Proceedings
Foundation Eng. Congress
1981
ASCE Geotechnical + Construction
Div. Spec. Conf.

A USE OF GEOSTATISTICS TO PREDICT THE OCCURRENCE OF COLLAPSING SOILS

Edward A. Nowatzki[1], M.ASCE, Molla M. Ali[2], A.M.ASCE
and Donald E. Myers[3]

ABSTRACT: Collapse susceptible soils are those that are stable under load while they are dry, but undergo large decreases in volume virtually instantaneously, when they become wetted to near saturation moisture contents while under the same load. Their occurrence is widespread throughout the world, mostly in arid and semi-arid regions. In the United States they are found predominantly in the southwest and midwest. Locally their distribution can vary widely, both areally and with depth. The ability to estimate the probability of occurrence of such soils with a known degree of certainty would be invaluable to planners, developers, and geotechnical engineers in the areas where these soils are known to exist.

The geostatistical method of Indicator Kriging (IK) was used to estimate the probability distribution of the collapse-related soil parameter, γ_d, within Tucson, AZ. The parameter γ_d is defined as the insitu dry unit weight of soil. Test data from over 400 different locations with almost 1000 sampling points at depths ranging from 0 to 12.16 m were obtained from local consultants and previous researchers. These data were organized into a massive data base and used for the geostatistical analyses. Probability contour plots having known variance of estimation were generalized for the probability of occurrence of soils having "low", medium", and "high" collapse susceptibility. The probability and estimated variance contour maps, when superimposed on a grid map of the study area for each level of collapse susceptibility, can be used to estimate the probability of occurrence of collapse susceptible soils with known variance of estimation even for areas where no test data are available. Contour plots of the probability of encountering soils having high collapse susceptibility based on the collapse-related parameter γ_d are presented in this paper for illustration.

---

1 - Associate Professor, Department of Civil Engineering and Engineering Mechanics, University of Arizona, Tucson, AZ 85721.

2 - Staff Engineer, Western Technologies Inc., 3737 E. Broadway Road, Phoenix, AZ 85036.

3 - Professor, Department of Mathematics, University of Arizona, Tucson, AZ 85721.

INTRODUCTION

The use of probabilistic models and statistical techniques in various disciplines of geotechnical engineering has increased rapidly in the recent past. These techniques are not limited to data analysis alone but include reliability assessment of earth structures and other constructed facilities, risk assessment for regulatory control, economic optimization, and project feasibility (Beacher, 1984).

Theories of statistics and probability are used in these analyses for an assessment of the uncertainty in the prediction of the performance of the structures which were designed on the basis of average parameters obtained from a limited number of laboratory and/or field tests. Generally the number of samples tested is small, therefore the statistical estimation error plays an important role in the reliability of the prediction.

In order to examine the variability of a given soil parameter over an area, classical statistical methods rely upon an analysis of variance to quantify the variability without regard to direction. The sample size necessary to estimate the mean within some specified confidence interval is usually calculated on the basis of a frequency distribution of the observations (Vieira et al, 1981). Regardless of the sampling plan, these methods provide an incomplete description of the variability of the parameter of interest because there is no link between the calculated variances and the physical distance between sampling points. In other words, a knowledge of the frequency distribution of the observations alone does not provide any information about the variability of the observations with respect to spatial locations. The spatial variability of a soil parameter given in terms of position coordinates provides an additional characterization not derived from the frequency distribution of observations alone.

Geostatistics provides the tools necessary for an adequate description of the spatial variability of a soil parameter as well as an unbiased estimation of its value at non-sampled locations. While the application of geostatistical methods to typical geotechnical engineering problems may be questionable because of the limited number of samples generally tested in most projects, there are examples of their application in cases where large amounts of data are available from reliable sources, as is the case in this study.

BACKGROUND AND SCOPE

For this study collapse-related soil data were collected from the job files of consulting engineering firms in Tucson, AZ. and from the reports of previous researchers (e.g. Sabbagh, 1982). An attempt was made to consider as much as possible of the entire area of Tucson. Data was collected for 411 different locations having 992 sampling points. Sample depths ranged from the surface to approximately 12.2 meters. These data were reduced and arranged

Table 1. Data Set Characteristics.

Data Set	Depth (m)	Number of Samples
1	0 - 0.30	125
2	0.30 - 0.6	286
3	0.61 - 0.91	254
4	0.91 - 1.22	100
5	1.22 - 1.83	104
6	1.83 - 12.2	123
7	0 - 12.2	219

into seven data sets according to depth as shown in Table 1. In addition to sampling location grid coordinates, all data sets contain values for each of the following parameters: D = depth of sample (m), C_p = percent collapse following saturation under load (Jennings and Knight, 1957), e_0 = insitu void ratio, n_0 = insitu porosity, S_0 = insitu degree of saturation (%), γ_d = insitu dry unit weight (kN/cu m), and w_0 = insitu moisture content (%). Data set seven was derived from the first six data sets for sampling points that included the above parameters plus three additional parameters, R, A, and PL where R = Gibbs' (1961) collapse parameter, A = Alfi's (1984) collapse parameter, and PL = plastic limit.

The main goal of this study was to obtain the probability distribution of each of the above-listed collapse criteria and collapse-related soil parameters throughout the Tucson Basin for each of three pre-defined cut-off values corresponding to "low," "medium," and "high" collapse susceptibility. Because of space limitations only selected results for the collapse-related soil parameter γ_d will be presented here.

Prior to the presentation of the results, a brief description of the geostatistical methods used in the analysis will be given. These include methods for estimating the probability distribution, the principles of Indicator Kriging, estimation of the cumulative probability function, and the concept of the indicator variogram.

ESTIMATION OF PROBABILITY DISTRIBUTION

Various techniques for estimating regionalized variables have appeared in the geostatistical literature over the past 10 to 15 years. These techniques can be grouped into two broad categories: parametric techniques and nonparametric techniques. In parametric estimation, assumptions are made concerning the distribution of the variable under study. Multi-Gaussian Kriging (MK) and Disjunctive Kriging (DK) are examples of parametric estimation procedures. MK is based on the assumption of multivariate normality after a suitable transformation. DK is based on the hypothesis of bivariate normality after an appropriate transformation.

These techniques share a distinct disadvantage in that the assumption of multivariate or bivariate normality may not hold for all applications. Additionally, there are no available statistical tests with which to investigate the validity of a multivariate distribution hypothesis. Consequently, it is possible that these techniques will be applied to problems which do not conform to the initial assumption. In such cases, parametric methods will yield erroneous results. Moreover, both of these methods are difficult to apply because of their mathematical complexity (Sullivan, 1984).

In contrast, nonparametric methods do not require assumptions concerning the distribution of the variable under study (Journel, 1983). Indicator Kriging (IK) and Probability Kriging (PK) are examples of nonparametric estimation procedures. They are distribution-free, nonlinear methods in which the variables at each sampling location are transformed into a distribution function. This function is then used to estimate the spatial distribution of the variable as well as its probability distribution within a defined region. When a distribution is skewed or has "heavy tails", as is the case for the lognormal and gamma distributions, linear techniques such as MK and DK yield poor estimates of the probability distribution. For this reason, nonlinear estimation methods such as IK or PK are preferable. In addition, nonparametric solutions are easier to obtain than parametric solutions because the estimates are derived from solutions of linear kriging systems which are nearly identical to ordinary kriging systems. In light of these considerations, IK was used to estimate the probability distributions of all collapse criteria and collapse-related soil parameters for each data set.

INDICATOR KRIGING

The theory and development of nonparametric estimators of spatial distributions is similar to the theory and development of nonparametric estimators of the local mean value. The major differences involve the variables estimated and the types of data used. At a given location, x, a random variable, z(x), can be defined. The set of these random variables is called a random function. At a given point in space, z(x) is then a random variable. Over a region, however, the random function incorporates the complete spatial correlation structure of any subset of the random variables. In that case the random function describes both the random and structured aspects of the variable of interest.

Before an optimal estimator can be defined, some knowledge of the spatial law of the random function is necessary. Because the data is from only one realization of the random function, and because it is impossible to infer properties of a random function from a single realization, some form of stationarity must be assumed. Generally, stationarity of the bivariate distribution of z(x) and z(x+h) for various values of the distance vector, h, is invoked for nonparametric estimation of spatial distributions. This has the mathematical form:

$$F_{x,x+h}(z,z') = \Pr\,[Z(x) < z' \,,\; Z(x+h) < z'] \qquad (1)$$

One consequence of bivariate stationarity is that the bivariate distribution of two random variables $z(x_1)$ and $z(x_2)$ depends only upon the magnitude of the distance, h, separating the two variables and not on the particular locations x_1 and x_2. This property implies stationarity of the univariate distribution, i.e., the random variables representing the variable of interest at each particular location $z(x_1)$, $z(x_2)$, ..., $z(x_n)$ are identically distributed and independent (Sullivan, 1984). With the stationarity of the bivariate distribution established, nonparametric estimators of the probability distribution can be developed.

THE INDICATOR FUNCTION

The spatial distribution of a random function with point support within a region A can be defined mathematically as:

$$\phi(A,z_c) = \frac{1}{A}\int_{x \in A} i(x,z_c)dx \qquad (2)$$

where $\phi(A,z_c)$ = spatial distribution of point values within the region A for a cut-off z_c (i.e. the proportion of point values within the region A less than z_c); $i(x,z_c)$ = an indicator that takes values of 1 or 0 as follows: $i(x,z_c) = 1$ if $z(x) \le z_c$, $i(x,z) = 0$ if $z(x) > z_c$; where $z(x)$ = the observed value of the variable under study at location x, and z_c = cut-off value of the variable under study.

The spatial distribution $\phi(A,z_c)$ can be considered as a realization of a random variable $\Phi(A,z_c)$, given by

$$\Phi(A,z_c) = \frac{1}{A}\int_{x \in A} I(x, z_c)dx \qquad (3)$$

where $I(x,z_c)$ is an indicator random function that takes values of 1 or 0 according to the same criteria as given previously for $i(x;z_c)$. The expected value of this random variable is given by:

$$E[\Phi(A,z_c)] = \frac{1}{A}\int_{x \in A} E[I(x,z_c)]dx$$

$$= F(z_c) = Pr\ [Z(x) < z_c] \qquad (4)$$

where $F(z_c)$ = the stationary univariate distribution function. Thus the spatial distribution function, $\phi(A,z_c)$, is a realization of a random function, $\Phi(A,z_c)$, whose expected value is equal to the value of the stationary univariate cumulative distribution function, $F(z_c)$. The indicator variables change as the cut-off values change. If the cut-off value increases, more of the sampled values fall below the cut-off value causing more of the indicator variables to take the value 1.

ESTIMATION OF CUMULATIVE PROBABILITY FUNCTION

Transforming the raw data by indicator variables is done for the purpose of estimating the cumulative probability function $\phi(A,z_c)$, which represents the proportion of points in an area A where the value of a variable $z(x)$ is below the cut-off value z_c. Hence the value of $\phi(A,z_c)$ can be taken as the probability that a parameter is below the cut-off value, z_c.

To estimate $\phi(A,z_c)$, a linear estimator, $\phi^*(A,z_c)$, is used:

$$\phi^*(A,z_c) = \sum_{\alpha=1}^{n} \lambda_\alpha i(x_\alpha, z_c) \qquad (5)$$

To insure that ϕ^* is unbiased, it is required that:

$$\sum_{\alpha=1}^{n} \lambda_\alpha = 1 \qquad (6)$$

where the weights λ_α are obtained from a kriging system using the variogram for the indicator variable.

As an alternative to imposing the unbiasedness condition given in Eq. 6 and using variograms, indicator covariance and an estimate for the mean, $F^*(z_c)$, could be used. In this case, simple indicator kriging is used rather than ordinary indicator kriging. The final form of the estimator then becomes:

$$\phi^*(A,z_c) = \sum_{\alpha=1}^{n} \lambda_\alpha i(x_\alpha, z_c) + [1 - \sum_{\alpha=1}^{n} \lambda_\alpha] F^*(z_c) \qquad (7)$$

The estimation variance of this estimator is given by:

$$\sigma^2_{ik} = \bar{C}_1(A,A,z_c) - \sum_{\alpha=1}^{n} \lambda_\alpha \bar{C}_1(x_\alpha, A, z_c) \qquad (8)$$

where

$$\bar{C}(A,A,z_c) = E[\,\Phi(A,z_c)\,\Phi(A,z_c)] - [F(z_c)]^2 \qquad (9a)$$

$$\bar{C}(x_\alpha, A, z_c) = E[\,\Phi(A,z_c)\ I(A,z_c)] - [F(z_c)]^2 \qquad (9b)$$

For each region, $\phi^*(A,z_c)$ is a function of the cut-off value z_c. If a series of cut-off values is applied, as was the case in this study for low, medium, and high collapse susceptibility, a series of estimates will be obtained. In each series, as the cut-off value increases , $\phi^*(A,z_c)$ also increases because the percentage of points below the cut-off increases. This characteristic of the estimator function $\phi^*(A,z_c)$ is the same as that of the indicator variables, $i(x,z_c)$ noted previously. However, since $\phi^*(A,z_c)$ is to be used as a cumulative probability function, the following order relationships must also be satisfied: $\phi^*(A,-\infty) = 0$ and $\phi^*(A,+\infty) = 1$; $0 \leq \phi^*(A,z_c) \leq 1$ for all regions of A and for all cut-off values z_c; and $\phi^*(A,z_c)$ is increasing, i.e. $\phi^*(A,z_c) \leq \phi^*(A,z'_c)$ if $z_c \leq z'_c$.

The order relationships above are not satisfied if either of the following conditions occurs: (a) $\phi^*(A,z_c)$ estimated by the method of indicator kriging is decreasing, i.e., $\phi^*(A,z_c) > \phi^*(A,z'_c)$; or (b) $\phi^*(A,z_c)$ has values less than 0 or greater than 1. In short, if an estimated distribution has "order relation problems," it is not a valid distribution function. When an order relation problem occurs, a simple method of resolving it is to set $\phi^*(A,z_c)$ equal to $\phi^*(A,z'_c)$.

INDICATOR VARIOGRAM

A natural way to compare the values of a soil parameter Z(x) at two points in space is to consider the difference in the values. For a set of pairs of sample points a certain separation distance apart, the absolute average of the difference in values of the soil parameter can easily be obtained. For mathematical reasons, the squared distances are usually considered instead, and the dissimilarity function is chosen as:

$$2\gamma(h) = \text{Ave}\,[Z(x) - Z(x+h)]^2 \qquad (10)$$

The term $2\gamma(h)$ is known as the "variogram". Being a function of the distance vector, it expresses how the value of a parameter varies with distance in a given direction. If the data for the parameter exhibits directional anisotropy, then the $\gamma(h)$ function will also depend on direction as well as separation distance and should be written as $\gamma(h,\theta)$. The variogram can also be interpreted as the elementary estimation variance of a variable Z(x) by another variable Z(x+h) at a distance h units from x. As such it is expressed as:

$$E([Z(x+h) - Z(x)]^2) = 2\gamma(h) \qquad (11)$$

The variogram can be estimated by:

$$2\gamma^*(h) = \frac{1}{N(h)} \sum [Z(x_i) - Z(x_i+h)]^2 \qquad (12)$$

where $[(x_1, x_1+h); \ldots; (x_{N(h)}, x_{N(h)}+h)]$ are N(h) pairs of samples separated by the distance vector h. Because of the factor two, $\gamma(h)$ is called the "semi"-variogram, and $\gamma^*(h)$ is called the experimental "semi"-variogram. This function bears the same relationship to γ as a histogram does to a probability distribution (Clark, 1979).

Thus the sample variogram represents the measure of the difference between the values of a parameter at locations a distance h units apart. The easiest way to display these differences is in the form of a graph, with the value of the variogram as the ordinate and the distance between pairs of samples as the abscissa. A typical plot is shown in Figure 1 for the so-called "Spherical Model Variogram". When h = 0, two samples are at exactly the same location and the difference between the two values of the parameter of interest is expected theoretically to be zero. If the distance between the two samples increases a little, then some difference in the value of the parameter is expected. Therefore, the sample variogram takes a small positive value. As the sampling points move farther apart, the differences in the values of the parameter can be expected to rise. In the ideal case, when the distance becomes very large, the values of the parameter become independent of one another. The variogram value then becomes nearly constant since it

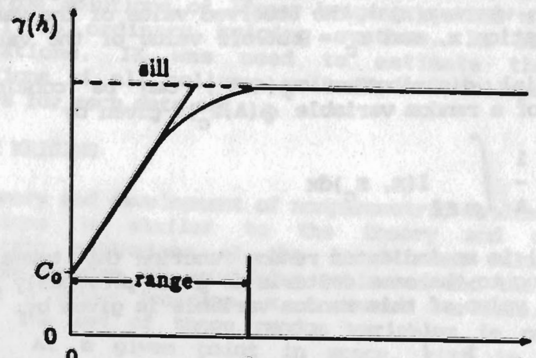

Figure 1. Spherical Model Variogram.

represents the difference between sets of independent samples. The shape of the variogram gives an indication of the spatial dependence of the soil parameter being investigated. If, for all values of separation distance greater than zero, $\gamma(h)$ is found to be a constant, say C_0, then the variogram indicates that the observations are spatially independent. If, on the other hand, the value of $\gamma(h)$ approaches C_0 gradually or in some consistent manner as the value of the lag increases, then the observations are spatially dependent. The distance at which samples become independent of one another is called the "range of influence" of a sample and is usually denoted by the symbol "a". The value of $\gamma(h)$ at which the graph levels off is usually denoted by the symbol "C" and is called the "sill" of the variogram. In geostatistical analyses it is often found that $\gamma(h)$ does not tend toward zero when h approaches zero. This means that the variability between two values, Z(x) and Z(x+h), taken at two close points may be significant and may increase as the size of the discontinuity at the origin of $\gamma(h)$ increases. This local variability is a random phenomenon. However, as the distance h increases, the variability often becomes more continuous. This discontinuity of the variogram at the origin is known as the "nugget" effect. Essentially it describes the variogram of a purely random phenomenon. All of these concepts are illustrated in Figure 1 for the Spherical or Matheron Variogram Model (Clark, 1979)

The theory of regionalized variables, as described by Journel and Huijbergts (1978), is the foundation of the entire field of geostatistics. The modelling of variograms is the first and most essential step in applying the technique of kriging to obtain unbiased estimations. With kriging, a considerable amount of computation is necessary to obtain an adequate estimate of the variogram because of the empirical and subjective nature of the estimation process. Theoretically, the variogram Z(x) is defined by Knudsen and Kim (1978) as:

$$\gamma(h) = \frac{1}{2} \text{Var} [Z(x+h) - Z(x)] \quad (13)$$

The formal definition of the variogram is given by:

$$\gamma(h) = \frac{1}{2V} \int_V [Z(x+h) - Z(x)]^2 dx \quad (14)$$

where Z(x) is the value of a particular parameter at the point x and Z(x+h) is its value at a point located a distance h from x. The variogram $\gamma(h)$ is a function of the distance h between the points x and x+h in one-, two-, or three-dimensional space. In practice, variograms are computed by using a discrete number of points obtained at incremental distances. Therefore Equation 14 can be written as:

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} [Z(x_i) - Z(x_i+h)]^2 \quad (15)$$

where N_h represents the number of sample pairs each having a separation distance h. All samples are assumed to lie on a straight line along which the computation is being performed. The basic unit h used for the interval in Equation 15 is known as the "class size." A tolerance on the class size is required so that data points closer to one class interval than another can be considered during the computation. Once the distributions of the indicator variables at sampled points have been determined, the indicator variogram for the distribution can be estimated by:

$$\gamma_1(h,z_c) = \frac{1}{2N_h} \sum_{i=1}^{N_h} [i(x+h,z_c) - i(x,z_c)]^2 \quad (16)$$

The indicator variogram can also be interpreted as a bivariate probability where:

$$\gamma_1(h,z_c) = 0.5 \, (\Pr [Z(x) > z_c \text{ and } Z(x+h) \le z_c]) + 0.5 \, (\Pr [Z(x) \le z_c \text{ and } Z(x+h) > z_c]) \quad (17)$$

The indicator variograms $\gamma_1(h,z_c)$ are estimated for each cut-off value of the collapse-related soil parameter, γ_d, obtained for low, medium, and high collapse susceptibility. Figure 2 shows the indicator functions for the three cut-off values. These functions can be expressed mathematically as:

a) For high collapse susceptibility,

$$i(x,z) = \begin{cases} 1 \text{ if } Z \ge z_{ca} \\ 0 \text{ if } Z < z_{ca} \end{cases} \quad (18a)$$

b) For medium collapse susceptibility,

$$i(x,z) = \begin{cases} 1 \text{ if } z_{ca} \le z_{cc} \le z_{cb} \\ 0 \text{ otherwise} \end{cases} \quad (18b)$$

c) For low collapse susceptibility or noncollapse,

$$i(x,z) = \begin{cases} 1 \text{ if } Z \le z_{cb} \\ 0 \text{ if } Z > z_{cb} \end{cases} \quad (18c)$$

Table 2. Cut-Off Values of Dry Unit Weight (γ_d) for High, Medium, and Low Collapse Susceptibility.

Collapse Susceptibility	Value of γ_d (kN/cu. m.)
High (HC)	$\gamma_d < 14.3$
Medium (MC)	$14.3 < \gamma_d < 15.5$
Low (LC)	$\gamma_d < 15.5$

The cut-off values of γ_d for low (LC), medium (MC), and high (HC) collapse susceptibility used in this study are listed in Table 2. They are the same as those proposed by Sabbagh (1982) based on conventional volumetric-gravimetric relationships among parameters for which others had determined collapse susceptibility cut-off values. The statistical significance of each of these cut-offs was evaluated by Ali (1987).

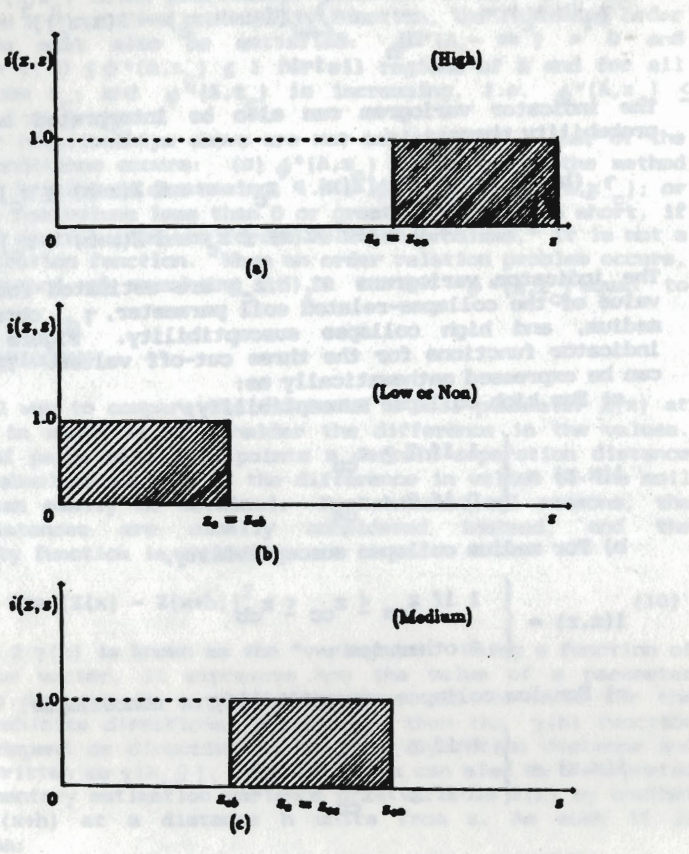

Figure 2. Indicator Functions for Selected Cut-Off Levels: (a) High Collapsing (HC), (b) Noncollapsing (NC), (c) Medium Collapsing (MC).

Two types of variogram models were found to be appropriate for all the collapse-related soil parameters and collapse criteria investigated. These are the Spherical Model with range a, sill C, and nugget C_0; and the Pure Nugget Model with nugget value C_0. The indicator variograms for each parameter were obtained for each cut-off level following the same procedure as is followed for ordinary variograms. All of the geostatistical computations, including variogram estimation and kriging, were performed using the computer program BLUEPACK (1982).

RESULTS AND DISCUSSION

Equations 7 and 8 were used with the indicator variogram to develop contour plots of the probability that values of the collapse-related soil parameters and collapse criteria considered in the study would have a value in the ranges corresponding to low, medium, or high collapse susceptibility, as well as to generate contour plots for the associated variance of estimation for each condition. The results for high collapse susceptibility predicted on the basis of γ_d are presented in Figure 3 for Data Set 1, which pertains to sample depths within 0.31 meters of the surface. This depth range corresponds approximately to founding elevations of most residential and light commercial buildings in Tucson. The solid contours in Figure 3a represent the probability that the value of γ_d is greater than or equal to the cut-off value for high collapse susceptibility. The broken contour lines in Figure 3b show the corresponding variance of estimation. These contours indicate the reliability of the estimation of the probabilities shown by the solid contours.

The cross-hatched zones in Figure 3a indicate the areas in Tucson where there is at least a 60% probability that a soil having a high potential for collapse based on the value of γ_d will be encountered within 0.31 meters of the surface. The broken contour lines in Figure 3b suggest that the variance of the estimate ranges between 0.6 and 0.7 which can be considered moderate to good. The accuracy of the estimate depends largely upon the development of an accurate variogram.

SUMMARY AND CONCLUSIONS

The technique of Indicator Kriging was applied to selected collapse criteria and collapse-related soil parameters determined from laboratory tests on soil samples obtained throughout Tucson, AZ., in order to model associations among variables and to estimate the probability that the value of a selected parameter is above or below a certain critical value corresponding to low, medium, or high collapse susceptibility. The following conclusions can be drawn based on values of the collapse-related parameter, γ_d:

1. The zones of high collapse susceptibility as predicted by geostatistical analyses coincide with those suggested by geomorphological considerations, i.e. collapse susceptible soils occur most frequently within the flood planes of the major ephemeral water courses in the Basin.

2. The results of this study confirm the findings of previous researchers, such as Crossley (1968), who delineated areas within Tucson where the occurrence of collapse susceptible soils could be expected based strictly on observations of structural distress.

3. The principles of geostatistics can be applied successfully to geotechnical engineering problems where large amounts of data are available from a reliable source.

4. Geostatistics is a valuable tool for characterizing and modelling the three-dimensional spatial variability of parameters related to soil behavior.

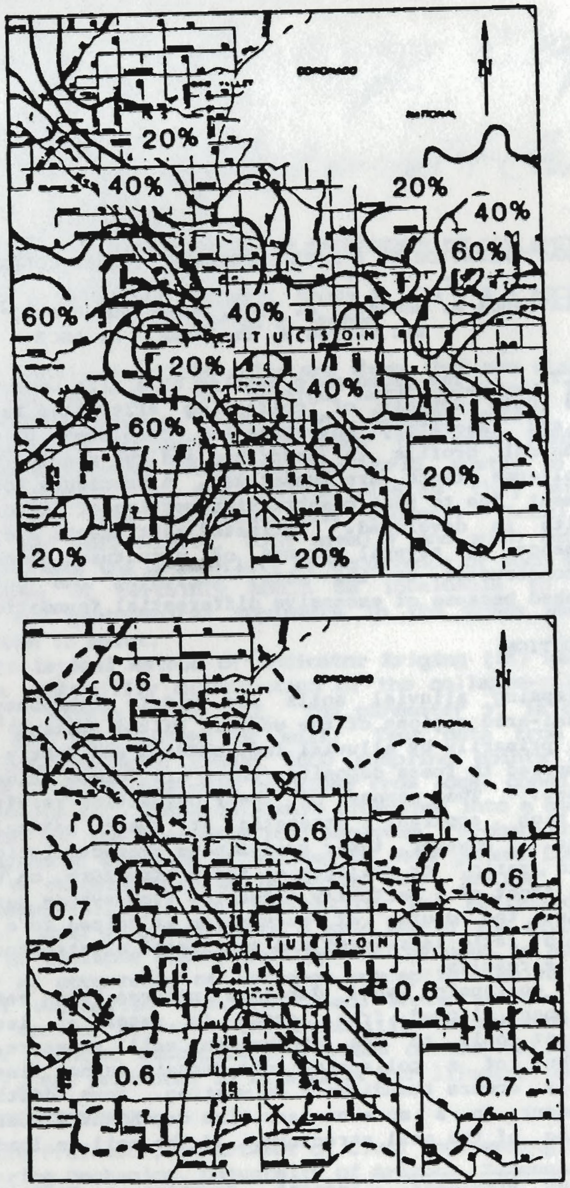

Figure 3. Contours for Soils Within 0.31m of Surface Having "High" Collapse Susceptibility Based on Dry Unit Weight Value
(a) Probability of Occurrence
(b) Estimation Variance

5. The collapse criteria and collapse-related soil parameters identified in this study can be considered as regionalized variables having spatial structures that can best be fitted by a spherical model variogram. The range of influence of the structure in this study varied from 8.9 to 12.9 km. This distance is large relative to the usual distances over which soils are sampled for conventional construction projects. Therefore, the application of geostatistical concepts for the estimation of collapse criteria and collapse-related soil parameters is justified.

6. The method of Indicator Kriging provides a means whereby contour plots can be developed of the probability that certain collapse criteria and collapse-related soil parameters are above or below predetermined cut-off values for low, medium, or high collapse susceptibility. Such plots, along with the contour plots of their estimation variance, can be used effectively by planners, developers, and geotechnical engineers to predict the occurrence of potentially poor foundation soils in areas where no samples have been taken. Information of this type can be used to establish the extent and nature of geotechnical investigations that should be undertaken in those areas.

7. The development of the variogram involves subjective evaluations and is not done by theoretically rigorous analyses. A better and more accurate picture of the probabilities and corresponding estimation variances can be obtained by having the data plotted for closer probability intervals. Development of such plots requires either larger sized paper or enlargement of smaller scale plots to a suitable scale. The probability contour plots for low, medium, and high collapse susceptibility could be combined to obtain a clear understanding of the distribution of collapse susceptible soils in Tucson. A three-dimensional picture could be obtained for any given level of collapse susceptibility by combining plots for all data sets. The reliability of the results could be improved if the data base was expanded and increased with information obtained from soil investigations related to future development in the area.

REFERENCES

1. AAS Alfi, Mechanical and Electron Optical Properties of a Stabilized Collapsible Soil in Tucson, Arizona, PhD Dissertation, University of Arizona, Tucson, Arizona, 1984, 310 pp.
2. MM Ali, A Probabilistic Analysis of the Distribution of Collapsing Soil in Tucson Using Kriging Method, PhD Dissertation, University of Arizona, Tucson, Arizona, 1987, 252 pp.
3. GB Beacher, Geostatistics, Reliability, and Risk Assessment in Geotechnical Engineering, Geostatistics for Natural Resource Characterization, G. Verly et al (eds.), D. Reidel, 1984, pp. 731-744.
4. BLUEPACK, Centre de Geostatistique Fontainebleau, France, 1982.
5. I Clark, Practical Geostatistics, Applied Science Publishing, LTD, Essex, London, 1979, 129 pp.

6. RW Crossley, A Geologic Investigation of Foundation Failures in Small Buildings in Tucson, Arizona, MS Thesis, University of Arizona, Tucson, AZ, 1969, 58 pp.
7. HJ Gibbs, Properties Which Divide Loose and Dense Uncemented Soils, US Bureau of Reclamation, Report No. EM-608, Denver, CO, 1961.
8. JE Jennings & K Knight, The Additional Settlement of Foundations Due to Collapse of Structure of Sandy Subsoil on Wetting, 4th ICSMFE, Vol. I, London, England, 1957, pp. 316-319.
9. AG Journel, Non-Parametric Estimation of Spatial Distributions, Math. Geology 15, No. 3, 1983, pp. 445-468.
10. AJ Journel & CJ Huijbregts, Mining Geostatistics, Academic Press, New York, NY, 1978, 600 pp.
11. HP Knudsen & YC Kim, A Short Course on Geostatistical Ore Reserve Estimation, Department of Mining and Geological Engineering, University of Arizona, Tucson, AZ, 1978.
12. AO Sabbagh, Collapsing Soils and Their Clay Mineralogy in Tucson, Arizona, MS Thesis, University of Arizona, Tucson, AZ, 1982, 269 pp.
13. JA Sullivan, Non-Parametric Estimation of Spatial Distribution, PhD Dissertation, Stanford University, Palo Alto, CA, 1984, 385 pp.
14. SR Vieira, DR Nielson & JW Bigger, Spatial Variability in Field Measured Infiltration Rate, Journal of the Soil Science Society of America, Vol. 45, 1981, pp. 1040-1048.