APPLICATION OF KRIGING WITH TREND TO SOIL CHARACTERIZATION IN TOP HOLE DESIGN

^a Silva, P. H. M. ¹; ^a Várady Filho, C. A. F.; ^a Santos, J. P. L.; ^a Lima Júnior, E. T.; ^b Dias, R.; ^b Cutrim, F. S.

^a Federal University of Alagoas, Scientific Computation and Visualization Laboratory, Maceió – AL, Brazil

^b PETROBRAS

Received: 13.05.2020 / Revised: 18.05.2020 / Accepted: 25.05.2020 / Published on line: 21.07.2020

ABSTRACT

In top hole designs, geotechnical characterization is fundamental in the study of the axial load capacity of the soil and its implications on the design of the conductor casing. The present work estimates soil parameters from CPTu (piezocone penetration test) data using geostatistical methods, with evaluation of trends, variograms, and kriging. Results suggest that universal and residual kriging are good tools for soil characterization in top hole section design, considering their capacity to support spatiality of the phenomenon and the mapping of the uncertainties involved.

KEYWORDS

kriging; geotechnical characterization; axial load capacity; soil; oil well projects

¹ To whom all correspondence should be addressed.

Address: Federal University of Alagoas, Computation and Visualization Laboratory, Campus A.C. Simões, Av. Lourival Melo Mota, s/n, Tabuleiro do Martins - Maceió - AL, Brazil.

ZIP Code: 57072-970 | Phone number: +55 82 99991-7218 | e-mail: pedroengpetroleo@outlook.com

doi:10.5419/bjpg2020-0009

1. INTRODUCTION

The challenge of production in deep and ultradeep waters requires greater complexity and robustness in oil well projects. In the first stage of the well construction, research on spatial distribution of soil parameters is of paramount importance. Another major interest falls on the geotechnical characterization of soil to assess the load capacity, following **API-RP-2A regulation (API, 2011)**.

According to **Lacasse et al. (2007)**, soils have variable properties due to its formation process, which generates uncertainties in their mechanical characteristics. Several strategies have been proposed to develop numerical models to infer soil parameters in locations of unknown interest considering spatial continuity, associated with the natural phenomenon of its formation, as seen in **DNV-RP-C207 (DNV, 2012)**.

Classical statistics, although assessing good measurements, fail to describe spatial information associated with the values. In this aspect, geostatistics appears as an option to address this gap and, by using the concept of variogram, it can describe the influence of a given phenomenon through space. Its analysis allows identifying, qualifying, and understanding spatial variation of the phenomenon from apparently random and independent sample data.

The geostatistical predictors in Kriging method use the information of the sampled neighborhood to incorporate spatial dependence of a stochastic process, improving the analysis without generating bias and optimizing variance. Literature presents several types of kriging with emphasis in ordinary kriging, which considers the average floating throughout the area, and kriging with trend, for when a trend is modeled in the data.

In this sense, the present work aims to apply two kriging methods with trend evaluation (Universal and Residual Kriging) in nine sample sets of undrained shear strength (S_u) for the generation of topological surfaces that represent the continuous distribution of this parameter in offshore soils. We also compare universal and residual kriging with ordinary kriging to assess differences in behavior. Cross validation is used to evaluate residues obtained for each of the three techniques.

2. MATERIALS AND METHODS

The calculation of soil parameters may, instead of considering data from a single CPT test, use several test results correlated with their respective geolocations. Usage of the kriging technique will enable unbiased estimates with the lowest variance.

According to **Matheron (1963)**, the starting point for the development of geostatistics was the inability of classical statistics to consider the spatial aspect of a phenomenon, which is the most important feature in a geological study.

The basic element of geostatistics is the regionalized variable, since its variation characterizes the spatial phenomenon that originated it. These variables have casual and structured characteristics, i.e. they can assume locally any value according to a probability function and globally they have a structure that can be treated by a spatial function (**Journel & Huijbregts, 1978**).

Regionalized variables are based on the definitions of random process and data stationarity. In a random process, the Z value of a property at a certain point x , $Z(x)$, is included in an infinite range of possibilities, but can be spatially correlated at some scale. In this sense, values of the property z_i obtained for specific points x_i are called a random process of a variable performed at the respective points. As the phenomena of soil formation present natural continuity even with variabilities according to their direction, the concept of stationarity is introduced, which underlies the geostatistical processes and allows the assumption that there is a certain measurable degree of variation in nearby locations.

In general, the premise can be represented as the first and second order moments being zero (intrinsic hypothesis). Equation 1 points to the modelling of the random process in a trending part μ , and ϵ is the natural random variation based on location. Equation 2 defines the first order moment of the same variable at point x to the variable Z .

$$Z(x) = \mu + \epsilon(x), \quad (1)$$

$$E\{Z(x)\} = \mu. \quad (2)$$

Matheron (1963) proposed the minimization of the error, so that its variance should be null. The semi-variogram is defined as:

$$\gamma(h) = \frac{1}{2} E\{[Z(x) - Z(x+h)]^2\}. \quad (3)$$

in which h describes the length domain, so-called lag. The theoretical variogram is only one of the possible tools. In another aspect, the experimental variogram is estimated from data performed in a region and usually is computed using the formula of the Method of Moments, as described by **Matheron (1965)**.

$$\gamma(h) = \frac{1}{2 \cdot m(h)} \sum_{j=1}^{m(h)} \{Z(x_j) - Z(x_j+h)\}^2 \quad (4)$$

with $m(h)$ as the number of internal comparisons (pairs) to the lag. With the increment of h , it is possible to calculate variogram values as a function of distance. The point combinations, then, depend on the lag used and the directions taken.

Another option for variogram construction is the development of the so called robust variograms which are more accurate in the presence of outliers, as treated by **Cressie and Hawkins (1980)** and **Dowd (1984)**.

As variograms create discrete data, fitting methods are required. The fit of a curve describes the main characteristics of the sequence while reducing point fluctuations (or residuals). The curve should represent a mathematical expression capable of describing the random process with the change of lags and ensure positive variances in estimates. In this aspect, the spherical (SPH) model, according to Equation 5, shows better results in several applications in the literature.

$$\gamma = \begin{cases} C_0 + C \cdot \left\{ \frac{3h}{2r} - \frac{1}{2} \cdot \left(\frac{h}{r} \right)^3 \right\}, & 0 \leq |h| \leq r \\ -C_0 + C, & |h| > r \\ 0, & |h| = 0 \end{cases} \quad (5)$$

with C_0 is the nugget effect, $C + C_0$ is the sill, C is the structure variance, and r is the range.

Once the variogram is modelled, estimation by kriging is the next step. The kriging nomenclature is a generic term for a set of fit methods using least squares to provide the best linear estimates without bias (Best Linear Unbiased Estimation -

BLUE), aiming at the lowest possible variance. Ordinary kriging, as established by **Matheron (1963)**, is one of the simplest techniques to implement. It presents good results for several situations but does not take trending into account. Equation 6 presents the system of equations.

$$\begin{cases} \sum_{i=1}^N \lambda_i \gamma(x_i - x_j) + \psi(x_0) = \gamma(x_j - x_0) \\ \sum_{i=1}^N \lambda_i = 1 \end{cases} \quad (6)$$

in which λ_i and ψ_i are the weights calculated by kriging and used to estimate data.

For cases where a trend in the data is detected, the methodology of the kriging technique is different. Trend consists in maintaining a certain behavior, such as mean parameter value or constant/linear gradient, associated to a specific direction. In the case of geological layers, the trend consists in the representation of the drift of the respective layer. **Matheron (1969)** developed the Universal Kriging technique to consider the presence of trends.

It is a method similar to ordinary kriging, applied when the regionalized variable is not stationary, but presents a trend and its residues contain the intrinsic hypothesis, indicating a constant mean and, consequently, a variogram that impedes the precise modeling of the spatial correlation structure of the considered data.

A non-stationary regionalized variable can be considered as consisting of two components:

- (i) The trend, which is the expected mean value of this variable within a certain neighborhood, and which varies systematically;
- (ii) The residual, which is the difference between the actual values and the displacement value caused by the trend.

In these cases, the kriging is performed on the residuals, so that the structure of the variation without trend can be described. More directly, if the regionalized variable represents a non-stationary process, one works on the residual stationarity of the variable. In universal kriging, the local adjustment is done through a n -order polynomial in terms of the coordinates of the estimated points. The trending in Equation 1 can be represented as:

Table 1. Hole data: UTM locations and undrained resistance values for depths of 10, 20, and 30 meters.

X (m)	Y (m)	Undrained Shear Strength by Hole		
		Z(u) (kPa)		
		10m	20m	30m
637500	7610300	17.46	33.63	47.70
635000	7615300	17.45	33.39	53.22
642000	7601300	15.93	27.44	42.11
636000	7603300	18.44	35.17	50.76
649000	7606300	16.50	33.21	48.85
650000	7615300	16.88	35.38	53.34
641000	7604300	16.81	35.17	44.14
632000	7610300	16.76	32.36	47.76
644000	7609300	17.79	32.86	48.48

$$Z(x) = \frac{1}{N} \sum_{i=1}^N \beta_k f_k(x) + \epsilon(x) \quad (7)$$

with k known, f_i functions, and β_i coefficients can be calculated. An example is modeling a trend as an inclined plane:

$$Z(x) = \beta_0 + \beta_1 x + \beta_2 x + \epsilon(x) \quad (8)$$

There is some criticism regarding the application of universal kriging, because although it is mathematically correct, difficulties arise in separating the phenomenon into two components, a deterministic trend and a random fluctuation based only on the difference between distances.

The kriging of the residues is an option for the cases of non-stationary variables, for their drift can be filtered by a polynomial surface of low degree adjusted to the dataset. According to **Neuman and Jacobson (1984)**, the estimation by kriging of the residues consists of estimating the residues and, subsequently, adding them to the drift, represented by the trend surface obtained by the polynomial adjustment.

When conducting the analysis of the results, cross validation, also known as cokriging, it provides a means to choose, among several methods, the one that best fits the sample. This technique consists of, for a sample with N points, performing kriging using $N - 1$ points, estimating the value of N -th point and calculating the value of the residue between calculated and actual values. The process must be performed with all points.

For undrained resistance data from 9

geotechnical holes made in soil considered homogenous, according to Table 1, the methodology above was performed to estimate the spatial distribution of this variable in the evaluated quota plan from a discretization of the locations in a 100x100 mesh, and subsequent cross validation to quantify the uncertainties involved. According to the company's information security policy, the data was uncharacterized. The routine was implemented in the Python computer language.

3. RESULTS

The variograms generated followed Matheron's methodology, from 20 lag values; and can be visualized for the depths of 10 and 30 meters, in Figures 1 and 2, respectively.

Figure 1. Variogram for depth of 10m.

Figure 2. Variogram for depth of 30m.

The variogram for the depth of 30 meters allows visualization of a trend in the data, indicating a non-stationarity of the shear strength values at this depth, which points to the prediction by one of the kriging techniques with trend. On the other hand, for a depth of 10 meters, the data seems to stabilize at a theoretical level close to 1.15 kPa^2 , which allows the application of the ordinary kriging technique.

The application of ordinary (OK), universal (UK), and residuals (RK) kriging methods, associated to the cross validation for the depths of 10, 20, and 30 meters presented in Figures 3 to 8.

From the results for the spatial distributions of the undrained resistance at the same depth level, two distinct forms of distribution can be visualized, so that the universal kriging technique differs from the others.

Ordinary and residue kriging present the same amplitude of variation of the resistance and similar contours, diverging basically in a greater flatness for the forms of residue kriging. Nevertheless, the results of universal kriging presented periodic behavior. Such behaviors are due to the characteristics of the techniques, which should be highlighted:

1. Universal kriging estimates drift and residuals at the same time, i.e. there is an adjustment of the trend each time the kriging is made at one point, not fitting a global surface to the data set. In addition, the variance of the technique includes both trend and residue estimates;
2. The kriging of the residues assumes that there would be no error in this operation with subsequent correction by adding the trend;

Figure 3. Estimated S_u surface area by OK for the depth of 10m.

Figure 4. Estimated S_u surface area by OK for the depth of 30m.

Figure 5. Estimated S_u surface area by RK for the depth of 10m.

3. Ordinary kriging is a special case of universal kriging when there is no constant change in trend.

In Figure 9, the proximity of the values obtained by the three techniques used implies that there is no differentiation between the techniques. Another aspect it is worth mentioning is that:

Figure 6. Estimated S_u surface area by RK for the depth of 30m.

Figure 8. Estimated S_u surface area by UK for the depth of 30m.

Figure 7. Estimated S_u surface area by UK for the depth of 10m.

1. The relative proximity of the results indicates a relative improvement in accuracy when inserting a more complex model (in this case, with drift modeling);

2. The data received did not undergo any filling, interpolation, or any other method to mitigate the noises (other than the above-mentioned de-characterization). Thus, the technique still had to deal with the data noise. When modelling the drift, most of the noise is found in the residues;
3. The maximum residual estimated by residual kriging can be used as a margin of error for the estimate by providing a confidence interval;
4. Although the quantifications of uncertainty indicate the best technique, the choice should consider the geological and geotechnical knowledge and accumulated experience, so that the geostatistical techniques can assist in the decision-making process, offering additional technical basis.

Figure 9. Residual in each sampling point, in kPa, after cross validation, for depths of 10, 20, and 30 meters (from the top).

4. CONCLUSIONS

The results corroborate the use of geostatistical techniques as estimators of the geotechnical parameter assessed. In this sense, it should be noted that the variances found in the data suggest the need for more information to satisfy stationary conditions and the intrinsic hypothesis.

Within this context, data acquisition is an important process of data analysis. The modeling of uncertainties, as flexible as it may be, brings risks associated with the application of numerical tools and subsequent data estimation. The influence of uncertainties is significant primarily in modeling linear trends from the data found. One possible way to circumvent the variability found in trends is to use, for example, more robust quadratic and/or transcendental modeling.

ACKNOWLEDGMENTS

The authors of the present study thank Petrobras for the technical and financial support it has expended to carry out this work.

5. REFERENCES

- API. **Geotechnical and foundation design considerations**. 1st Edition. API RP 2GEO. Washington: American Petroleum Institute, 2011.
- Cressie, N.; Hawkins, D. M. Robust estimation of the variogram: I. **Journal of the International Association for Mathematical Geology**, v. 12(2), p. 115-125, 1980. <https://doi.org/10.1007/BF01035243>
- Dowd, P. A. **The variogram and kriging: Robust and resistant estimators**. In: Verly, G.; David, M.; Journel, A. G.; Marechal A. (eds) *Geostatistics for natural resources characterization*. Dordrecht: Springer, p. 91-106, 1984. https://doi.org/10.1007/978-94-009-3699-7_6
- Journel, A. G.; Huijbregts, C. J. **Mining geostatistics**. London: Academic press, 1978.
- Lacasse, S.; Nadim, F.; Rahim, A.; Guttormsen, T. R. **Statistical description of characteristic soil properties**. In: *Offshore Technology Conference*, OTC-19117-MS, Houston, Texas, 2007. <https://doi.org/10.4043/19117-MS>
- Matheron, G. *Principles of geostatistics*. **Economic geology**, v. 58(8), p. 1246-1266, 1963. <https://doi.org/10.2113/gsecongeo.58.8.1246>
- Matheron, G. **Les variables régionalisées et leur estimation: Une application de la théorie des fonctions aléatoires aux sciences de la nature**. Paris: Masson et Cie, 1965.
- Matheron, G. **Le krigeage universel**. Paris: École Nationale Supérieure des Mines de Paris, 1969.
- Neuman S. P.; Jacobson, E. A. Analysis of nonintrinsic spatial variability by residual kriging with an application to regional groundwater levels. **Mathematical Geology**, v. 16(5), p. 499-521, 1984. <https://doi.org/10.1007/BF01886329>
- DNV. DNV-RP-C207: **Statistical Representation of Soil Data**. In: Det Norske Veritas: Høvik, Norway 2012.