Mathematical Geology, Vol. 22, No. 1, 1990

Probabilistic Analysis of Collapsing Soil by Indicator Kriging[1]

Molla M. Alli,[2] Edward A. Nowatzki,[2] and Donald E. Myers[3]

Collapsing soils, which undergo a large decrease in bulk volume virtually instantaneously upon saturation and/or load application, are found in arid and semi-arid regions of the world. In the western and midwestern U.S., problems resulting from collapsing soils are being recognized due to rapid industrial and urban developments. A probabilistic analysis of the distribution of such soils would be a rational approach for quantifying risk involved for a project in an area where such soils are found. Indicator kriging was applied to seven sets of collapse and collapse-related soil parameters to obtain the probability that a certain parameter is more or less than a predefined critical value for low, medium, and high collapse susceptibility. Results are presented in the form of probability contour plots with known variance of estimation of the probability. The ability to predict the probability of occurrence of collapse and collapse-related soil parameters for different critical values with a known degree of certainty is invaluable to planners, developers, and geotechnical engineers.

KEY WORDS: indicator transforms, indicator kriging, collapsing soils.

INTRODUCTION

The variety of soil materials encountered in engineering problems is virtually limitless. Soils may be residual or transported. Transported soils may be alluvial (stream borne), aeolian (wind borne), glacial (ice borne), or colluvial (gravity transported). Alluvial soils are deposited when a mountain runoff flows into a valley or onto a plain. When such soils are deposited in an arid or semiarid environment, insufficient time may have occurred for them to consolidate under their own weight due to rapid evaporation rates. They become partially saturated with large voids. Application of typical foundation loads on such soils will cause only minor deformations as long as the degree of saturation remains small. As soon as the soil becomes saturated, large deformations take place due

---

[1]Manuscript received 3 October, 1988; accepted 1 June 1989.

[2]Department of Civil Engineering, University of Arizona, Tucson, Arizona 85721, U.S.A.

[3]Department of Mathematics, University of Arizona, Tucson, Arizona 85721, U.S.A.

0882-8121/90/0100-0015$06.00/1

to the reduction of volume and collapse of the intergranular structure. The rate of volume change depends on the rate at which water is available to the soil. If water is rapidly available, the subsequent volume change and deformation is rapid and the phenomenon is referred to as collapse. Thus a collapsing soil may be defined as a soil that undergoes a large decrease in bulk volume virtually instantaneously upon saturation or load application or both (Mitchell, 1976).

This paper presents a probabilistic investigation of the distribution of collapsing soils throughout the city of Tucson, Arizona. Contour plots of selected collapse criteria expressed as a function of measurable physical soil properties were developed. These plots provide information about the partial distribution of specific characteristics and quantify risks involved in any building project in a particular area. Variables considered in the analysis are mainly of two categories, established collapse criteria and collapse-related soil parameters. The two types of parameters with their abbreviations used in this investigation are defined as follows.

Collapse Parameters

C_p = % collapse = $\Delta e_c/(1 + e_0)$ where

- Δe_c = change in void ratio following saturation of specimen under a pressure P = 2000 kPa.
- e_0 = initial void ratio = initial void volume (V_{v0})/solids volume (V_s).

R = Gibb's collapse ratio = w_s/w_l where

- w_s = saturation moisture content = weight of water at saturation (W_{ws})/weight of solids (W_s).
- w_l = liquid limit moisture content = weight of water at liquid limit (W_{wl})/weight of solids (W_s).

LL = liquid limit (ASTM D-423, 1980).

A = Alfi's collapse parameter = $(e_0 - e_l)\gamma_w/[(1 + e_0)\gamma_d w_0]$ where

- e_l = void ratio at liquid limit = void volume at liquid limit (V_{vl})/solids volume (V_s).
- w_0 = initial moisture content = initial weight of water (W_{w0})/weight of solids (W_s).
- γ_d = initial dry unit weight = weight of solids (W_s)/total volume (V_T).
- γ_w = unit weight of water.

Collapse Related Soil Parameters

γ_d = initial dry density (defined previously).

e_0 = initial void ratio (defined previously).

w_0 = initial moisture content (defined previously).
n_0 = initial porosity = initial void volume (V_{v0})/total volume (V_T).
s_0 = initial degree of saturation = initial water volume (V_{w0})/initial void volume (V_{v0}).
PL = Plastic Limit (ASTM D-424, 1980).

STUDY AREA AND DATA SET

Collapse-related data have been collected from several consulting engineers' offices in Tucson and from reports of previous research conducted on collapsing soils in Tucson (e.g., Sabbagh, 1982). An attempt was made to consider as much of the entire area of Tucson (Fig. 1) as possible. Within a period of approximately 2 years, data were assembled for 411 locations with 992 sample points. Seven to 10 soil parameters were determined for each sample point that ranged from the surface to depths of about 40.0 ft. Data sets (Table 1) were established arbitrarily for ranges of depths; data sets one through six each have seven parameters D, C_p, e_0, n_0, s, γ_d, and w_0, whereas data set seven has three additional parameters R, A, and PL.

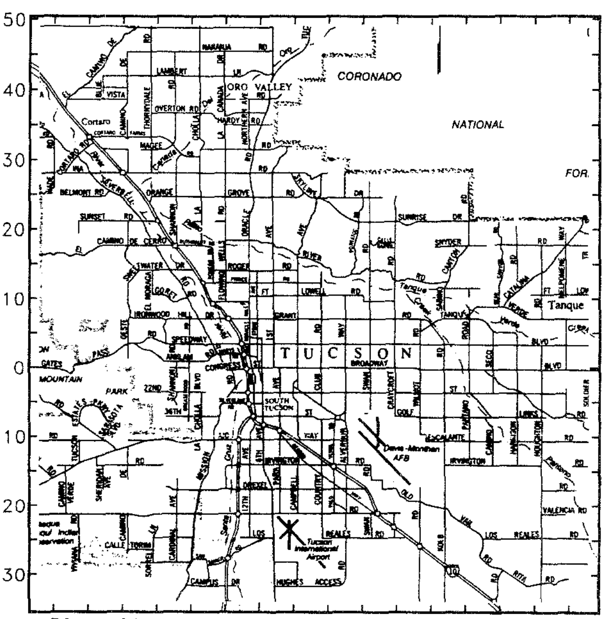

Fig. 1. Map of Tucson showing area covered in this investigation.

Table 1. Data Sets Used in the Analysis

Data set no.	Range of depths (ft.)	No. of data (N)
1	0.0–1	125
2	1.0–2	286
3	2.0–3	254
4	3.0–4	100
5	4.0–6	104
6	6.0–40	123
7	0.0–40	219

ESTIMATION OF PROBABILITY DISTRIBUTION

Various estimation techniques for estimating regionalized variables have been introduced in the geostatistical literature over the past 10–15 years. These can be broadly grouped into two categories, parametric and nonparametric. In the case of parametric methods, assumptions concerning the distribution of the variable under study are made. Multigaussian kriging (MK), disjunctive kriging (DK), and lognormal kriging (LK) are examples of parametric estimation procedures. All of these techniques share a distinct disadvantage in that the parametric hypothesis which forms the basis of the techniques may not hold for any given application. Additionally, there are no statistical tests which adequately investigate the validity of a multivariate distribution hypothesis. These techniques possibly will be applied in instances which do not conform to the initial hypothesis. In such cases, they will yield poor results. A further drawback of parametric methods is that many practitioners find them difficult to comprehend and apply due to their mathematical complexity (Sullivan, 1984).

Nonparametric methods, in contrast to parametric methods, require no assumptions concerning the distribution of the variables. The basis of these techniques is the indicator transform, which essentially transforms the variables at each location into a distribution function. The transformed data can be used to estimate the spatial distribution as well as the probability distribution of the variable within a region. In addition to not requiring any distribution hypothesis, nonparametric solutions are easier to obtain because the estimates are derived from solutions of linear kriging systems which are nearly identical to the ordinary kriging systems.

As indicated before, the main goal of the present investigation is to obtain the probability distribution of each collapse and collapse-related parameter at its three predefined cut-off levels [i.e., at low (NC), medium (MC), and high (HC) collapse susceptibility level]. Several geostatistical methods indicated above can be used for estimating the probability distributions. MK is based on an assumption of multivariate normality after a transformation. DK is based on the assumption of bivariate normality after an appropriate transformation. The requirement of bivariate or multivariate normality may not be satisfied for a specific application and these two methods are difficult to apply due to their mathematical complexity. On the other hand, indicator kriging (IK) and probability kriging (PK) are nonlinear distribution-free methods and generally are preferable to linear estimation techniques. When the distribution is skewed or has heavy tails (e.g., lognormal, gamma), modeling the variogram is more difficult. As a result, linear techniques are poor for estimating the probability distribution. In light of these considerations, IK was used to estimate the probability distributions of collapse and collapse-related parameters at different depths.

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 (Journel, 1983). The major differences between these two types of estimators are in variables estimated and types of data used. A random variable can be defined at each and every location, x_i, within the region of interest. The set of these random variables determines a random function. As a set of random variables, the random function expresses both local and regional behavior of the variable of interest. At a given point in space, the random function $Z(x)$ is a random variable, but over a region, the random function incorporates the complete spatial correlation structure of any subset of the random variables. Thus, the random function describes both random and structured aspects of the variable.

For any kind of estimation, some knowledge of the spatial law of the random function is necessary before an optimal estimator can be defined. Because the data are from only one realization and because any property of a random function cannot be inferred from a single realization without a model, a stationarity assumption must be included in the model to allow inference of the properties of the random function. The type of stationarity invoked for nonparametric estimation of spatial distributions is stationarity of the bivariate distribution of $z(x)$ and $z(x + h)$ for various values of the distance vector h; that is,

$$F_{x,x+h}(z, z') = F_h(z, z') \tag{1}$$

One consequence of bivariate stationarity is that the bivariate distribution of two random variables $z(x_1)$ and $z(x_2)$ depends on the magnitude of vector h separating these two random variables and not on the particular locations, x_1 and x_2. This implies stationarity of the univariate distribution; that is, the random variables representing the variable of interest of each particular location

$z(x_1)$, $z(x_2)$, $\cdots$, $z(x_n)$ are identically distributed, nonindependent random variables (Sullivan, 1984).

With the stationarity of the bivariate distribution established, nonparametric estimators of the probability distribution can be developed.

THE INDICATOR FUNCTION

The following material is based on Journel (1983). The spatial distribution of a random variable on a point support within a region A can be defined mathematically as:

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

where $\phi(A, z_c)$ is the spatial distribution of point values within region A for cutoff z_c (i.e., the proportion of point values within region A less that z_c) is an indicator defined as:

$$i(x, z_c) = \begin{cases} 1, & \text{if } z(x) \leq z_c \\ 0, & \text{if } z(x) > z_c \end{cases} \tag{3}$$

$z(x)$ is the observed value at location x and z_c is the cutoff value.

The spatial distribution $\phi(A, z_c)$ can be considered as a realization of a random variable $\Phi(A, z_c)$, where $\Phi(A, z_c)$ is the following integral transform of the random variable $z(x)$,

$$\Phi(A, z_c) = \frac{1}{A} \int_{x \in A} I(x, z_c)\, dx \tag{4}$$

where $I(x, z_c)$ is the indicator random function defined as:

$$I(x, z_c) = \begin{cases} 1, & \text{if } z(x) \leq z_c \\ 0, & \text{if } z(x) > z_c \end{cases} \tag{5}$$

The expected value of this random variable can be determined as:

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

$$\begin{aligned} E[\Phi(A, z_c)] &= \frac{1}{A} \int_{x \in A} \left\{ (1)\Pr[z(x) \leq z_c] + (0)\Pr[z(x) > z_c] \right\} dx \\ &= \Pr[z(x) \leq z_c] \text{ for stationarity of } z(x) \end{aligned} \tag{6}$$

$$= F(z_c) \tag{7}$$

where $\Pr[z(x) \leq z_c]$ = Probability that $z(x) \leq z_c$ and $F(z_c)$ = Stationary univariate cumulative distribution function.

The spatial distribution function $\phi(A, z_c)$ can thus be seen as 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 distribution of indicator variables at sampled points, $i(x, z_c)$, will be the same with only two values, 0 and 1. If the observed value is less than the cutoff value, the indicator value will be 1, otherwise, it will be 0. Clearly, the indicator variable $i(x, z_c)$, will change as the cutoff value increases. More samples have values less than the cutoff value; therefore, more indicator variables have the value 1.

The distribution of indicator variables is called the Bernoulli distribution and is similar to that given in Eq. (7). The expected value is

$$E[I(x, z_c)] = (1)\Pr[Z(x) \leq z_c] + (0)\Pr[Z(x) > z_c] = F(z_c)$$

and its variance

$$\mathrm{Var}[I(x, z_c)] = F(z_c)\{[1 - F(z_c)]\}.$$

ESTIMATION OF CUMULATIVE PROBABILITY FUNCTION

The purpose of transforming raw data by indicator variables is to use the indicator variables to estimate the cumulative probability function $\phi(A, z_c)$. The $\phi^*(A, z_c)$ functions are linear combinations of the indicator function. This function is the exact proportion of values, within any area A, of a variable $z(x)$, less than the cutoff value z_c. If probability measure is used, the value of $\phi^*(A, z_c)$ can be taken as the probability that an estimated parameter is less than the cutoff value.

The form of the estimator of $\phi(A, z_c)$, $\phi^*(A, z_c)$ is given by,

$$\phi^*(A, z_c) = \sum_{\alpha=1}^{n} \lambda_\alpha * i(x_\alpha, z_c) \tag{8}$$

with the constraint

$$\sum_{\alpha=1}^{n} \lambda_\alpha = 1 \tag{9}$$

for unbiasedness, where λ_α are the weights.

Simple kriging can be used to find the weights in Eq. (8) using the indicator variables $i(x, z_c)$ and indicator variogram. The final form of the estimator is:

$$\phi^*(A, z_c) = \sum_{\alpha=1}^{n} \lambda_\alpha * i(x_\alpha, z_c)\left(1 - \sum_{\alpha=1}^{n} \lambda_\alpha\right) F^*(z_c) \tag{10}$$

The estimation variance of this estimator is:

$$\sigma_{ik}^2 \overline{C}_i(A, A, z_c) - \sum_{\alpha=1}^{n} \lambda_\alpha * \vec{C}_i(x_\alpha, A, z_c) \tag{11}$$

where

$$\overline{C}(A, A, z_c) = E[\Phi(A, z_c) * \Phi(A, z_c)] - [F(z_c)]^2$$

$$\overline{C}(x_\alpha, A, z_c) = E[\Phi(A, z_c) * I(x_\alpha, z_c)] - [F(z_c)]^2$$

For each region, $\phi^*(A, z_c)$ is a function of the cutoff value z_c. If a series of cutoff values is applied, a series of estimates will be obtained. As the cutoff value increases, $\phi^*(A, z_c)$ increases, because the percentage of points with values smaller than the cutoff grade is increased. Because $\phi^*(A, z_c)$ is to be a cumulative probability function, the following order relations must be satisfied.

i. $\phi^*(A, -\infty) = 0$ and $\phi^*(A, +\infty) = 1$;
ii. $0 \leq \phi^*(A, z_c) \leq 1$ for all regions of A and for all cutoff values z_c; and
iii. $\phi^*(A, z_c)$ is nondecreasing, i.e., $\phi^*(A, z_c) \leq \phi^*(A, z_c')$, if $z_c \leq z_c'$.

If either of the following occurs

i. $\phi^*(A, z_c)$ estimated by IK is decreasing, i.e., $\phi^*(A, z_c) > \phi^*(A, z_c')$, or
ii. $\phi^*(A, z_c)$ has negative values or values greater than 1,

the order relations are not satisfied. In short, if an estimated distribution has order relation problems, it is not a valid distribution function. Whenever a order relation problem occurs, one simple method of resolving it is to set $\phi^*(A, z_c') = \phi^*(A, z_k)$ as shown (Fig. 2).

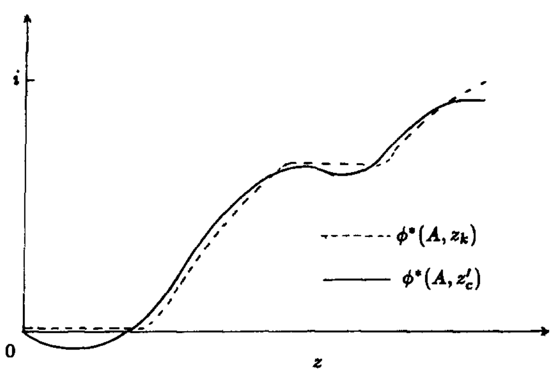

Fig. 2. Solving the order relations problem.

INDICATOR VARIOGRAM

The indicator variogram can be estimated by

$$\gamma_i(h, z_c) = \frac{1}{2N_h} \sum_{i=1}^{N_h} [i(x + h, z_c) - i(x, z_c)]^2 \tag{12}$$

where N_h = number of pairs.

The indicator variogram also can be interpreted as a bivariate probability

$$\begin{aligned}\gamma_i(h, z_c) = {} & 0.5\{\Pr[z(x) > z_c, z(x + h) \leq z_c]\} \\ & + 0.5\{\Pr[z(x + h) > z_c, z(x) \leq z_c]\}\end{aligned} \tag{13}$$

The indicator variograms $\gamma_i(h, z_c)$ are estimated for each cutoff value obtained for low, medium, and high collapse susceptibility criteria (Fig. 3). Mathematically, these functions (a), (b), and (c) can be expressed as:

For high collapse susceptibility,

$$i(x, z) = \begin{cases} 1, & \text{if } z \leq z_{ca} \\ 0, & \text{if } z > z_{ca} \end{cases} \tag{14a}$$

For low or noncollapse susceptibility,

$$i(x, z) = \begin{cases} 1, & \text{if } z \leq z_{cb} \\ 0, & \text{if } z > z_{cb} \end{cases} \tag{14b}$$

For medium collapse susceptibility,

$$i(x, z) = \begin{cases} 1, & \text{if } z_{ca} \leq z_{cc} \leq z_{cb} \\ 0, & \text{if otherwise} \end{cases} \tag{14c}$$

Critical values for parameters R, C_p, and n_0 (Table 2) were obtained from Sabbagh (1982); other critical values were derived from conventional volumetric-gravimetric relationships among the parameters.

Because moisture content (w) and degree of saturation (s) are related by the general expression $se = wC_s$, and because critical values for these parameters could not be derived individually from critical values of the other available parameters, the ratio w/s, designated as S_0 for initial conditions, was used as a measure of the critical value for moisture content [i.e., $S_0(w_0/s_0) = e_0/C_s$].

Indicator variograms for each parameter were obtained for each cutoff level. The variograms are summarized (Tables 3–5). Data set 1 through data set 6 each have five parameters (C_p, γ_d, n_0, S_0, e_0). Data set 7 has three additional parameters (R, A, PL). Two types of models were found to be appropriate for the parameters: the spherical model with range a, sill C, and nugget C_0; and

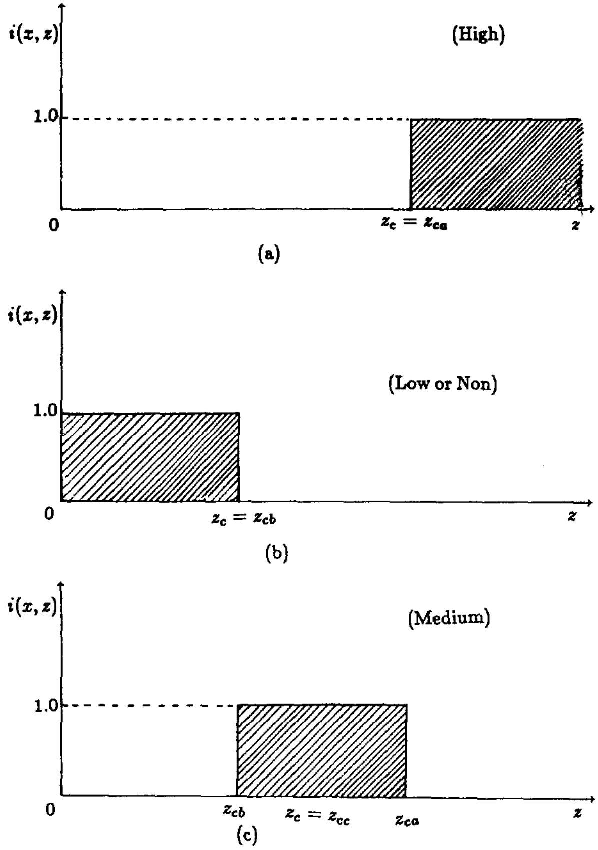

Fig. 3. Indicator functions for different cutoff levels: (a) high collapsing (HC), (b) noncollapsing (NC); (c) medium collapsing (MC).

Table 2. Critical Values for HC, NC, and MC Soil Parameters

Parameters	High collapsing (HC)	Noncollapsing (NC)	Medium collapsing (MC)
R	≥ 1.4	< 1.0	$1.0 \leq R < 1.4$
$C_p(\%)$	>5	≤ 2	$2 < C_p \leq 5$
A	$e_0 > 0.67, A > -0.67$	$e_0 < 0.67, A < -0.67$	—
$n_0(\%)$	≥ 45	<40	$40 \leq n_0 < 45$
e_0	≥ 0.82	<0.67	$0.67 \leq e_0 < 0.82$
γ_d, (pcf)	≤ 91.0	>99.0	$91.0 \leq \gamma_d \leq 99.0$
$S_0(w_0/s_0)$	≥ 0.308	<0.253	$0.253 < S_0 \leq 0.308$
PL	≥ 23	<19	$19 \leq PL < 23$

Table 3. Indicator Variogram Model Parameters of all Data Sets at HC Cutoff Levels

Data set	Parameter Z	Nugget C_0	Range a	Sill C
	C_p	0.16	25.0	0.25
	γ_d	0.21	20.0	0.24
1	n_0	0.20	25.0	0.26
	S_0	0.18	—	—
	e_0	0.19	30.0	0.25
	C_p	0.22	—	—
	γ_d	0.19	—	—
2	n_0	0.22	—	—
	S_0	—	—	—
	e_0	0.22	—	—
	C_p	0.18	15.0	0.24
	γ_d	0.16	15.0	0.19
3	n_0	0.22	30.0	0.26
	S_0	—	—	—
	e_0	0.22	30.0	0.26
	C_p	0.24	—	—
	γ_d	0.19	—	—
4	n_0	0.17	40.0	0.26
	S_0	—	—	—
	e_0	0.17	40.0	0.27
	C_p	0.18	20.0	0.22
	γ_d	0.16	—	—
5	n_0	0.19	20.0	0.22
	S_0	—	—	—
	e_0	0.18	20.0	0.22
	C_p	0.16	30.0	0.21
	γ_d	0.15	—	—
6	n_0	0.22	—	—
	S_0	—	—	—
	e_0	0.22	—	—
	C_p	0.22	25.0	0.25
	γ_d	0.18	20.0	0.25
	n_0	0.18	15.0	0.24
7	S_0	—	—	—
	e_0	0.17	20.0	0.25
	R	0.12	25.0	0.18
	A	0.12	25.0	0.18
	PL	0.11	25.0	0.22

Table 4. Indicator Variogram Model Parameters of all Data Sets at NC Cutoff Levels

Data set	Parameter Z	Nugget C_0	Range a	Sill C
1	C_p	0.12	25.0	0.18
	γ_d	0.18	20.0	0.22
	n_0	0.07	20.0	0.12
	S_0	0.16	—	—
	e_0	0.17	20.0	0.12
2	C_p	0.18	—	—
	γ_d	0.21	20.0	0.25
	n_0	0.13	40.0	0.18
	S_0	0.18	15.0	0.22
	e_0	0.12	35.0	0.17
3	C_p	0.21	—	—
	γ_d	0.21	30.0	0.26
	n_0	0.14	40.0	0.27
	S_0	0.22	—	—
	e_0	0.15	45.0	0.26
4	C_p	0.24	—	—
	γ_d	0.19	30.0	0.22
	n_0	0.17	30.0	0.24
	S_0	0.21	—	—
	e_0	0.18	25.0	0.24
5	C_p	0.21	—	—
	γ_d	0.19	25.0	—
	n_0	0.10	30.0	0.23
	S_0	0.21	—	—
	e_0	0.10	30.0	0.23
6	C_p	0.25	—	—
	γ_d	0.20	—	—
	n_0	0.22	—	—
	S_0	0.22	—	—
	e_0	0.22	—	—
7	C_p	0.14	20.0	0.19
	γ_d	0.16	20.0	0.24
	n_0	0.06	15.0	0.10
	S_0	0.14	25.0	0.19
	e_0	0.07	15.0	0.11
	R	0.17	20.0	0.26
	A	0.05	—	—
	PL	0.12	15.0	0.22

Table 5. Indicator Variogram Model Parameters of all Data Sets at MC Cutoff Levels

Data set	Parameter Z	Nugget C_0	Range a	Sill C
1	C_p	0.11	30.0	0.15
	γ_d	0.18	15.0	0.22
	n_0	0.23	—	—
	S_0	—	—	—
	e_0	0.21	—	—
2	C_p	0.19	15.0	0.22
	γ_d	0.21	—	—
	n_0	0.21	—	—
	S_0	—	—	—
	e_0	0.21	—	—
3	C_p	0.21	—	—
	γ_d	0.16	15.0	0.21
	n_0	0.17	20.0	0.21
	S_0	—	—	—
	e_0	0.17	20.0	0.21
4	C_p	0.20	—	—
	γ_d	0.20	—	—
	n_0	0.13	20.0	0.20
	S_0	—	—	—
	e_0	0.13	20.0	0.20
5	C_p	0.14	15.0	0.21
	γ_d	0.18	—	—
	n_0	0.18	—	—
	S_0	—	—	—
	e_0	0.18	—	—
6	C_p	0.21	—	—
	γ_d	0.05	15.0	0.10
	n_0	0.11	20.0	0.16
	S_0	—	—	—
	e_0	0.11	20.0	0.16
7	C_p	0.17	—	—
	γ_d	0.15	15.0	0.20
	n_0	0.16	20.0	0.22
	S_0	0.005	—	—
	e_0	0.15	20.0	0.21
	R	0.17	20.0	0.26
	A	—	—	—
	PL	0.17	—	—

the pure nugget model with nugget value C_0. In some cases, the calculated critical value is greatly different from the available measured values. Computation of the variograms is not appropriate in such cases.

PRESENTATION OF RESULTS

Equations (10) and (11) were used with indicator variograms to develop contour plots for each of the collapse criteria and collapse-related parameters. A summary of all of the probabilistic analyses for which contour plots could be developed (Table 6) uses an asterisk to indicate that indicator variograms (third column) were modeled for these parameters for HC, NC, and MC, or that indicator kriging (fourth column) was applied to estimate probability contour plots with the associated kriging variance. For parameters showing a pure nugget model, kriging is not advantageous; the absence of an asterisk indicates that no analysis was performed. Presentation of all these contour plots would require a large amount of space, only selected cases have been presented here (asterisks in fifth column of Table 6).

The estimated probability contour plots (a) (Figs. 4 and 5) and their associated variance of estimation (b) for parameters C_p and γ_d, (Data Set-1) show the probability that the value of the parameter is greater than or equal to the specified critical value for collapse in high collapse susceptibility. Regions of collapse-susceptible soils within the top foot of the surface can easily be detected from these plots.

Similar plots for NC critical values (Figs. 6 and 7) for C_p and γ_d (data set 1) show the probability that the value of the parameter is less than or equal to the specified critical value for collapse in low collapse susceptibility.

Results for MC critical values (Figs. 8 and 9) for the parameters C_p and γ_d (data set 1) show the probability that the value of the parameter is greater than the specified critical value for low collapse susceptibility but less than the critical value for high collapse susceptibility.

Results in the form of contour plots have been obtained for parameters of a data set which show some spatial structure other than the pure nugget model. The contour plots give a probability measure, and the associated variance of estimation helps to indicate how good the estimation of probabilities was. Of course, accurate estimation requires an accurate estimation of the variogram.

The usefulness of contour plots in this study is demonstrated by a numerical example. Estimated probability contours of C_p (data set 1) for HC (Fig. 10a) indicate that the zone defined by the 60–80% probability contours (shaded area) passes around the area adjacent to the Santa Cruz River. This suggests that a 60–80% probability exists that the value of C_p, at 0–1 ft. depth within this area, is greater than the HC critical value of C_p as listed (Table 2).

Table 6. Summary of Probabilistic Analysis on all Data Sets at Three Cutoff Levels

Data set	Parameter Z	Indi. Varg.			Indi. Krig.			Result		
		HC	NC	MC	HC	NC	MC	HC	NC	MC
	C_p	*	*	*	*	*	*	*	*	*
	γ_d	*	*	*	*	*	*	*	*	*
1	n_0	*	*	*	*	*	*	*	*	—
	S_0	*	*	—	—	—	—	—	—	—
	e_0	*	*	—	*	*	*	*	*	—
	C_p	*	*	*	—	—	*	—	—	—
	γ_d	*	*	*	—	*	—	—	—	—
2	n_0	*	*	*	—	*	—	—	—	—
	S_0	—	*	—	—	*	—	—	*	—
	e_0	*	*	*	—	*	—	—	—	*
	C_p	*	*	*	*	—	—	—	—	—
	γ_d	*	*	*	*	*	*	—	—	—
3	n_0	*	*	*	*	*	*	—	—	—
	S_0	—	*	—	—	—	—	—	—	—
	e_0	*	*	*	*	*	*	—	—	—
	C_p	*	*	*	*	*	*	—	—	—
	γ_d	*	*	*	—	*	—	—	—	—
4	n_0	*	*	*	*	*	*	—	—	—
	S_0	—	*	—	—	—	—	—	—	—
	e_0	*	*	*	*	*	*	—	—	—
	C_p	*	*	*	*	—	*	—	—	—
	γ_d	*	*	*	—	*	—	—	—	—
5	n_0	*	*	*	*	*	—	—	—	—
	S_0	—	*	—	—	—	—	—	—	—
	e_0	*	*	*	*	*	—	—	—	—
	C_p	*	*	*	*	—	—	—	—	—
	γ_d	*	*	*	—	—	*	—	—	—
6	n_0	*	*	*	—	—	*	—	—	—
	S_0	—	*	—	—	—	—	—	—	—
	e_0	*	*	*	—	—	*	—	—	—
	C_p	*	*	*	—	—	*	—	—	—
	γ_d	*	*	*	*	*	*	—	—	—
	n_0	*	*	*	*	*	*	—	—	—
7	S_0	—	*	—	—	—	—	—	—	—
	e_0	*	*	*	*	*	*	—	—	—
	R	*	*	*	*	*	*	*	*	*
	A	*	*	—	*	—	—	—	—	—
	PL	*	*	*	*	*	—	*	*	—

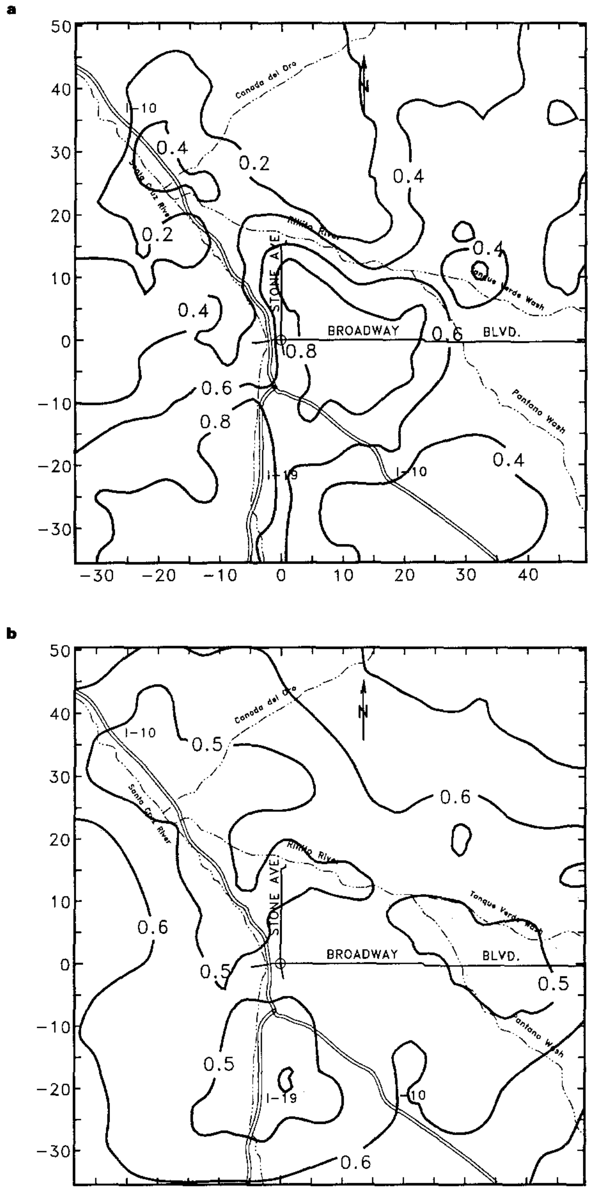

Fig. 4. Contour plots of (a) estimated probability and (b) associated kriging variance of C_p (data set 1) for HC by indicator kriging.

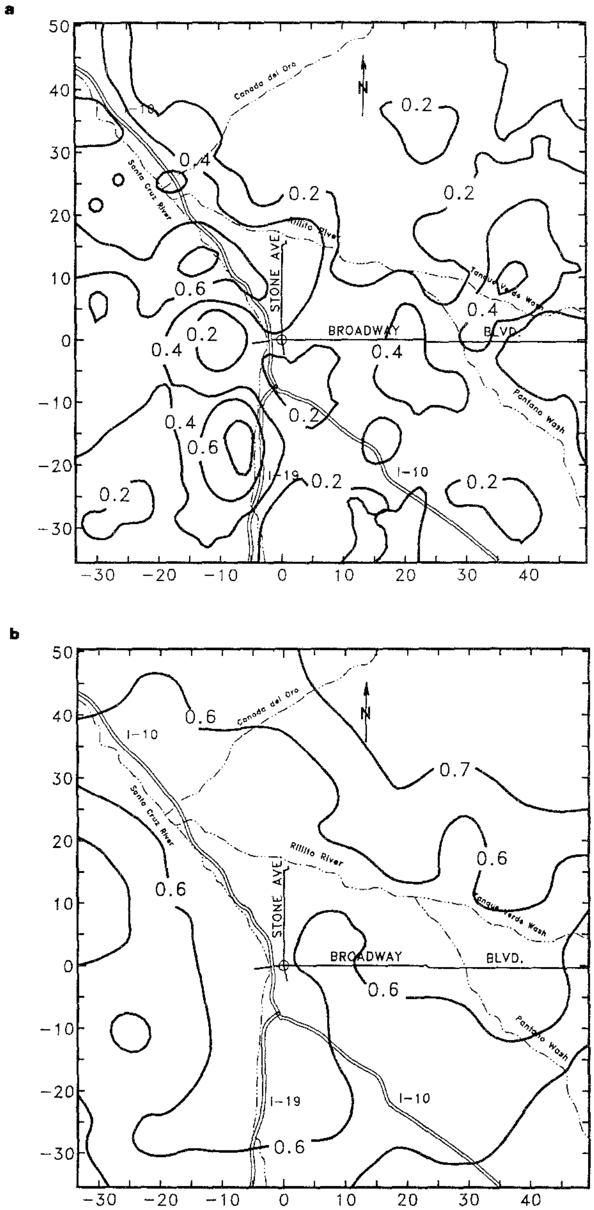

Fig. 5. Contour plots of (a) estimated probability and (b) associated kriging variance of γ_d (data set 1) for HC by indicator kriging.

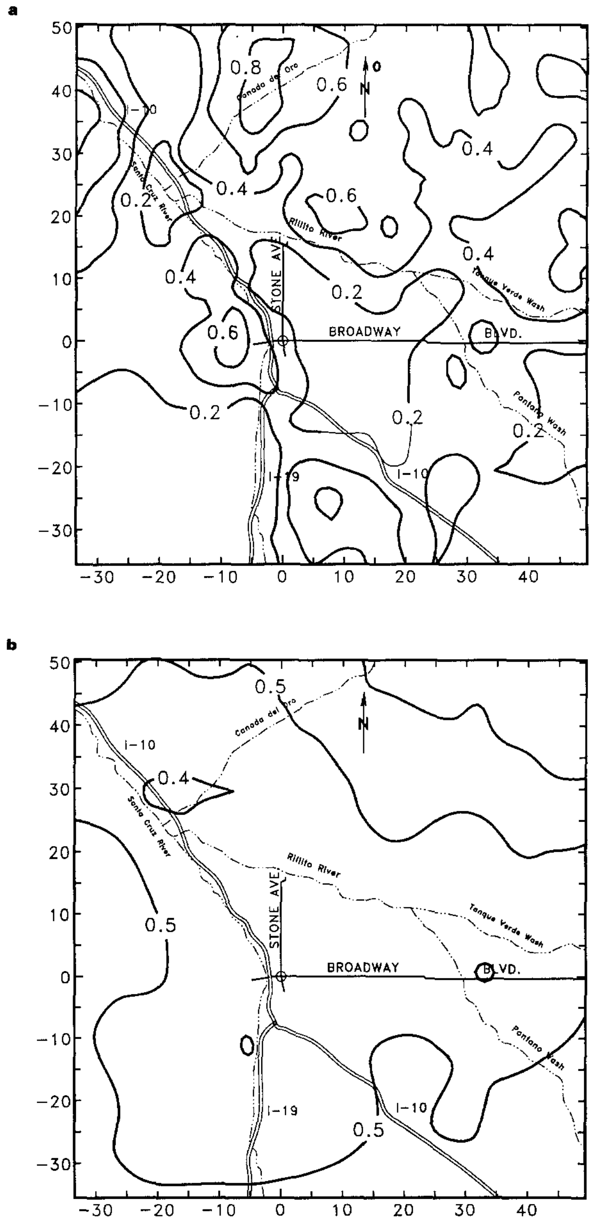

Fig. 6. Contour plots of (a) estimated probability and (b) associated kriging variance of C_p (data set 1) for NC by indicator kriging.

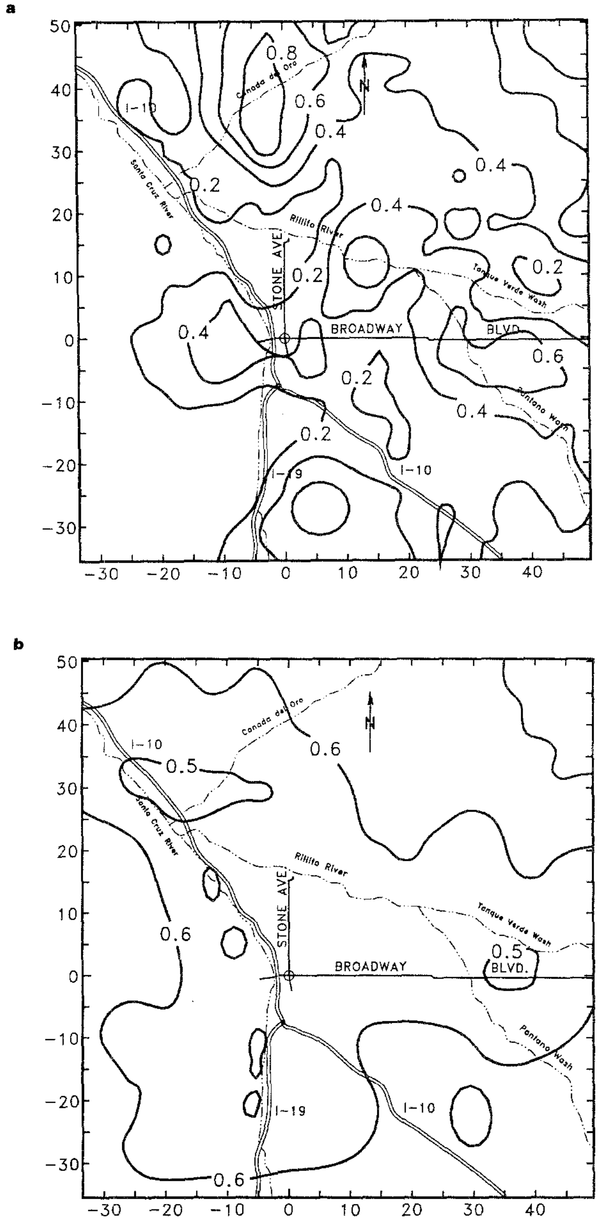

Fig. 7. Contour plots of (a) estimated probability and (b) associated kriging variance of γ_d (data set 1) for NC by indicator kriging.

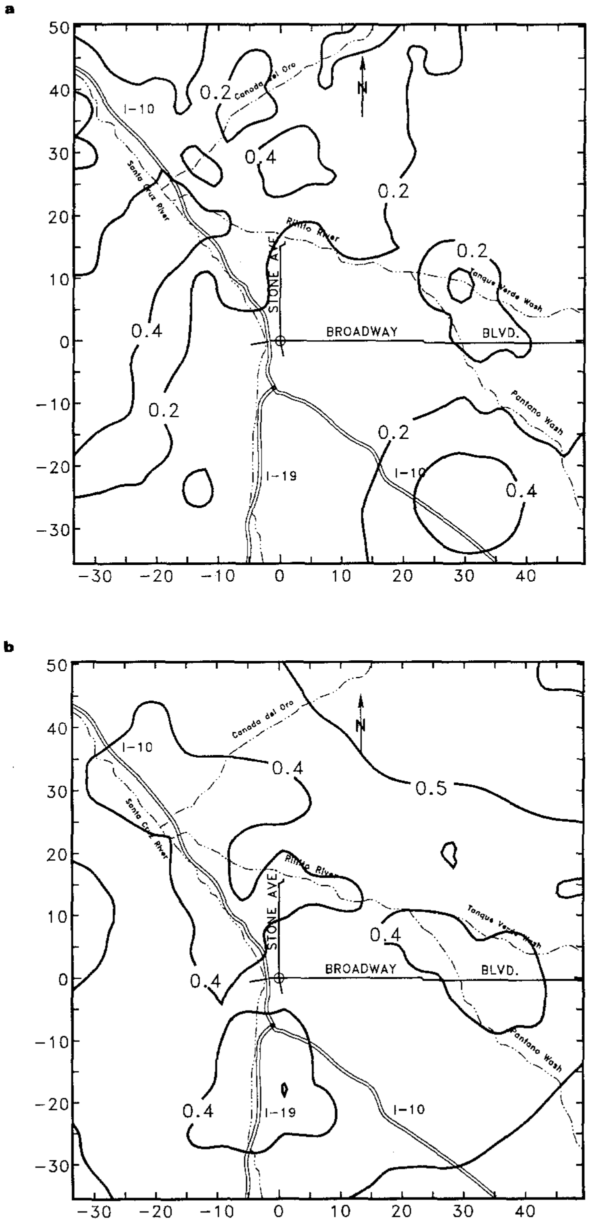

Fig. 8. Contour plots of (a) estimated probability and (b) associated kriging variance of C_p (data set 1) for MC by indicator kriging.

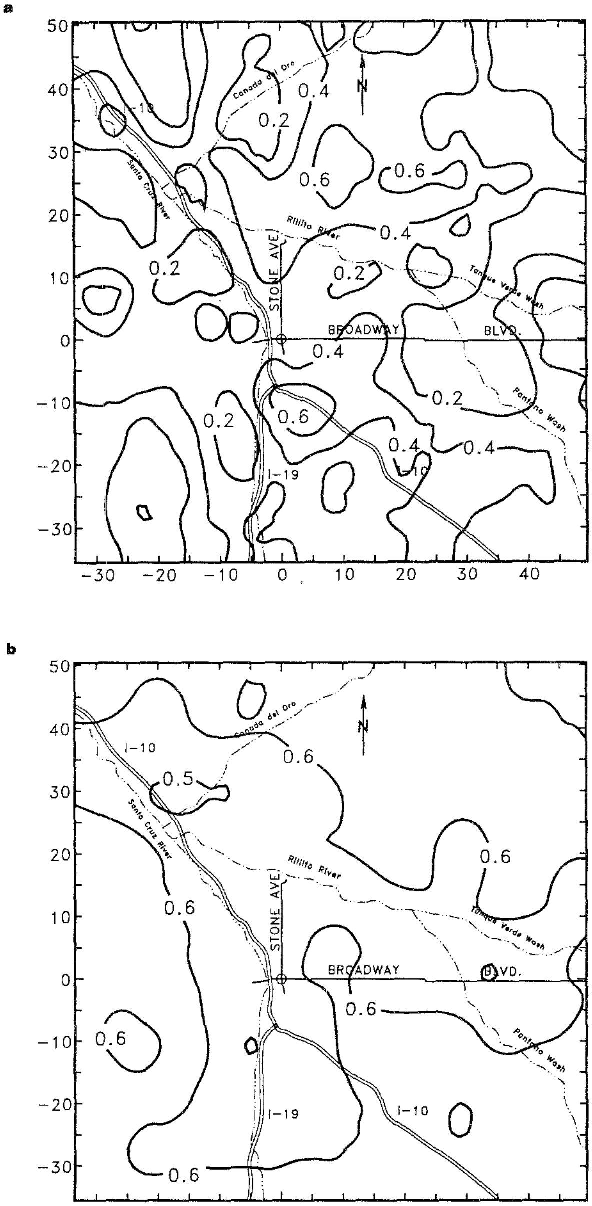

Fig. 9. Contour plots of (a) estimated probability and (b) associated kriging variance of γ_d (data set 1) for MC by indicator kriging.

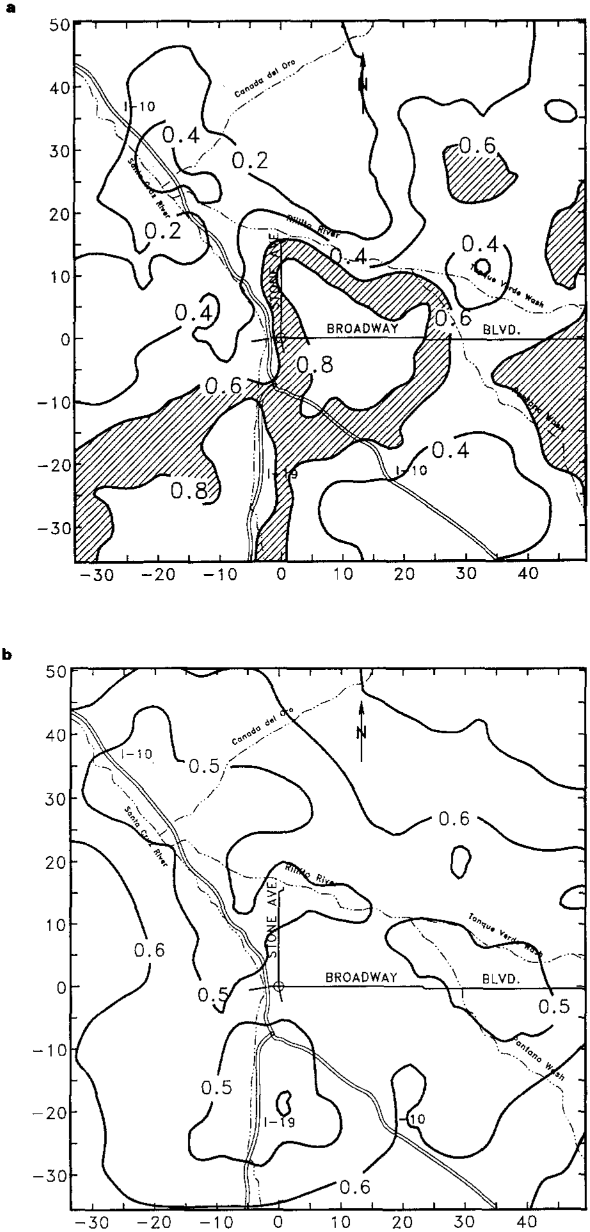

Fig. 10. Example of use of (a) probability plot to determine zone where probability of encountering soils having high collapse within top foot is between 60–80% with (b) variance of the estimation of probability.

Associated with this estimation is the estimation variance (Fig. 10b), indicating reliability. The variance has a narrow range lying between 0.5 and 0; as such, it can be considered as a moderate value. Accuracy of the estimation depends largely upon accurate development of the variogram. As discussed, the development of the variogram involves subjective evaluations and is not a rigorous analysis.

A more accurate picture of probabilities can be obtained by plotting data for smaller probability intervals. Production of such plots would require larger size paper, or enlarging small scale plots to a more suitable scale. By combining probability contour plots, a complete and clear understanding of the collapse-susceptibility can be obtained for all depths and for each separate region of Tucson.

SUMMARY AND CONCLUSIONS

Indicator kriging was applied to collapse and collapse-related soil parameters to model associations among the variables, and to estimate the probability that certain parameter is greater than or less than a critical value. Based on this study, the following conclusions can be drawn regarding the collapsing soil parameters for the Tucson area.

1. The principles of geostatistics can be applied successfully to geotechnical problems when large amounts of data are available from a reliable source.
2. Geostatistics is a valuable tool for characterizing and modeling spatial variability of geotechnical parameters.
3. The collapse and collapse-related soil parameters can be considered as regionalized variables having spatial structures that can be fitted by a spherical model variogram. The range of influence of the structure varied from 5.5–8.0 miles. This distance is large relative to usual distances over which soils are sampled. Therefore, the application of geostatistical concepts for estimation of collapse and collapse-related soil parameters is justified.
4. Indicator kriging provides a means whereby contour plots may be obtained for probabilities that certain collapse and collapse-related parameters are less than or greater than predetermined critical values for low, medium, and high collapse susceptibility. Such plots with their associated estimation variances confirm the subjective findings of a previous investigator (Crossley, 1968) for the areas of collapsing soil in Tucson.

REFERENCES

Abdullatif, A. A. 1969, Physical Testing of Engineering Properties of Collapsing Soil in the City of Tucson, (unpublished) M.S. thesis, University of Arizona, 135 p.

Anderson, F., 1968, Collapsing Soils and Their Basic Parameters in an Area of the Tucson, Arizona Vicinity, (unpublished) M.S. thesis, University of Arizona, 104 p.

ASTM, 1980, *Natural Building Stones: Soil and Rock:* Annual Book of ASTM Standards, Part 19, American Society for Testing and Materials, Philadelphia, 634 p.

Crossley, R. W., 1969, A Geologic Investigation of Foundation Failures in Small Buildings in Tucson, Arizona, (unpublished) M.S. thesis, University of Arizona, 58 p.

Journel, A. G., 1983, Non-Parametric Estimation of Spatial Distribution: Math. Geol., v. 15, p. 445–465.

Mitchell, J. K., 1976, *Fundamentals of Soil Behavior:* John Wiley & Sons, New York, 422 p.

Sabbagh, A. O. 1982, Collapsing Soils and their Clay Mineralogy in Tucson, (unpublished) M.S. thesis, University of Arizona, 269 p.

Sullivan, J. A., 1984, Non-Parametric Estimation of Spatial Distribution, (unpublished) Ph.D. thesis, Stanford University, 281 p.