCF	62.64	93.36	156.6	219.24	281.88
build 115′ a_3 @ 2000′	11 caverns 1010′ × 2000′	2000 MMCF	66.805 MMCF	53.66	107.32	187.81	241.47	321.96
build 25′ a_4 @ 5000′	19 caverns 1030′ × 990′	2000 MMCF	12.464 MMCF	49.11	98.22	147.33	196.44	245.55
build 55′ a_5 @ 5000′	7 caverns 600′ × 540′	2000 MMCF	12.464 MMCF	39.89	79.78	139.62	219.40	279.23
build 115′ a_6 @ 5000′	5 caverns 535′ × 300′	2000 MMCF	12.464 MMCF	36.53	91.33	164.39	237.45	288.77

gas storage caverns (Einstein et al., 1977). The hypothetical example involves a storage facility with a total capacity of 2000 MMCF consisting of a number of individual caverns.

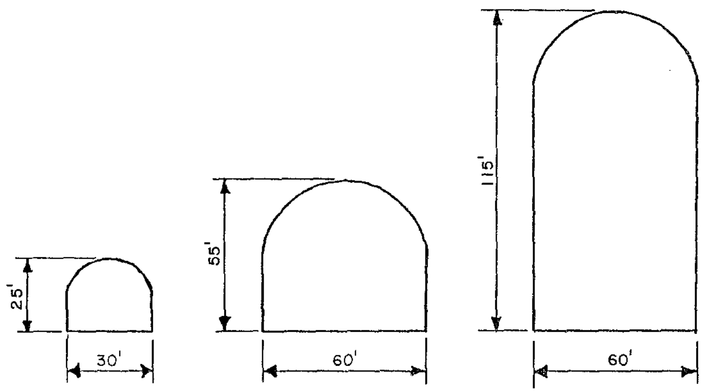


Fig. 24. Possible cavern cross-section

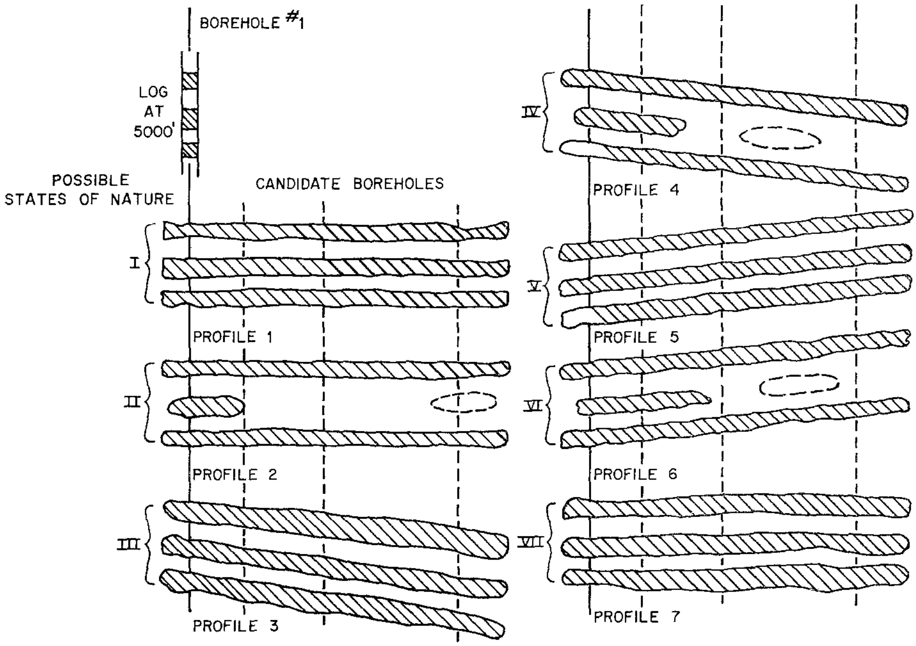


Fig. 25. Results and inferences from first boring

Three cavern sizes and shapes (Fig. 24) and 2 depths (2000′ and 5000′) are considered. Greater depths make it possible to provide the storage capacity with a smaller cavern volume. Construction cost is affected by cavern size, depth and total cavern volume as summarized in Table 9. A first boring has been completed and its log is plotted in Fig. 25.

The log indicates, at a depth of 5000 ft. (1500 m), two sandstone layers of 80 ft. (24 m) and 60 ft. (18 m) thickness. These are interbedded with shale layers, each 100 ft. (30 m) thick. Jointing appears to be low throughout the rock at this depth. From this log one infers, e. g., the seven profiles of Fig. 25 as possible states of nature. (Other inferences are possible but considered unlikely in this case.). The seven profiles denote differences in major structural aspects of the geology. Other variations within each profile, e. g., in the degree of jointing anticipated, or in the size and frequency of shale lenses, can be superimposed upon the profiles.

The questions at this point are: 1) whether to drill a second boring, and 2) if so, where? Three potential locations are considered, however, any number could be. The candidate location are:

(1) 200′ (~60 m) to the right of the first boring;
(2) 500′ (~150 m) to the right of the first boring;
(3) 1000′ (~300 m) to the right of the first boring.

Three reliability matrices will have to be developed, one for each location. They include both physical or mechanical uncertainties in taking the boring, and inferential uncertainties in extrapolating boring results. Since this latter uncertainty is a function of boring location, reliability matrices will be different for different locations.

As an illustration, Table 10 shows the reliability matrix for the proposed boring located 500′ to the right of the initial boring.

Table 10 indicates that there is thought to be a general 40% reliability that the boring results will indicate the profile correctly. The remaining 60% in each case is distributed over 6 other profiles. Profiles having features similar to the actual profile are assigned higher likelihoods, and vice versa. This process is subjective; judgment is required in assigning these reliabilities.

Because of the limited information obtained from this first boring, the seven profiles in Fig. 25 are taken to be equally likely. Their prior probabilities are therefore:

$$P\,[\text{Profile 1}] = 0.15$$
$$P\,[\text{Profile 2}] = 0.15$$
$$P\,[\text{Profile 3}] = 0.15$$
$$P\,[\text{Profile 4}] = 0.15$$
$$P\,[\text{Profile 5}] = 0.15$$
$$P\,[\text{Profile 6}] = 0.15$$
$$P\,[\text{Profile 7}] = 0.10$$
$$\Sigma \ldots\ldots\ldots = 1.0$$

With a small computer program to conduct preposterior analyses, the following results are obtained for the value of a second boring:

$$\text{EVSI}_{200'\text{ rights}} = \$\ 0.0 \text{ millions}$$
$$\text{EVSI}_{500'\text{ rights}} = \$\ 1.4 \text{ millions}$$
$$\text{EVSI}_{1000'\text{ rights}} = \$\ 6.2 \text{ millions}$$

The results show that it is most beneficial to sink the borehole 1000′ to the right. In a general sense, borings spaced at greater distances from the initial boring shed more light on regional characteristics of the geology.

Table 10. *Example Reliability Matrix*
Boring Indicates Profile #:

True profile	1	2	3	4	5	6	7
Profile 1	0.4	0.1	0.15	0.05	0.15	0.05	0.1
Profile 2	0.15	0.4	0.05	0.15	0.05	0.15	0.05
Profile 3	0.15	0.05	0.4	0.1	0.15	0.05	0.1
Profile 4	0.05	0.15	0.05	0.4	0.15	0.15	0.05
Profile 5	0.15	0.05	0.15	0.05	0.4	0.1	0.1
Profile 6	0.05	0.15	0.05	0.15	0.1	0.4	0.1
Profile 7	0.15	0.05	0.15	0.05	0.15	0.05	0.4

Borings spaced more closely tend to confirm findings in a local area, but do not reveal the regional characteristics as well. At this initial point, an indication of the regional geology is apparently more valuable than confirmation of results from the first boring. Obviously, these results vary from case to case.

In analogous manner the position of a 3rd boring, 500′ to the left or to the right of the initial one was examined. This resulted in EVSI of $ 2.8 millions for the boring located 500′ to the left and an EVSI = 0 for the other. Details of the procedure and the example, as well as other alternatives and sensitivity studies are given in Einstein et al. (1977).

IV. Closing Comments

It was possible to show that statistical methods in exploration make it possible to formally describe variability in engineering geologic conditions and in the data collected, but also to plan exploration approaches. Of particular interest are the methods that allow the use of subjective assessments of uncertainty.

In part II of this paper, it will be shown how the results of exploration programs with statistical descriptions can be used in design and construction.

References

Ashley, D. B., Veneziano, D., Einstein, H. H., Chan, M. H. (1981): Geologic Prediction and Updating in Tunnelling — A Probabilistic Approach. Proceedings, 22nd U. S. Symposium on Rock Mechanics. MIT.

Baecher, G. B. (1978): Analyzing Exploration Strategies. In: C. H. Dowding (ed.), Site Characterization and Exploration, ASCE/NSF.

Baecher, G. B., Lanney, N. A. (1978): Sampling for Joint Persistence. 19th U. S. National Symposium on Rock Mechanics.

Barnett, V. (1973): Comparative Statistical Inference. New York: John Wiley & Sons.

Cruden, S. M. (1977): Describing the Size of Discontinuities. Int. J. of Rock Mechanics and Mining Sciences *14*, 133—137.

Einstein, H. H., Vick, G. G. (1977): Geologic Model for a Tunnel Cost Model. Proceedings, 2nd Rapid Excavation and Tunneling Conference, San Francisco.

Einstein, H. H., Markow, M. J., Chew, K. (1977): Assessing the Value of Exploration for Underground Gas Storage. Proceedings, Rockstore.

Einstein, H. H., Labreche, D. A., Markow, M. J., Baecher, G. B. (1978): Decision Analysis Applied to Rock Tunnel Exploration. Engineering Geology *12*, 143—161.

Einstein, H. H., et al. (1980): Risk Analysis for Rock Slopes in Open Pit Mines, Parts I—V, USBM Technical Report J0275015.

Einstein, H. H., Baecher, G. B. (1982): Probabilistic and Statistical Methods in Engineering Geology — Problem Statement and Introduction to Solution. Rock Mechanics, Suppl. XII, 47—62.

Epstein, B. (1954): Truncated Life Tests in the Exponential Case. Annals of Mathematical Statistics *25*, 555.

Fisher, R. A. (1931): The Truncated Normal Distribution. British Association of Advanced Science, Math. Tables *I*, XXXIII—XXXIV.

Hald, A. (1949): Maximum Likelihood Estimators of the Parameters of a Normal Distribution Truncated at a Known Point, Skand. Aktuar Tidskr. *32*, 119.

Kendall, M. G., Stuart, A. (1967): The Advanced Theory of Statistics, Hafner.

Larsson, I. (1952): A Graphic Testing Procedure for Joint Diagrams. American Journal of Science *250*, 586—593.

Moavenzadeh, F., Markow, M. J. (1974—1978): Tunnel Cost Model, Series of 10 reports.

Priest, S. D., Hudson, J. (1981): Estimation of Discontinuity Spacing and Trace Lengths Using Scan Line Surveys. Int. J. Rock Mech. Min. Sci. *18*, 183—198.

Sander, B., et al. (1954): Einführung in die Gefügekunde. Wien: Springer-Verlag.

Snow, D. (1966): Disc'n. — Theme III, Int. Cong. Rock Mechanics, Lisbon.

Staël von Holstein, C. S. (1974): A Tutorial in Decision Analysis. In: R. A. Howard, et al. (eds.), Reading in Decision Analysis, SRI.

Terzaghi, R. (1964): Sources of Error in Joint Surveys. Geotechnique *15*, 287—304.