# Indicator kriging geostatistical methodology applied to geotechnics project planning

RODOLFO MOREDA MENDES[1] & REINALDO LORANDI[2]

[1] *Geological Institute, Secretariat of Environment. (e-mail: rodolfo.mendes@igeologico.sp.gov.br)*
[2] *Civil Engineering Department/UFSCar. (e-mail: lorandi@power.ufscar.br)*

**Abstract:** Indicator kriging technique using SPRING 4.0 (Geographic Information System - Brazil/INPE) is presented to support preliminary investigation for urban geotechnics projects. It is applied to estimate the cumulative probability distribution function of three variables - penetration resistance index $N_{SPT}$, groundwater table and potential collapse of the soil, in order to produce indicative maps showing the probabilistic occurrence of those variables. The variable values were obtained from 241 boreholes in the São José do Rio Preto urban area (Brazil), and results of consolidation laboratory tests on undisturbed soil samples collected in this region. The indicative maps of the three variables were integrated to produce a probabilistic map for the quantitative evaluation of the more favourable sites for deep foundations. Integration of these variables assisted the construction site selection for urban infrastructure and the results demonstrated the potential for applying this method.

**Résumé:** La technique de kriging pour indication, implémentée dans SPRING 4.0 (Système de Renseignement Géographique – Brésil / INPE), est présentée comme un engin à soutenir la recherche préliminaire pour des projets géotechniques urbaines. Ça c'est appliqué pour estimer la fonction de distribution de probabilité accumulée à trois variables – l'indice de résistance à la pénétration, $N_{SPT}$, le niveau d'eau souterraine et le potentiel de collapse du sol. L'objectif est l'obtention des cartes indicatives de la probabilité d'occurrence de ces variables là. Les valeurs des variables ont été obtenues à partir de 241 orifices dans la région urbanisée de São José do Rio Preto (Brésil). Ils résultent des essais de consolidation au laboratoire, performers en indéformables échantillons de sol extraites de cette région là. Les cartes indicatives concernant les trois variables ont été intégrées pour produire une carte probabilisme à être utilisé dans l'évaluation quantitative des lieux les plus favorables à l'implémentation des fondations profondes. L'intégration des ces variables a aidé la sélection des lieux pour la construction de l'infrastructure urbaine. Les résultats positifs ont démontré le potentiel de la méthodologie applique.

**Keywords:** Geographic information systems, geotechnical maps, foundations, planning, penetration test.

## INTRODUCTION

Geographical Information Systems are now widely used in both public organisations and private companies. GIS technology is gradually being introduced as a tool to aid decision-making in several scientific fields. Some researchers are now investigating quantitative characterisation to evaluate the quality of the results produce using GIS. In recent years, mathematical models have been implemented in GIS, with the aim of representing several environmental features, to aid interpretation and providing better understanding so more informed decisions can be made.

However, it is necessary to understand the uncertainties associated to spatial data that occur in mathematical modelling within a GIS environment. The uncertainties generated during mathematical modelling are also in the final products (e.g. maps and charts), which are used in the decision-making processes. If such uncertainties are not taken into consideration, the quality of the information contained in the final product might affect all the interpretations based on that product. It is, therefore, important to understand the uncertainties in mathematical models and to estimate the uncertainties associated with spatial attributes contained in final products. Estimating the uncertainties of spatial attributes has been reported in the literature (e.g. Arbia 1993, Burrough 1992, Goodchild & Guoging 1992, Goodchild 1993, Heuvelink & Stein 1989; Heuvelink, Burrough & Stein 1993, and Heuvelink 1998), as it adds high quality information to data and final products made available by GIS.

In this context, kriging geostatistical method may be used to build probabilistic models of the uncertainties related to the values of the attributes. Linear kriging may be used to infer the parameters of a Gaussian probabilistic model (mean and variance), while indicator kriging, a non-linear kriging method, is used to build a discreet approximation of a generic probabilistic model. Kriging geostatistical algorithms allow the construction of models of random distribution variables, which are used to infer of attribute's values and to estimate the uncertainties associated to such values (Goovaerts 1997, Deutsch & Journel 1998).

Indicator kriging is characterized by geostatistical non-linear procedures to model the variability of spatial attributes. These procedures are used to infer a discreet approximation to the model of distribution of probability of the attribute, which is used to model uncertainties of its value. Models of uncertainty are obtained directly from the distribution. They are independent from the selected estimator and are related to the variability of the attribute. Indicator kriging estimator is a linear kriging estimator applied on a set of elements which attributes values have been modified according to a non-linear transform. Indicator kriging requires a non-linear transform, called codification by means of indication, which transforms each value of the set into indicator values (Deutsch & Journel 1998).

 

The main advantage of the indicator kriging technique is that it is non parametric (i.e. the data distribution for the variable is not assumed). The distribution function can be estimated, which makes it feasible to determine the uncertainties and infer the attributes value where there are no samples. Indicator kriging is able to model attributes with high spatial variability. Atypical data may be included (Journel, 1983).

In this paper, indicator kriging technique in GIS SPRING 4.0 (Câmara *et al.,* 1996) has been used to generate a probabilistic map of three variables. This map is used during evaluation and as an indicator of the most favourable site for the construction deep foundations in the study area.

# CHARACTERIZATION OF THE STUDIED AREA

## *Location*

The area studied is the urban area of central of São José do Rio Preto (Brazil) with 81.0 km$^2$, located in the western region of the State of São Paulo, between 20°45’52’’ and 20°50’48’’ south, and 49°20’18’’ and 49°25’26’’ west (Figure 1).

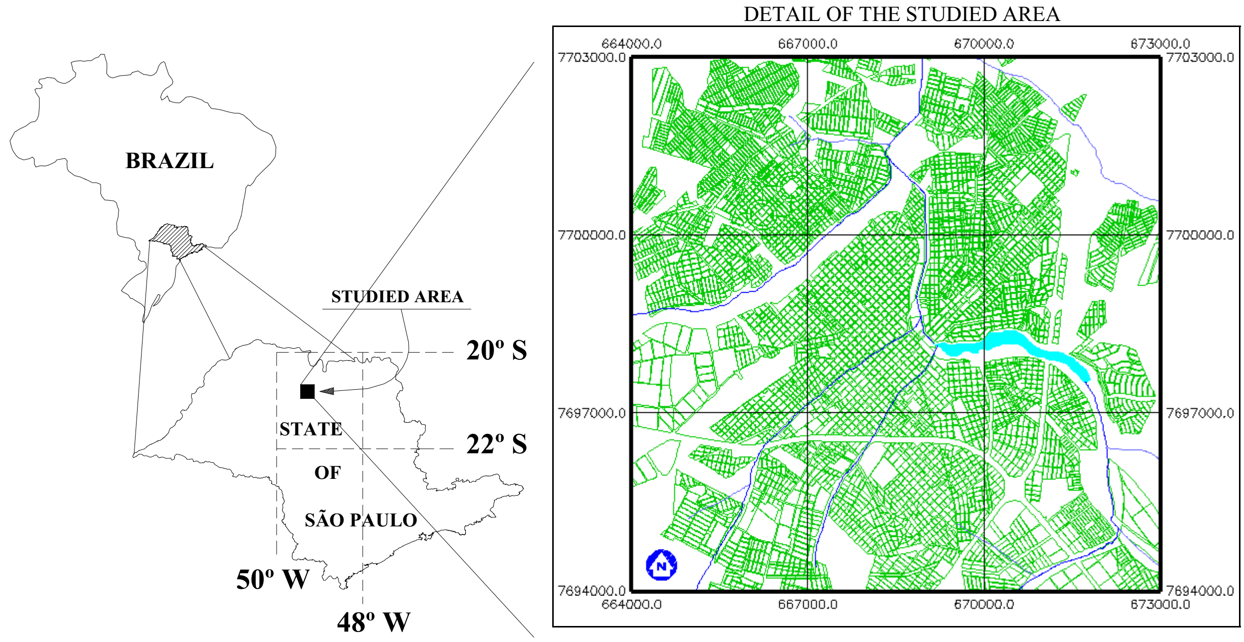


**Figure 1.** Location of the studied area

The area has a humid tropical climate (Arid, 1966), with an annual mean of temperature of about 25.4C. The vegetation cover is 10.2% permanent crops, 18.9% annual crops, 52.9% grassland, 7.9% forest and 9.9% unproductive and reforested lands.

## *Geomorphology*

This area is in the Central Western High Plain, which is between 400 and 700 metres above sea level, has mean gradients generally between 2% and 10% and typically comprises wide and low hills with convex peaks. The mean distance between rivers ranges between 1.75 and 3.75 km.

The studied area has predominantly ‘soft’ topography, with undulate topography, with wide low ridges and about 100 m difference between the highest and lowest levels (470 meters to 570 meters). Long, narrow, low ridges form watersheds and divide the land into the main basins (Barcha, 1980).

## *Geology*

The solid geology in the study area is represented by Bauru Group (Upper Cretaceous), with the Adamantina Formation, which covers most of the area, above, and the Santo Anastácio Formation below. Regionally, Adamantina Formation is between 58 to 140 m thick. It is generally a red and brown fine sandstone and forms benches. The greatest thickness of this formation is found in the ridges between rivers in the western region (Barcha, 1980).

The Santo Anastácio Formation is typically a reddish brown, fine to medium sandstone and is found in lower lying valley bottoms.

At surface, most of the developed area has a thick layer of reddish yellow latosol, whilst the slopes of the main valley and in the north eastern border of the region have reddish yellow latosol with argisol, indicating an intermediate stage of soil development. Gleisols occur along the bottom of the main valleys, and are associated with organic clays of low bearing capacity and near surface water table. Other soil, such as argisols and cambisols, are also present but are of limited extent (Augusto Filho, Ridente Jr. & Alves 1999).

# GEOSTATISTICAL METHODOLOGY

## *General concepts*

Geostatistics models the values of an attribute within a region of the earth surface (e.g. region A) as a random function. For each position $u \in A$, the value of the attribute of a spatial datum is modelled as a random variable Z (u). It means that, for the position *u*, the random variable Z (u) may take several values for the attribute, each value with an associated probability of occurrence. For the *n* sampled positions, $u_\alpha$, with $\alpha$ =1, 2,...n, $z(u_\alpha)$ are considered deterministic, or they may be considered random variables for which the measured values have 100% probability of occurrence. A random variable is characterized by its cumulative distribution function F (u;z), defined as:

$$F(u;z) = \Pr ob\{Z(u) \leq z\} \quad (1)$$

The function of univariate conditioned accumulated distribution, Z (u), is given by:

$$F((u;z)/(n)) = \Pr ob\{Z(u) \leq z/(n)\} \quad (2)$$

A random function (or a random field) is a set of random variables defined over the region *A*. The function of multivariate accumulated distribution, which characterizes the random function composed by *k* random variables, is defined by:

$$F(u_1 .....,u_k ; z_1 ,.....,z_k ) = \Pr ob\{Z(u_1) \leq z_1 ,.....,Z(u_k) \leq z_k \} \quad (3)$$

The function of univariate conditioned accumulated distribution F (u; z|(n)) models the local uncertainty of a random variable. The function of multivariate accumulated distribution models the global uncertainty of a random field. The model for the distribution of a random variable, or a random function, may be determined by two approaches: parametric and non parametric. The former establishes a model of distribution that is determined by a limited set of values (e.g. model of normal distribution, which is determined by values of mean and variance of the distribution). The latter does not assume any probabilistic distribution model. Hence, it is not determined by a limited set of values. In this case the function of distribution of probability is obtained by a set of estimated values, which represents a discreet approximation to the model of distribution. Once they are determined, probabilistic models of distribution are used to estimate the values of an attribute where there aren't any, as well as for modelling the uncertainties of the values of the attribute. Many texts on geostatistics include this method but for specific examples see Journel & Huijbregts (1978), Journel (1989) and Isaaks & Srivastava (1989).

## *Experimental variogram*

The variogram forms the basis for weighted interpolation used in kriging. It allows the quantitative representation of variation of a parameter in space. Generally, the variogram is a function of increment h (the further are the samples, the more different are their values – Huijbregts 1975). The first step in geostatistical estimation is the construction of an experimental variogram, which is determined from the sample set, by:

$$\gamma(h) = \frac{1}{2 \cdot N(h)} \sum_{(\alpha,\beta)/h_{\alpha\beta}=h} \left(z(u_\alpha) - z(u_\beta)\right)^2 \quad (4)$$

where $\gamma$(h) is the variogram function, N (h) is the number of pairs of samples selected by *h* vector; $\alpha$ and $\beta$ determine the position of the random variable in the sample space.

The following graphics express the spatial behaviour of the random variable, besides showing dimension of the influence zone and occasional aspects of anisotropy. It is expected that observations (data, sampling) geographically close to each other are more alike to each other than those at greater distances. Thus, it is expected that $\gamma$(h) increases with distance *h*. Figure 2 shows a fit of the theoretical variogram to the experimental variogram with features close to the ideal.

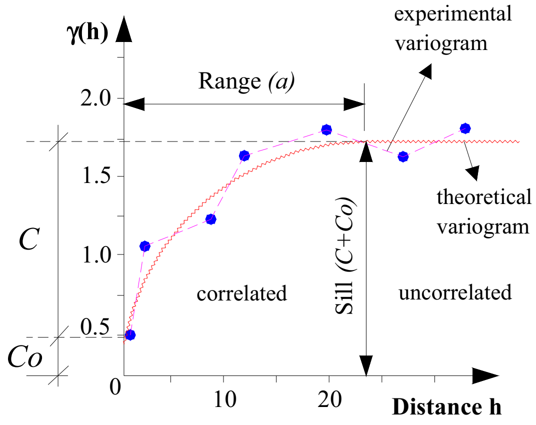


**Figure 2.** Theoretical variogram fitted to an experimental variogram

Figure 2 shows the parameters of the variogram:

Range (*a*): indicates the distance between locations beyond which observations appear independent, i.e. variance no longer increases.

Sill ($C + C_0$): is the value of the variation chart, which corresponds to its range (*a*). From this point forward, one assumes that there is no more spatial dependence between the samples, as variance of the differences between the pairs of samples becomes constant with distance.

Nugget ($C_0$): describes the unexplained variance of the variable when modelled for very short distances between samples, i.e. shorter than the sampling distance.

## *Theoretical variogram*

After several iterations of the experimental variogram in several directions, it is necessary to adjust the mathematical model to represent the variable as realistically as possible. It is important that the mathematical model represents the trend of the variogram with relation to distance *h*. Estimates obtained from kriging will then be more precise and reliable. Several theoretical models are used to adjust the variogram including exponential, spherical and Gaussian models (Figure 3).

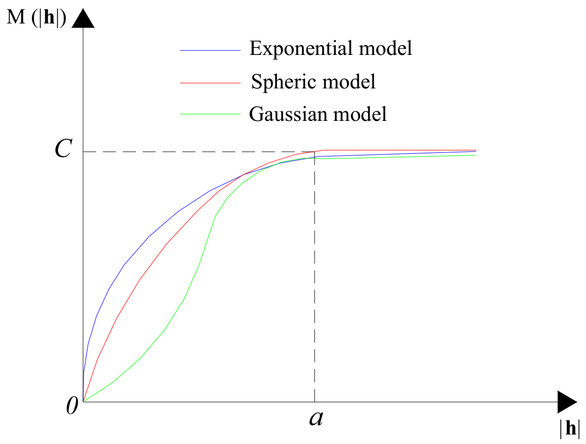


**Figure 3.** Theoretical variogram of adjustment

The adjustment procedure is iterative. In this process, the user makes a primary adjustment and checks if the theoretical model is suitable. In practice, experimental variogram present values of nugget effect ($C_0$) greater than zero. The data here is best represented by the exponential model, which equation is given by:

$$\gamma(h) = 0 \ ; \ \text{for } h = 0$$

$$\gamma(h) = C_0 + C \cdot \left[ 1 - e^{\left(\frac{-h}{a}\right)} \right] \ ; \ \text{for } h \neq 0 \qquad (5)$$

where $C_0$, *C* and *a* are the basic parameters of the adjustment model. This mathematical model reaches the sill asymptotically. Practical range is defined as the distance for which the value of the model is 95% of the sill.

## *Indicator kriging estimates*

As has already been stated, the function of univariate conditioned accumulated distribution, F (u; z|(n)), models the uncertainties of the values of Z (u) in positions *u,* which have not been sampled. An approximation of the function of conditioned accumulated distribution may be obtained by the indicator kriging procedure. This kriging is an estimation technique with the same basis of the linear kriging, but it is applied to attributes with non-Gaussian distribution, which are transformed according to a non-linear mapping, codification by means of indication. Applied on a set of numeric sample values, $Z(u=u_a)$, codification by means of indication generates, for a cutting value $z_k$, a sample set by means of indication $I(u=u_a; z_k)$, of the kind:

$$I(u;z_k) = \begin{cases} 1 \xrightarrow{if} Z(u) \leq z_k \\ 0 \xrightarrow{if} Z(u) > z_k \end{cases}$$

Codification by means of indication is applied on the whole sample set, creating, for each cutting value, a sample set by means of indication for which sample values are transformed either in 0 or 1. The *K* cutting values $z_k$, k = 1, 2,..., *K* are defined according to the number of samples. Definition of a variation computing model depends on the existence of a minimum distribution of zeros (0) and ones (1) within the set of samples codified by means of indication. It can be shown that the best variation computing is obtained for a cutting value equal to the median computed from the sample set (Isaaks & Srivastava 1989). Thus, it is possible to use a single cutting value, equal to or close to the median, in order to generate a single codification by means of indication, known as codification by means of indication by median. Use of a single cutting value reduces computation time, however, this may be restricting if the use of the variation recording model, close to the median, is too different from those obtained by other cutting values (Isaaks & Srivastava 1989).

In the same way one uses the sample set of the attribute to infer values for the numeric random variable Z (u), the set of sample data by means of indication is used to infer values for the random variables by means of indication $I(u;z_k)$ , with $u \neq u_a$. The conditional expectancy of the numeric random variable obtained by means of indication $I(u;z_k)$ is calculated by:

$$E\{I(u;z_k)/(n)\} = 1 \cdot \Pr ob\{I(u;z_k) = 1/(n)\} + 0 \cdot \Pr ob\{I(u;z_k) = 0/(n)\} =$$

$$= 1 \cdot \Pr ob\{I(u;z_k) = 1/(n)\} = F^*(u;z_k/(n)) \qquad (6)$$

This equation presents an extremely important result concerning the inference of the probability distribution of a random variable. It means that the conditional expectancy of I (u ; $z_k$) yields, for the cutting value $z = z_k$, an estimate of the function of conditioned accumulated distribution, F* (u; $z_k$|(n)), for numeric attributes. Conditional expectancy E {I (u; $z_k$|(n))} may be estimated by algorithms of kriging, which employ numeric random variables, codified by means of indication. Kriging by simple indicator is simple linear kriging applied to a sample set codified by means of indication at the cutting values $z = z_k$, and it has the following formulation:

$$F_S^*[((u;z_k)/(n))] = \sum_{\alpha=1}^{n(u)} \lambda_{S\alpha}(u;z_k) i(u_\alpha;z_k) + \left[1 - \sum_{\alpha=1}^{n(u)} \lambda_{S\alpha}(u;z_k)\right] F^*(z_k) \qquad (7)$$

where F* ($z_k$) is the average of the random function of the stationary region, and the weights $\lambda_{s\alpha}$ (u ; $z_k$) are determined in order to minimize the variance of the estimative error. Considering the sum of the pondering weights equal to 1, a simpler version of kriging by means of indication is obtained. It is the kriging by means of ordinary indication, which estimative expression comes to be:

$$F_o^*[(u;z_k/(n))] = \sum_{\alpha=1}^{n(u)} \lambda_{o\alpha}(u;z_k) i(u;z_k) \qquad (8)$$

The weights $\lambda_{s\alpha}$ (u ; $z_k$) are obtained by solving the following system of equations of kriging by means of ordinary indication:

$$\sum_{\beta=1}^{n(u)} \lambda_{o\beta}(u;z_k) C_1(h_{\alpha\beta};z_k) + \phi(u;z_k) = C_1(h_\alpha;z_k) \qquad \forall \alpha = 1,2,......,n(u) \qquad (9)$$

$$\sum_{\beta=1}^{n(u)} \lambda_{o\beta}(u;z_k) = 1 \qquad (10)$$

where $\phi$ (u ; $z_k$) is Lagrange's multiplier, $h_{\alpha\beta}$ is the separation vector (which is defined by the positions $u_\alpha$ and $u_\beta$), $h_\alpha$ is the vector defined between the positions $u_\alpha$ and $u_\beta$, $C_1$ ($h_{\alpha\beta}$ ; $z_k$) is the self-covariance defined by $h_{\alpha\beta}$, and $C_1$ ($h_\alpha$ ; $z_k$) is the self-covariance defined by $h_\alpha$ in $z = z_k$. The self-covariances are determined by the theoretical variation computing model defined by set I for $z = z_k$.

Kriging by means of indicator, either simple or ordinary, yields, for each cutting value $z_k$, an estimate, which is the best least square estimate of the conditional expectancy of the random variable I (u ; $z_k$). By using this property, it is possible to compute the estimates for the values of the function of accumulated distribution of Z (u) for several values of z = $z_k$ belonging to the domain of Z (u). The set of values of the functions of accumulated distribution, estimated at the cutting values, is considered a discrete approximation of the real function of accumulated distribution of Z (u). The greater the number of cutting values, the better the approximation.

The approximation is complemented by the definition of an adjustment function for the distribution, which should be used to infer the cumulative distribution function for values different to the cutting values. A linear adjustment is the simplest to be defined, but functions of higher degree may be used. Thus, indicator kriging is defined as non-parametric, as it does not assume the probability distribution of the random variable. On the contrary, it allows the construction of a discrete approximation for the function of accumulated distribution of Z (u). Discrete probability values may be used directly to estimate typical statistical values of the distribution, such as: mean, variance, median, moments and others.

## GEOTECHNICAL INFORMATION

A total of 1237 standard penetration test N-values (SPT N values) from a regional study carried out during 1991 to 2000 (Mendes & Lorandi, 2003a) (Figure 4) were evaluated and 241 SPT N representative values were selected based on local topographic and geological conditions. Analyses indicated that there was little vertical variability for relatively high distances (greater than 500 meters). As the distance between boreholes is less than this distance, it is assumed that the geological and topographic factors are spatially well represented. Information on the position of the water table was obtained from an index of resistance to penetration (Mendes & Lorandi, 2003b). Particle size analysis (sand, silt and clay) was carried out on samples from depths of 6.0, 8.5, 11.0 and 13.5 meters.

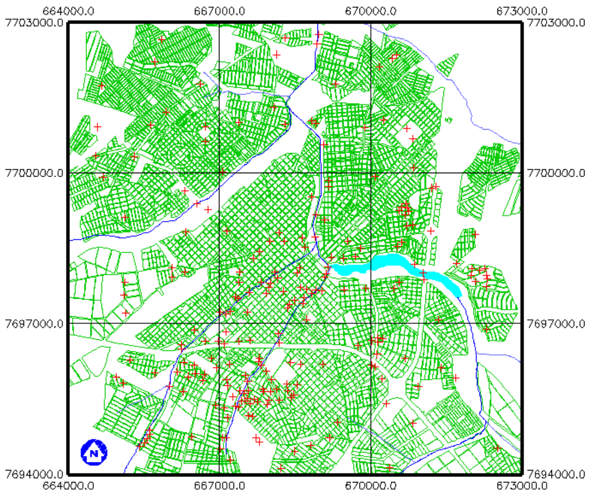


**Figure 4.** Distribution of SPT probing boreholes over the studied area.

All the field and laboratory data with grid reference and depth were input into a spreadsheet (Excel of *Microsoft Office*™). The data were loaded into GIS – SPRING 4.0 environment (Camara *et al.* 1996) and geo-referred in UTM (Universal Transverse Mercator) coordinates still within this system. Figure 4 depicts the spatial distribution of SPT boreholes used in the study, along with a representation of the urban network.

Next phase involved the selection of relevant foundation type and depth and the acceptable range of variables for the variables used for a probabilistic map. The selected foundation is one that is most frequently used in the study area, which has a base at 8.5 meters bellow the surface (Figure 5).

The penetration resistance index was defined as a function of the $N_{SPT}$ values from the zone beneath the foundation including the pressure bulb, and the allowed variability for the foundations. Estimates of allowable pressure at the base of the foundation using Brazilian professional practice form theoretical calculations or empirical correlations of the penetration resistance index. The empirical expression selected to calculate the values of allowable pressure to support the foundation is given below and may be used for any sort of soil (Albieiro & Cintra 1996).

$$\sigma_{ad} = 20 \cdot N_{72} + \sigma'_{vb} \quad (11)$$

where $\sigma_{ad}$ is the allowable pressure at the base of the foundation (in kPa), $\sigma'_{vb}$ is the effective vertical geostatic stress (in kPa), and $N_{72}$ is the average value of the penetration resistance index of the SPT sampler, obtained according to Brazilian standards, within the active zone of the pressure bulb of the foundation, considered to be approximately 1.5 x the width of the base of the foundation (see Figure 5).

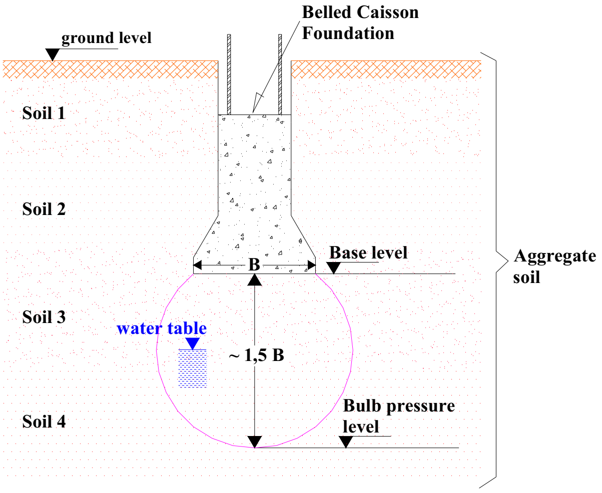


**Figure 5.** Structural element of the adopted foundation

At the analysis depth (8.5 metres) the effective vertical geostatic stress is considered to be approximately 130 kPa (for an average $\gamma$ of 15.0 kN/m$^3$). Hence, for an allowable pressure varying from 200 to 500 kPa, expression 11 yields a range of SPT values varying between 5 and 20 blows. The values used correspond to allowable pressures of soils described as normally dense or dense sands (as recommended by Brazilian Association for Technical Normalization – ABNT, 1996), which frequently occur within the studied region.

In order to identify where collapsible soils occur in the studied area, several tests were carried out on undisturbed soil samples. The aim is to avoid unacceptable rapid settlement of structures caused by abrupt collapse of soil. Collapsibility was evaluated using collapse potential (CP) calculated from consolidation tests followed by flooding. Collapse potential was defined from the equation proposed by Jennings & Knight (1957):

$$CP = \frac{\Delta e_c}{(1 + e_0)} \tag{12}$$

where $\Delta e_c$ is the change of void ratio caused by the flooding, and $e_0$ is the initial void ratio of the soil sample. Values of CP obtained during the loading stage of 50 kPa, along with some physical indexes are presented in Table 1.

**Table 1.** Values of Collapse Potential (CP) and of physical indexes of the tested samples

| Sample | Ground level (m) | Depth (m) | w (%) | LL (%) | IP (%) | S (%) | $e_0$ | $\gamma_d$ (kN/m³) | CP (%) |
|---|---|---|---|---|---|---|---|---|---|
| P1A | 497.8 | 1.0 | 5.8 | 31.4 | 11.2 | 16.6 | 0.96 | 13.9 | 0.2 |
| P1B | | 2.0 | 10.4 | 36.0 | 15.9 | 27.1 | 1.06 | 13.4 | 2.0 |
| P1C | | 3.0 | 8.2 | 37.0 | 16.0 | 22.8 | 0.98 | 13.9 | 0.7 |
| P2A | 545.6 | 1.0 | 5.5 | 30.0 | 10.6 | 15.5 | 0.97 | 14.0 | 1.8 |
| P2B | | 2.0 | 6.4 | 32.7 | 13.2 | 17.3 | 1.02 | 13.6 | 2.9 |
| P2C | | 3.0 | 5.9 | 35.5 | 11.9 | 15.5 | 1.04 | 13.5 | 5.7 |
| P3A | 529.9 | 1.0 | 5.8 | 30.0 | 12.3 | 17.0 | 0.93 | 14.0 | 7.3 |
| P3B | | 2.0 | 5.0 | 28.7 | 10.1 | 13.6 | 1.00 | 13.6 | 9.7 |
| P3C | | 3.0 | 4.3 | 28.0 | 9.2 | 11.2 | 1.07 | 13.5 | 3.5 |
| P4A | 490.7 | 1.0 | 5.9 | 29.0 | 9.3 | 18.1 | 0.90 | 14.4 | 2.3 |
| P4B | | 2.0 | 8.5 | 32.3 | 13.1 | 22.5 | 1.03 | 13.4 | 2.1 |
| P4C | | 3.0 | 5.9 | 32.0 | 11.0 | 16.4 | 1.00 | 13.8 | 1.1 |
| P4D | | 7.0 | 11.7 | 43.5 | 14.2 | 35.8 | 0.90 | 14.6 | 1.8 |
| P5A | 501.4 | 1.0 | 8.7 | 30.7 | 12.9 | 22.4 | 1.06 | 13.2 | 3.2 |
| P5B | | 2.0 | 7.6 | 31.0 | 13.3 | 26.1 | 0.81 | 15.2 | 4.0 |
| P5C | | 3.0 | 7.6 | 32.7 | 12.3 | 21.2 | 0.99 | 13.78 | 6.8 |

where w is water content, LL is liquid limit, IP is plasticity index, S is saturation, $e_0$ is initial voids ratio, $\gamma_d$ is dry density and CP is collapse potential.

Analysing the values of collapse potential (CP) from Table 1 show that a majority of samples had values greater than 2%, and some samples have values even greater than 5%, which suggests the presence of layers of collapsible soil in the studied area, according to the adopted criteria. Given the limited spatial representation of the collapse potential information of the studied area, a choice was made to represent this variable of the soil mass from the analysis of SPT data.

According to Ferreira *et al.* (1989), the presence of collapsible soils may be indicated by: low $N_{SPT}$ values (< 5 blows), discontinuous grain size distribution (absence of silt fraction), low saturation (<60%) and high porosity (> 40%). Regions of collapsible soil at a depth below the foundation (8.5m) were identified as having $N_{SPT}$ values less than 5 blows and silt content of less than 5%.

Water table depth is also important as if it is above the base of the foundations it may affect the construction of the foundation.

# RESULTS

## *Basic statistical parameters*

Table 2 and Figure 6 show the basic statistical parameters and histograms for the variables considered in the study.

**Table 2.** Basic statistical parameters of the analysed variables

| Statistics parameters | Water table | $N_{SPT}$ | Silt fraction |
|---|---|---|---|
| Number of samples | 241 | 241 | 241 |
| Mean | 8,8 | 20,6 | 19,4 |
| Standard deviation | 6,1 | 19,3 | 11,0 |
| Coefficient of variation | 0,69 | 0,94 | 0,57 |
| Minimum | 0,2 | 2,0 | 5,0 |
| 25 th | 3,7 | 8,0 | 10,0 |
| 50 th | 7,7 | 12,0 | 15,0 |
| 75 th | 13,6 | 20,0 | 30,0 |
| Maximum | 25,0 | 60,0 | 50,0 |

The cutting values for each variable were selected as a function of the relative frequency distribution (Figure 6). Thus, the cutting values used in the indicator kriging were for water table 4, 8, 11 and 14 meters; penetration resistance index ($N_{SPT}$) - 8, 12, 16 and 20 blows and silt fraction percentage of 15, 20, 25 and 30.

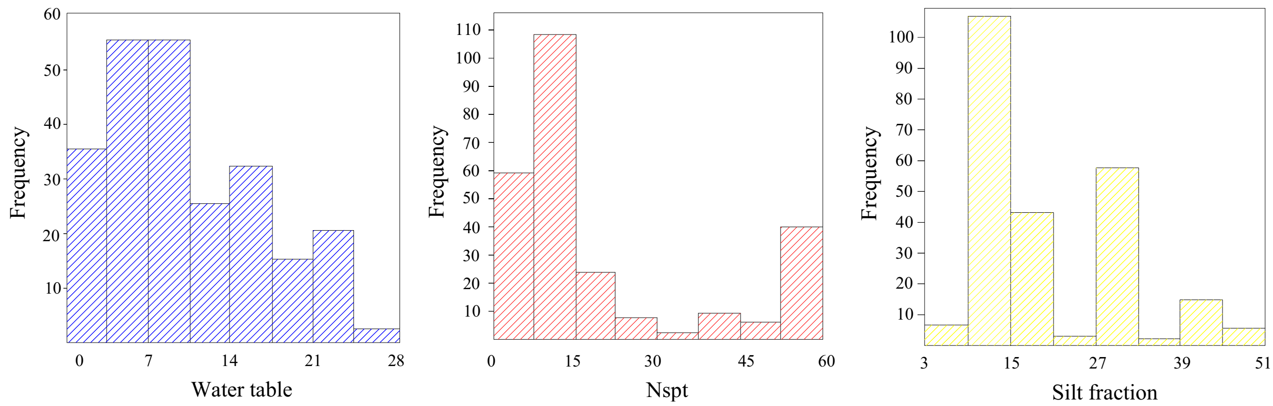


**Figure 6.** Histograms of the three variables: water table, $N_{SPT}$ and silt fraction

## *Variation computing analysis and indicator kriging estimates*

After the definition of the levels of cutting for the three variables, analyses of spatial variability were carried out in directions 0, 45, 90 and 135, which produced the mathematical variation model, shown on Figure 7. From the analyses, it was also possible to verify the isotropic spatial behaviour for all variables. For the estimating processes, indicator variation charts close to the average value of each variable. A summary of the results shown in Figure 7 are summarised in Table 3.

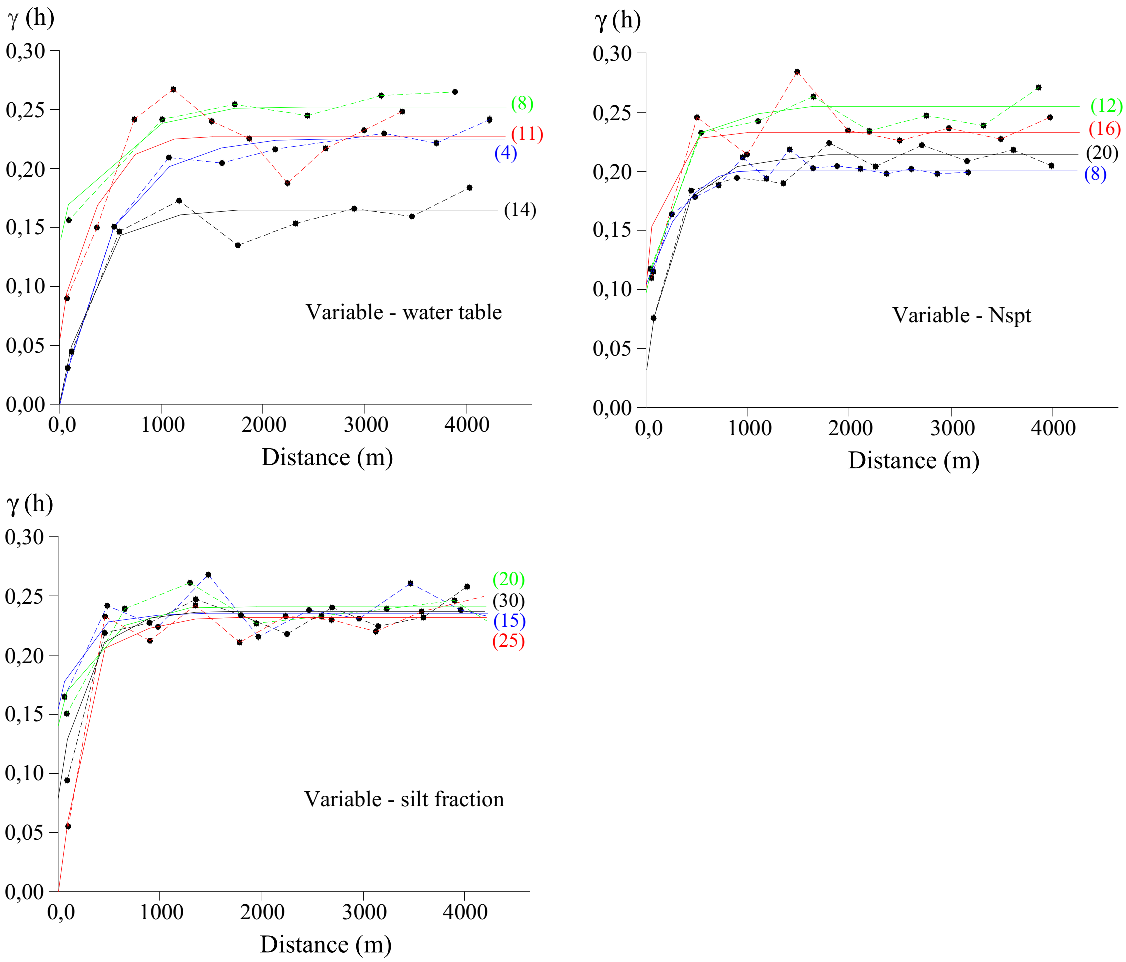


**Figure 7.** Indicator variogram for all analyzed cutting levels.

The values of variables and the respective uncertainties were estimated from the modulus of kriging by means of indication (implemented in GIS –SPRING 4.0, which is based on sub-routine "ik3d" of GSLIB / Deutsch & Journel 1998). It has been done for 25 m x 25 m regular grids, distributed over an 81.0 km$^2$ area. Values for each point of the grid were estimated by average values, which were inferred from accumulated distribution functions.

Indicative maps of the three variables were generated from the probabilistic distribution models from indictor kriging, for uncertainty values equal to the first moment and for a 90% probability of the occurrence of the variable values.

**Table 3.** Values obtained from modelled variogram

| **Variable – Water table** | | | | | **Variable - $N_{SPT}$** | | | | | **Variable – Silt fraction** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cuttof | Mod.* | $C_0$ | C | (a) | Cuttof | Mod.* | $C_0$ | C | (a) | Cuttof | Mod.* | $C_0$ | C | (a) |
| (4) | Exp. | 0,005 | 0,17 | 1700 | (8) | Exp. | 0,104 | 0,08 | 1100 | (15) | Exp. | 0,155 | 0,08 | 1000 |
| (8) | Exp. | 0,140 | 0,11 | 2100 | (12) | Exp. | 0,090 | 0,15 | 1700 | (20) | Exp. | 0,137 | 0,11 | 1100 |
| (11) | Exp. | 0,040 | 0,17 | 1300 | (16) | Exp | 0,102 | 0,12 | 1100 | (25) | Exp | 0,005 | 0,23 | 1500 |
| (14) | Exp. | 0,006 | 0,16 | 1600 | (20) | Exp. | 0,030 | 0,17 | 1800 | (30) | Exp. | 0,085 | 0,16 | 1200 |

* Mathematical model employed in fitting the indicative variogram (e.g. Exp. – Exponential model)

For the generation of probabilistic maps, indicative maps of the variables were classified according to intervals of probability of occurrence. Figure 8 shows the probabilistic maps for different ranges of the variables (i.e. water table, penetration resistance index - $N_{SPT}$ - and silt fraction), for a 90% global probability of occurrence.

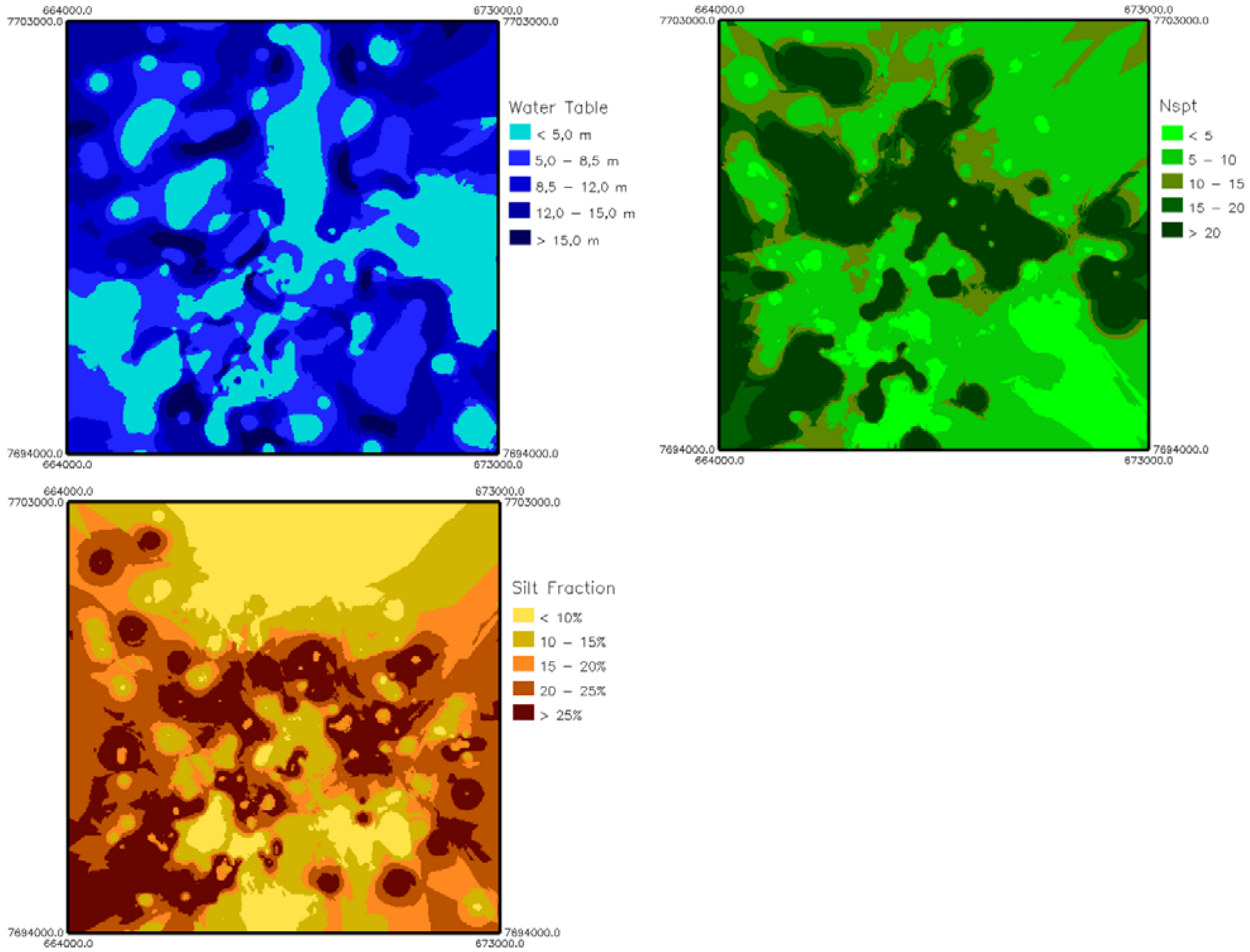


**Figure 8.** Probabilistic maps for the occurrence of variables values for each variable.

## *Probabilistic map for favourable foundation areas*

In order to identify areas that are favourable for the construction of deep foundations, a probabilistic map was constructed by combining the three probabilistic maps. The range of variable values used to indicate favourable foundation conditions are shown in Table 5.

**Table 5.** Range of variable values and favourability conditions

| Favourable Conditions | | | Unfavourable Conditions | |
|---|---|---|---|---|
| Water Table | Allowable Soil Pressure | Collapsibility | Water Table | Collapsibility |
| > 8,5 m | < 200 kPa<br>200 < $N_{SPT}$ < 300 kPa<br>300 < $N_{SPT}$ < 400 kPa<br>400 < $N_{SPT}$ < 500 kPa<br>> 500 kPa | $N_{SPT}$ > 5 blows<br>And<br>Silt fraction > 5% | < 8,5 m | $N_{SPT}$ < 5 blows<br>And<br>Silt fraction < 5% |

Unfavourable areas, relate to the occurrence of collapsible soils, the combination of probabilistic maps of the penetration resistance index ($N_{SPT}$ < 5 blows) and of percentage of silt fraction (< 5%) and water level. Figure 9 shows that the water level is the main constraining factor for the construction of the type of foundations considered in this study area (red area). Areas favourable for the construction of the foundations are those where all three variables indicate good ground conditions and are represented by values of allowable pressure (Figure 9).

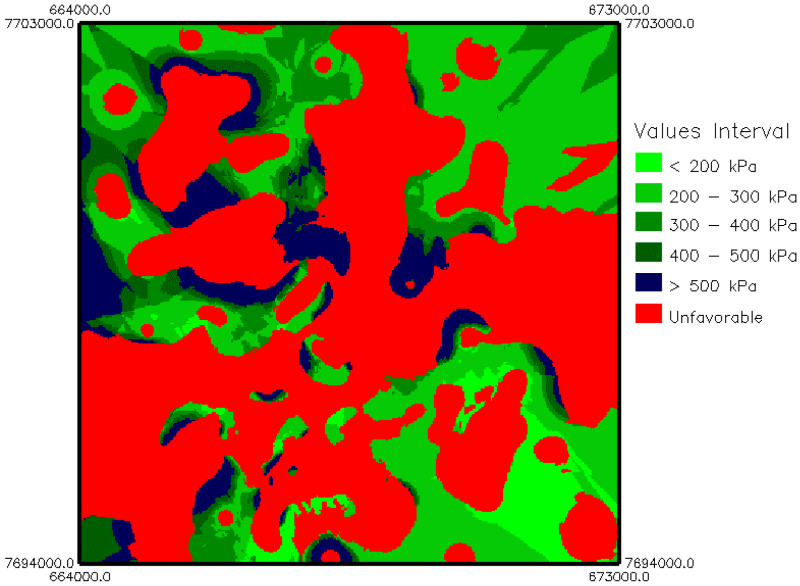


**Figure 9.** Probabilistic map showing the best and worst areas for the construction of the foundations.

Thus, in probabilistic terms, the best areas to construct the proposed foundations are in where the water table occurs at depths of greater than 8.5 meters, and for which the range of values of allowable pressure of the soil fits minimal design requirements. Construction of the foundations where water level is less than 8.5 m from the surface may also not be economically viable.

However, it must be highlighted that probabilistic maps present fields of variations of uncertainties, which are proportional to spatial behaviour of the analysed variables. Uncertainties increase in areas where a parameter is very variable. On the other hand, in areas where a parameter has little variation, or even no variation, uncertainty values are smaller.

The results presented show that estimation methods based on indicator kriging may be used to produce thematic maps, with information on uncertainties, from a sample set of numeric variables. In this case, the disadvantage associated with representing variables by homogeneous regions may be compensated by analysing probabilistic maps made during the decision making process of the final map generated by the GIS. The maps can be used to guide planning decisions and more focused geotechnical ground investigation.

# CONCLUSIONS

Geographical Information Systems (GIS) with indicator geo-statistical resources (such as the geo-statistics modulus implemented in SIG-SPRING 4.0) are a very powerful tool, as they allow the estimation of uncertainties associated with the variability of parameters, making possible the qualification of maps generated using these tools.

Probabilistic maps from indicator kriging can be used to show where the most favourable areas for constructing foundations at a specific depth (in terms of range of values). Similarly, other foundation depths could be studied using the same variables.

However, it must be stressed that the probabilistic map may be used to guide local geotechnical investigation, which may then be aimed at specific problems. In some cases, resulting probabilistic maps may eliminate some parts of the local investigational process, as some sites may be unsuitable for the proposed foundations.

**Acknowledgements:** Authors wish to thank CAPES (Coordination for the Perfection of Graduated Personnel Foundation) for the concession of an MSc scholarship to the first author.

**Corresponding author:** Prof. Rodolfo Moreda Mendes, Geological Institute, Secretariat of Environment, Av. Miguel Stefano nr. 3900, São Paulo, Zip Code: 04301-903, Brazil, Tel:+55(11)5077-2160. Email: rodolfo.mendes@igeologico.sp.gov.br.

# REFERENCES

ABNT 1996. NBR 6122 – Project and Execution of Foundation, Rio de Janeiro, 33p (in Portuguese).

ALBIERO, J.H. & CINTRA, J.C.A. 1996. Analysis and projects of deep foundations: belled caisson. *In:* HACHICH, W.C., FALCONI, F.F., SAES, J.L., FROTA, R.G.Q., CARVALHO, C.S. & NIYAMA, S. (eds.) *Foundations: Theory and Practice*. PINI, São Paulo, 302-327 (in Portuguese).

ARBIA, G. 1993. The use of SIG in spatial statistical surveys. *International Statistical Review*, **61**(2), 339-359.

ARID, F.M. 1966. *The Bauru Formation in the western region of the State of Sao Paulo.* Thesis, Sao Paulo State University (in Portuguese).

AUGUSTO FILHO, O., RIDENTE JR., J.L. & ALVES, C.F.C. 1999. Geotechnical compartment of the urban area of São José do Rio Preto (Brazil) by spatial generalization of percussion drilling data. *In: Proceedings of the 9th Brazilian Congress of Engineering Geology*, Águas de São Pedro. ABGE, São Paulo, 1-14 (in Portuguese).

BARCHA, S.F. 1980. Geological and hydrogeological aspects of the Bauru Formation in the western region of the State of São Paulo. Thesis, São Paulo State University (in Portuguese).

BURROUGH, P.A. 1992. Development of intelligent geographical information systems. *International Journal of Geographical Information Systems*, **6**(1), 1-11.

CAMARA NETO, G., SOUZA, R.C.M., FREITAS, U.M. & GARRIDO, J. 1996. SPRING: Integrating remote sensing and GIS by object-oriented data modeling. *Computers & Graphics*, **20**(3), 395-403.

DEUTSCH, C.V. & JOURNEL, A.G. 1998. *GSLIB Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 369p.

FERREIRA, R.C.; MONTEIRO, L.B.; PERES, J.E.E.; BENVENUTO, C. 1989. Some aspects on the behavior of Brazilian collapsible soils. *In: Proceedings of the 12th International Conference on Soil Mechanics and Foundation Engineering (ICSMFE), Rio de Janeiro*, suppl. vol., Balkema, Rotterdam, 117-120.

GOODCHILD, M.F. 1993. Data models and data quality: problems and prospects. *In:* GOODCHILD, M. F., PARKS, B. O. & STEYAERT, L. T (eds.) *Environmental Modeling with GIS*. Oxford University Press, New York, 94-103.

GOODCHILD, M.F. & GUOGING, S. 1992. Development and test of an error model for categorical data. *International Journal of Geographical Information Systems*, **6**(2), 87-104.

GOOVAERTS, P. 1997. *Geostatistics for natural resources evaluation*. Oxford University Press, New York, 481p.

HEUVELINK, G.B.M. 1998. Error propagation in environmental modeling with GIS. Taylor & Francis (Inc), New York, 127p.

HEUVELINK, G.B.M., BURROUGH, P.A. & STEIN, P.A. 1989. Propagation of errors in spatial modeling with GIS. *International Journal of Geographical Information System*, **3**, 303-322.

HEUVELINK, G.B.M. & BURROUGH, P.A. 1993. Error propagation in cartographic modeling using Boolean logic and continuous classification. *International Journal of Geographical Information Systems*, **7**(3), 231-246.

HUIJBREGTS, C.J. 1975. Regionalized variables and quantitative analysis of spatial data. *In:* DAVIS, J.C. & Mc. CULLAGH, M.J. (eds.) *Display and Analysis of Spatial Data*. John Wiley, New York, 38-53.

ISAAKS, E.H. & SRIVASTAVA, R.M. 1989. *An introduction to applied geostatistics*. Oxford University Press, New York, 561p.

JENNINGS, J.E. & KNIGHT, K. 1957. The additional settlement of foundations due to a collapse of structure of sand subsoil on wetting. *In: Proceedings of the 4th International Conference on Soil Mechanics and Foundation Engineering*, London, vol.1, Balkema, Rotterdam, 316-319.

JOURNEL, A.G. 1983. Nonparametric estimation of spatial distributions**.** *Mathematical Geology*, **15**(3), 445-468.

JOURNEL, A.G. 1989. Fundamentals of geostatistics in five lessons. In: Short Course in Geology, Vol. 8. American Geophysical Union, Washington DC, 40p.

JOURNEL, A.G. & HUIJBREGTS, C.J. 1978. *Mining geostatistics.* Academic Press, London, 600p.

MENDES, R.M. & LORANDI, R. 2003a. Application of geostatistics to estimate the unit points and skin resistances as guidelines in deep foundation projects. *Geotecnia-Geotechnic Portuguese Society Journal*, **99**, 5-32 (in Portuguese).

MENDES, R.M. & LORANDI, R. 2003b. Study of SPT behavior of penetration resistence under influence of groundwater level fluctuation, using geostatistcal methology. RMZ – Materials and Geoenvironment, 50 (1), 225-228.