

Doi: <http://dx.doi.org/10.1590/1809-4430-Eng.Agric.v38n2p260-269/2018>

RELATIONSHIP BETWEEN SAMPLE DESIGN AND GEOMETRIC ANISOTROPY IN THE PREPARATION OF THEMATIC MAPS OF CHEMICAL SOIL ATTRIBUTES

Luciana P. C. Guedes^{1*}, Miguel A. Uribe-Opazo², Paulo J. Ribeiro Junior³, Gustavo H. Dalposso⁴

^{1*} Corresponding author. Western Paraná State University / Cascavel - PR, Brasil. E-mail: luciana_pagliosa@hotmail.com

KEYWORDS

bootstrap; directional trend; geostatistics; spatial variability; tests of isotropy.

ABSTRACT

Spatial variability depends on the sampling configuration and characteristics associated with the georeferenced phenomenon, such as geometric anisotropy. This study aimed to determine the influence of the sampling design on parameter estimation in an anisotropic geostatistical model and the spatial estimation of a georeferenced variable at unsampled locations. Datasets were simulated with geometric anisotropy, considering five values for the anisotropic ratio (1, 2, 3, 4, 5), and three sampling designs: lattice, random and lattice plus close pairs. The simulation results were used as a reference to select anisotropic models to describe the spatial dependence structure in chemical soil properties. For each dataset (with either simulated or chemical soil properties), the values of the georeferenced variables at unsampled locations were estimated by kriging, considering estimated isotropic and anisotropic geostatistical models. The choice of the sampling design influenced the spatial estimation of the georeferenced variable and the quality of the estimation of the geostatistical anisotropic model. The incorporation of geometric anisotropy in the spatial estimation of simulated data sets and soil chemical properties produced differences in the spatial estimation and improved the level of detail of subregions in thematic maps.

INTRODUCTION

In a spatial variability study, the spatial dependence structure described by the semivariance function is considered anisotropic when it depends on the distance and direction separating the locations observed (Maity & Sherman, 2012). For transitive models (with sill) that describe the semivariance function, there are detailed reports in the literature that describe the concept of anisotropy and types, the geostatistical tools that permit the identification of anisotropy and the description of the anisotropic spatial models and their parameters (Boisvert et al., 2009; Facas et al., 2010; Cressie, 2015).

One of the most common forms of anisotropy is the geometric anisotropy. This type of anisotropy occurs when a transitive model of semivariance function presents constant values of nugget and sill in all directions but different ranges. When geometric anisotropy is identified, it must be incorporated as an intrinsic feature of the process that describes the spatial dependence structure to improve the accuracy of the spatial estimation of the

values of a georeferenced variable in unsampled locations, which results in the generation of maps that more accurately describe the spatial variability of the variable throughout the area under study (Guedes et al., 2008; 2013).

In an anisotropic spatial model, the spatial interpolation by kriging method assigns greater weight to sampled values located along the direction of greater spatial continuity. In addition, knowledge of the anisotropy of the spatial dependence structure facilitates planning of sampling configurations for further studies of the spatial variability of this variable in the same area.

Mucha & Blaszczyk (2012) used simulated data to determine that the influence of geometric anisotropy on spatial estimation depends on several factors associated with the spatial dependence structure, such as the distance over which the samples exhibit spatial autocorrelation (range), the degree of intensity of spatial dependence (relative nugget effect) and the distance between the sampling units. According to the studies conducted by

² Western Paraná State University / Cascavel - PR, Brasil.

³ Federal University of Paraná / Curitiba - PR, Brasil.

⁴ Federal Technological University of Paraná / Toledo - PR, Brasil.

Received in: 10-13-2016

Accepted in: 11-14-2017

Guedes et al. (2013), for simulated datasets, it follows that from the anisotropic ratio of 2, the spatial estimates in unsampled locations differ depending on the incorporation or omission of geometric anisotropy.

However, Mucha & Blaszczyk (2012) and Guedes et al. (2013) only considered systematic design (lattices) in their studies of spatial dependence. There are no studies that show the influence that the interaction between different samplings and the presence of anisotropy have on the estimation of the spatial dependence structure and on the estimation of values of the georeferenced variable in non-sampled locations.

Thus, considering simulated data and real data (soil chemical properties) with the presence of geometric anisotropy, the objectives of this work are the following: (a) to determine the implications of changing the sampling design for parameter estimation of the anisotropic spatial model; (b) to determine the implications of changing the sampling design for the spatial estimation of the georeferenced variable at unsampled locations; and (c) to evaluate different sampling designs to determine for which set of anisotropic parameters the isotropy assumption is unreasonable.

MATERIAL AND METHODS

Sampling designs with 100 sample points arranged in a regular area with a maximum limit of x and y coordinates equal to 100 were considered. Three sampling configurations were simulated: lattice 10 x 10, random and lattice 9 x 9 with the addition of 19 nearby points (lattice plus close pairs). The latter sampling design, the simulation study and its sample size were chosen because of the type of design used in the agricultural experiment for the study of spatial variability of soil properties performed in this study.

Datasets for each of these designs representing embodiments of multivariate stochastic processes were simulated for each of these configurations, assuming Gaussian stationary variables $Z = (Z(s_1), \dots, Z(s_n))^\top$, with a linear space model expressed by $Z(s_i) = \mu(s_i) + \epsilon(s_i)$ ($i = 1, \dots, n$), in which $\mu(s_i) = \mu$ is the deterministic term of the model, represented by a constant average μ ; $Z(s_i)$ corresponds to the observed values of the variable under study at the i^{th} known location, denoted by the vector $s_i = (x_i, y_i)^\top$ with ($i = 1, \dots, n$); $\epsilon(s_i)$ is the stochastic error with $E[\epsilon(s_i)] = 0$, and the change in space between points separated by Euclidean distance $h_{ij} = \|h_{ij}\|$, such that $h_{ij} = s_i - s_j$ is determined by a covariance function $C(h_{ij}) = Cov[\epsilon(s_i), \epsilon(s_j)] = \sigma_{ij}$, with ($i, j = 1, \dots, n$) (Kempen et al., 2012; De Bastiani et al., 2015).

Through the covariance function, we find the covariance matrix $\Sigma = [[\sigma_{ij}]]$, with particular parametric form given by $\Sigma = \varphi_1 I_n + \varphi_2 R$, where $R = R(\varphi_2) = [[r_{ij}]]$ is a $n \times n$ symmetric matrix with diagonal elements $r_{ii} = 1$ for $i = 1, \dots, n$. Thus, the covariance function is the function associated with

semivariance by $\gamma(h_{ij}) = C(0) - C(h_{ij})$ for many isotropic and stationary Gaussian processes, where $r_{ij} = \varphi_2^{-1} \sigma_{ij}$. Moreover, this covariance matrix has the following parameters: φ_3 is a function of the range ($a > 0$), φ_1 is the nugget effect ($\varphi_1 \geq 0$) and $\varphi_1 + \varphi_2$ is the sill ($\varphi_1 + \varphi_2$, with $\varphi_2 \geq 0$) (Uribe-Opazo et al., 2012).

However, in this study, we have assumed that the simulated georeferenced variable is stationary and exhibits a stochastic process with geometric anisotropy. Then, the covariance function is the function associated with the semivariance by the [eq. (1)], in which $\|Ah_{ij}\|$ expresses the Euclidean distance between pairs of points in n sampled locations, considering a linear transformation at those locations (Soares, 2014).

$$\gamma(\|Ah_{ij}\|) = C(\|0\|) - C(\|Ah_{ij}\|) \tag{1}$$

The semivariance function [eq. (1)] can be rewritten as a function of the Euclidian distance between two points $s_i = (x_i, y_i)^\top$ and $s_j = (x_j, y_j)^\top$, as shown in [eq. (2)].

$$\|Ah_{ij}\| = \|h_{ij}^*\| = h_{ij}^* = \left(u_{ij}^2 + \frac{v_{ij}^2}{F_a^2} \right)^{\frac{1}{2}} \tag{2}$$

where,

$$u_{ij} = [(y_i - y_j) \sin(\beta)] + [(x_i - x_j) \cos(\beta)],$$

β is the highest spatial continuity angle in π radians ($0 \leq \beta \leq \pi$) defined in the azimuth system;

$$v_{ij} = [(y_i - y_j) \cos(\beta)] - [(x_i - x_j) \sin(\beta)],$$

and

$$\varphi_4 = F_a = \frac{a_\beta}{a_{\beta + \frac{\pi}{2}}}$$

is the anisotropic ratio, where a_β is the spatial dependency distance (range) in the direction of higher spatial continuity (β) and $a_{\beta + \pi/2}$ is the spatial dependency distance (range) in the direction of lower spatial continuity ($\beta + \frac{\pi}{2}$), assuming that the directions of higher and lower spatial continuity are orthogonal.

The simulated data have an anisotropic semivariance function [eq. (1)] with an exponential model with parameter vector $\varphi = (\varphi_1, \varphi_2, \varphi_3, \varphi_4)^\top = (0, 10, 15, F_a)^\top$, $a = 60$ and β is equal to 0° , following the system of azimuth direction. For each sampling design, five datasets were simulated, in which F_a took on each of the values of 1, 2, 3, 4, and 5. Thus, to combine the sample values of the setting and anisotropic ratio, 1000 sets of data were simulated.

In each set of simulated data and in each of sampling designs, we estimated the parameter vector by the maximum likelihood method, considering the following situations: (1) anisotropic model with the

parameter vector $\theta = (\mu, \varphi^T)^T = (\mu, \varphi_1, \varphi_2, \varphi_3, \varphi_4)^T$, with $\beta = \mathbf{0}^o$ fixed; and (2) isotropic model with the parameter vector $\theta^* = (\mu, \varphi_1, \varphi_2, \varphi_3)^T$.

Considering the anisotropic model, to estimate asymptotic standard errors of θ , we used the inverse of the expected information matrix (Uribe-Opazo et al., 2012). The expected information matrix to an anisotropic process is given by [eq. (3)], where $k(\mu) = X^T \Sigma^{-1} X$ and $K(\varphi) = [(k_{bc}(\varphi))]$, with

$$k_{bc}(\varphi) = \frac{1}{2} \text{tr} \left(\Sigma^{-1} \left[\frac{\partial \Sigma}{\partial \varphi_b} \right] \Sigma^{-1} \left[\frac{\partial \Sigma}{\partial \varphi_c} \right] \right) \quad \text{for } b, c = 1, \dots, 4.$$

$$K(\theta) = \begin{pmatrix} k(\mu) & \mathbf{0} \\ \mathbf{0} & K(\varphi) \end{pmatrix} \quad (3)$$

Then, the first derivatives of Σ with respect to $\varphi_1, \varphi_2, \varphi_3$ and $\varphi_4 = F_a$ are given, respectively, by the following:

$$\frac{\partial \Sigma}{\partial \varphi_1} = I_n, \quad \frac{\partial \Sigma}{\partial \varphi_2} = R, \\ \frac{\partial \Sigma}{\partial \varphi_3} = \varphi_2 \left(\frac{\partial R}{\partial \varphi_3} \right) \quad \text{and} \quad \frac{\partial \Sigma}{\partial F_a} = \varphi_2 \left(\frac{\partial R}{\partial F_a} \right), \\ \frac{\partial R}{\partial \varphi_3} = \left[\left(\frac{\partial r_{ij}}{\partial \varphi_3} \right) \right] \quad \text{and}$$

$$\frac{\partial R}{\partial F_a} = \left[\left(\frac{\partial r_{ij}}{\partial F_a} \right) \right], \text{ for } i, j = 1, \dots, n.$$

The first derivatives of r_{ij} with respect to φ_3 and $\varphi_4 = F_a$ for the exponential and Gaussian covariance functions are given, respectively, by [eq. (4)] and [eq. (5)]:

$$\frac{\partial r_{ij}}{\partial \varphi_3} = \frac{r_{ij} h_{ij}^*}{\varphi_3^2} \quad \text{and} \quad \frac{\partial r_{ij}}{\partial F_a} = \frac{r_{ij} v_{ij}^2}{F_a^2 h_{ij}^* \varphi_3} \quad (4)$$

$$\frac{\partial r_{ij}}{\partial \varphi_3} = \frac{2r_{ij} (h_{ij}^*)^2}{\varphi_3^3} \quad \text{and}$$

$$\frac{\partial r_{ij}}{\partial F_a} = \frac{2r_{ij} v_{ij}^2}{F_a^3 \varphi_3^2} \quad (5)$$

$$\text{for } i, j = 1, \dots, n, \text{ and } \lim_{i \rightarrow j} \left(\frac{r_{ij} v_{ij}^2}{F_a^3 h_{ij}^* \varphi_3} \right) = \mathbf{0}.$$

The first derivatives of r_{ij} with respect to φ_3 and $\varphi_4 = F_a$ for the Matérn family covariance function (Matérn, 1986) are given, respectively, by [eq. (6)] and [eq. (7)]:

$$\frac{\partial r_{ij}}{\partial \varphi_3} = - \left(\frac{1}{\varphi_3} \right) \left[k r_{ij} + [2^{k-1} \Gamma(k)]^{-1} \left(\frac{h_{ij}^*}{\varphi_3} \right)^{k+1} K'_k \left(\frac{h_{ij}^*}{\varphi_3} \right) \right] \quad (6)$$

$$\frac{\partial r_{ij}}{\partial F_a} = - \left(\frac{v_{ij}^2}{F_a^2 h_{ij}^*} \right) \left[\left(\frac{k r_{ij}}{h_{ij}^*} \right) + \frac{1}{\varphi_3} [2^{k-1} \Gamma(k)]^{-1} \left(\frac{h_{ij}^*}{\varphi_3} \right)^k K'_k \left(\frac{h_{ij}^*}{\varphi_3} \right) \right] \quad (7)$$

$$\text{for } i, j = 1, \dots, n, \text{ where } \Gamma \text{ is the gamma function, } K'_k(t) = \frac{\partial K_k(t)}{\partial t} = - \left(\frac{1}{2} \right) [K_{k+1}(t) + K_{k-1}(t)],$$

$$K_k(t) = \left(\frac{1}{2} \right) \int_0^\infty x^{k-1} \exp \left[-t(w+w^{-1})/2 \right] dw \quad \text{is the modified Bessel function of the third type of order } k, \\ \text{with } k > \mathbf{0} \text{ fixed; and } \lim_{i \rightarrow j} \left(\frac{\partial r_{ij}}{\partial F_a} \right) = \mathbf{0}.$$

Considering the estimated geostatistical models, the values of the georeferenced variables at unsampled locations were estimated in the presence and absence of geometric anisotropy by the ordinary Kriging method.

The sets of estimated values of the georeferenced variables at unsampled locations (considering the anisotropic and isotropic models) were compared for the similarity measures: overall accuracy (\widehat{OA}) Kappa concordance index (\widehat{K}) and Tau concordance index (\widehat{T}), which aim to compare the thematic maps generated by the two interpolations, with the same rating from the error matrix, considering 10 intervals of values or classes (De Bastiani & Uribe-Opazo, 2012). Thus, using bootstrap resampling, confidence intervals were constructed for the mean similarity measures using the normal approximation (Efron & Tibshirani, 1993), considering the initial sample to perform the bootstrap resampling with 1000 values of each measure similarity obtained in the 1000 simulations (Fang & Wang, 2012; Rossoni et al. 2014).

In each set of simulated data and in each sampling design, tests of isotropy from Guan et al. (2004) (GSC) and Maity & Sherman (2012) (MS) were applied. For both tests, the null hypothesis that the georeferenced variable is isotropic was elaborated by considering a set of linear contrasts. Maity & Sherman (2012) elaborated that the null hypothesis is equivalent to $H_0: AC(t) = \mathbf{0}$ and the alternative hypothesis to $H_0: AC(t) \neq \mathbf{0}$. However, Guan et al. (2004) defined the null hypothesis as $H_0: A\gamma(t) = \mathbf{0}$ and the alternative hypothesis as $H_0: A\gamma(t) \neq \mathbf{0}$, for $t \in \Lambda$, where $\Lambda = (t_1, \dots, t_k)^T$ is a set of chosen spatial lags, and A is a matrix of contrasts, whose dimension depends on the definition of Λ . These tests propose to compare the values of covariance (Guan et al., 2004) and semivariance (Maity & Sherman, 2012) for pairs of lags described by Λ .

The tests can be used when the sampling region is any convex subset in $\mathbf{R}^d, d > 1$. The MS test was developed for sampling locations that are irregularly spaced. The covariance function is calculated by the kernel estimator, and the grid-based block bootstrap is used to calculate the test statistic. The GSC test was developed for gridded and non-gridded sampling locations. The semivariance is calculated by Matheron's estimator, and the test statistic is evaluated using a subsampling technique.

For each simulation, these tests were constructed based on the set of lags $\Lambda = ((1,0), (0,1), (5,0), (0,5), (15,0), (0,15))^T$. The choice of Λ was made by considering short lags and the direction of the weakest and strongest spatial correlation in each set of simulated data. In the MS test, the block length was equal to 10 and 200 resamples for each set of simulated data. In the GSC test, to preserve enough pairs

of sampling locations in each subblock, a block length of 2 was chosen.

The geostatistical analysis, identification of anisotropy and the analysis of the influence of the incorporation of anisotropy on the spatial estimation were also performed on a set of real data from a commercial area of grain production in Cascavel, Paraná, located at approximately 24.95° of South latitude and 53.37° of West longitude and at an average height of 650 m above sea level. The dominant soil type is the Rhodic Hapludox, which has a clay-like texture. The data refer to the crop year 2010/2011 and an area of 167.35 ha of experiments conducted by researchers from the Spatial Statistics Laboratory of the State University of West Paraná - Brazil.

The sampling configuration used in this study was a lattice plus close pairs, with 102 sampling points. All samples were georeferenced and localized with the aid of a signal receiving apparatus with a global positioning system (GPS) Geoexplore 3 set up for the Universal Transverse Mercator (UTM) coordinate system.

The area was planted with soybeans, and the data used in this study were related to soil chemical properties that exhibited geometric anisotropy in the geostatistical analysis: carbon ($C - g\ dm^{-2}$) and potential acidity ($H + Al^{+3} - cmolc\ dm^{-2}$). The data sets were obtained by performing routine chemical analysis in the soil analysis laboratory of Cooperativa Central de Pesquisa Agrícola Ltda. (COODETEC) of representative samples of each plot of approximately 50 grams, which were obtained by mixing four replicates of the parcel. $H + Al^{+3}$ were obtained by the Shoemaker, Mac Lean and Pratt (SMP) buffer method.

Simulated datasets and geostatistical analyses were prepared in the software R (R Development Core Team, 2016) using the geoR package (Ribeiro Jr. & Diggle, 2016). Tests of isotropy were conducted using the sm package (Bowman & Azzalini, 2015).

RESULTS AND DISCUSSION

Analysis of simulated data

For the simulated data and considering the anisotropic model for the semivariance function, the nugget (φ_1) and anisotropic ratio ($\varphi_4 = F_a$) parameters had a greater influence on the change in the sampling design. We obtained the worst results in terms of quality of the estimation of these parameters for the lattice design: the estimated values were more distant from the nominal values of the parameters φ_1 and F_a (Fig. 1-a and 1-b) and the estimated values of the standard errors of φ_1 and F_a were highest (Fig. 1-c and 1-d) compared with the random design and lattice plus close pairs.

By contrast, improved estimates of the φ_1 parameter were obtained by random design: the estimated values of φ_1 were closest to its nominal value (Fig. 1-a) and the lowest estimated values of the standard errors of φ_1 (Fig. 1-c).

The estimation of the φ_1 influences the determination of the weights assigned to the samples in the spatial estimation of a georeferenced variable at unsampled locations, performed by Kriging (Soares, 2014). Thus, the results for the estimation of the parameter φ_1 demonstrated that the use of random design in the study of a georeferenced variable with geometric anisotropy permits a more efficient description of the spatial variability, as expressed by the map.

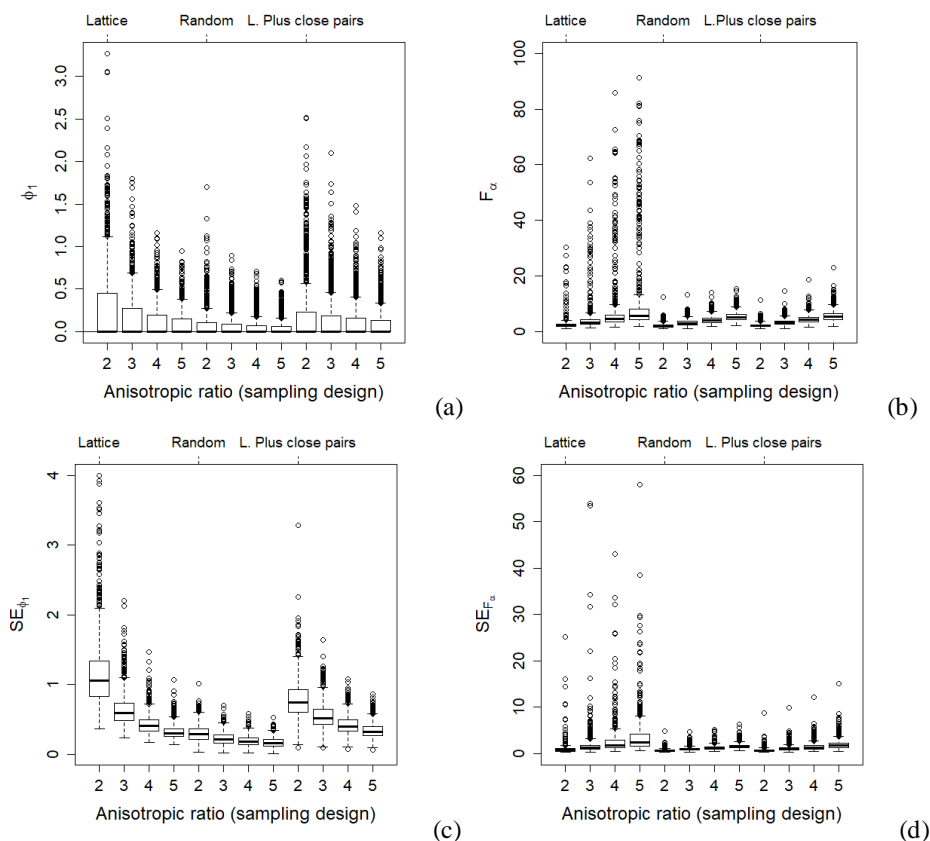


FIGURE 1. Boxplots for: estimated values of (a) nugget effect (φ_1) and (b) anisotropic ratio (F_α); and estimated values of the standard error (SE) of (c) nugget effect (φ_1) and (d) anisotropic ratio (F_α) for an anisotropic geostatistical model.

Figure 2 presents boxplot graphs of the similarity measures that compare the spatial estimation performed using the anisotropic and isotropic exponential models, both with estimated parameters. The dashed lines in these graphs correspond to ranges of values for these accuracy measures in which the spatial estimates were of similar intensity. For overall accuracy, the minimum level of

similarity required is $\widehat{\sigma_A} = 0.85$ (De Bastiani & Uribe-Opazo, 2012); and the concordance indices of Kappa (\widehat{K}) and Tau (\widehat{T}), as defined by Krippendorff (1980), were classified as low, medium and high similarity between the values estimated by the two anisotropic models at $[0.0, 0.67]$, $[0.67, 0.80]$ and $[0.80, 1.0]$, respectively.

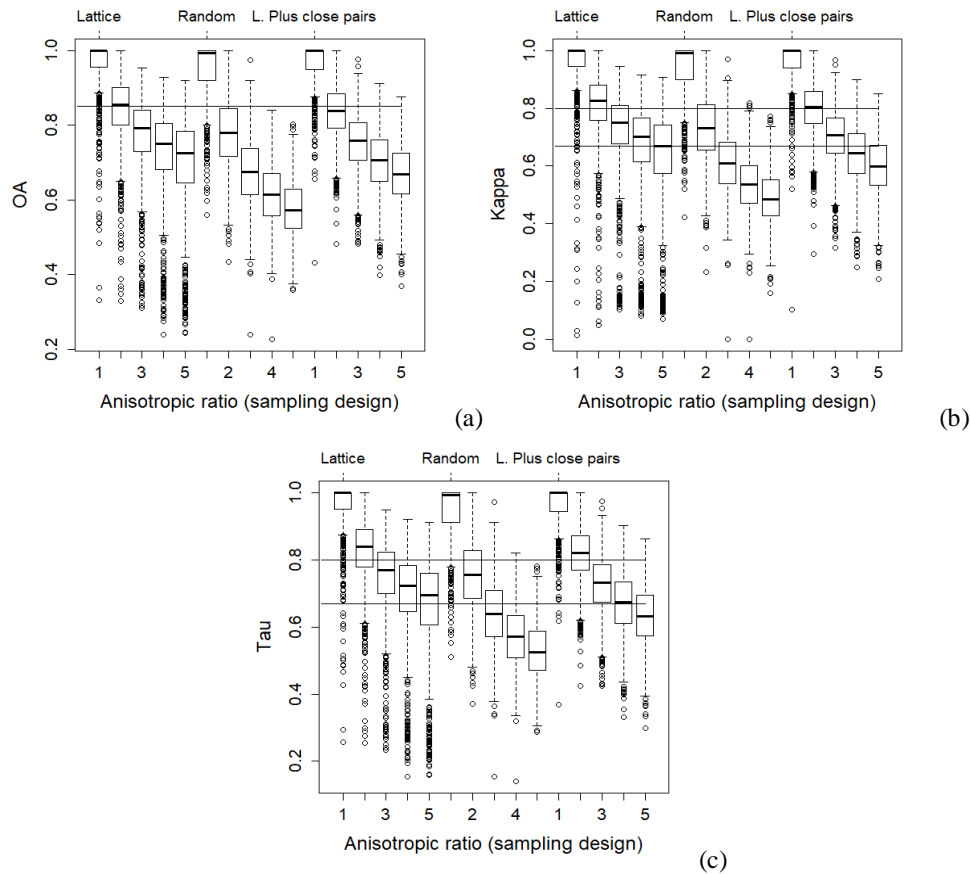


FIGURE 2. Boxplots for estimated values of the (a) overall accuracy (OA) and concordance index of (b) Kappa and (c) Tau regarding the comparison of spatial estimates using the anisotropic and isotropic geostatistical models both with estimated parameters for all simulations. L.: lattice.

These graphs illustrate that, for all studied sampling designs, the similarity measures decreased as the anisotropic ratio increased. For random sampling and an anisotropic ratio factor of 2, 94% of the simulations did not exhibit high similarity (Krippendorff, 1980; De Bastiani & Uribe-Opazo, 2012), indicating a relevant difference in the spatial estimation based on the incorporation or omission of the geometric anisotropy, from the anisotropic ratio of 2. For systematic sampling and lattice plus close pairs, this relevant difference occurred in more than 90% of the simulations for anisotropic ratio values (F_a) equal to 3.

However, all graphs (Fig. 2) demonstrated that, particularly for systematic sampling, there is high variability of these similarity measures and many small values (outliers) in approximately 10% of the simulations. For the random and lattice plus close pairs, outliers were observed in approximately 3.5% of the simulations.

These outliers occurred in simulations that produced an overestimation of the anisotropic ratio, in agreement with results obtained for the systematic sampling by Guedes et al. (2013). However, these authors emphasize that this overestimation of the anisotropic ratio occurred only for small sample sizes.

These conclusions are also supported by the confidence intervals constructed for the average of similarity measures using the bootstrap method and 95% confidence level (Dalposso et al., 2016). These confidence intervals are shown in Table 1 for all simulations grouped according to the anisotropic ratio, the sampling configuration and type of comparison performed. These confidence intervals indicated that, for random sampling

from the anisotropic ratio of 2, the confidence intervals for the mean accuracy measurements did not contain the minimum amounts required to be classified as high similarity for spatial estimates from the geostatistical models incorporating or omitting geometric anisotropy.

However, for the lattice plus close pairs and systematic sampling, the confidence intervals for the mean accuracy measurements did not contain the minimum amounts required to be classified as high similarity for anisotropic ratio (F_a) value of 3. These results provide evidence that there are differences in the classifications of generated thematic maps when considering the incorporation of anisotropy or not for the lattice plus close pairs and systematic sampling.

These differences can be explained by the fact that when the thematic maps are elaborated considering the presence of geometric anisotropy, these maps present subregions with greater continuity in the direction in which the anisotropy was identified (Guedes et al., 2008; 2013).

Table 1 also showed that the similarity measures decrease significantly as the anisotropic ratio increases at 5% significance, considering all studied sampling designs. Similar results were described by Rossoni et al. (2014), who considered the mean square error of the spatial estimation of a georeferenced variable for unsampled locations as a measure of the impact of the incorporation of geometric anisotropy in spatial estimation in simulations generated for random sampling.

TABLE 1. Confidence intervals for the average Overall Accuracy and Kappa and Tau concordance indexes obtained by the bootstrap method, comparing the spatial estimation using anisotropic and isotropic models, grouped according to the anisotropic ratio and the sampling design.

Design	Anisotropic ratio ($F_{1\alpha}$)		
	1 ^(a)	2	3
Overall Accuracy (\bar{OA})			
Random	[0.947, 0.956]	[0.774, 0.787]	[0.671, 0.682]
Lattice	[0.894, 0.904]	[0.834, 0.845]	[0.765, 0.778]
L. Plus Close Pairs	[0.964, 0.970]	[0.828, 0.836]	[0.747, 0.757]
Kappa Index (\bar{K})			
Random	[0.935, 0.945]	[0.725, 0.740]	[0.601, 0.615]
Lattice	[0.946, 0.959]	[0.795, 0.810]	[0.712, 0.729]
L. Plus Close Pairs	[0.954, 0.963]	[0.789, 0.801]	[0.692, 0.705]
Tau Index (\bar{T})			
Random	[0.941, 0.951]	[0.749, 0.763]	[0.634, 0.647]
Lattice	[0.954, 0.964]	[0.815, 0.828]	[0.739, 0.753]
L. Plus Close Pairs	[0.959, 0.967]	[0.808, 0.819]	[0.719, 0.730]

L.: lattice; ^(a) Isotropic case.

Comparing the sampling configurations for the confidence intervals of these similarity measures, mainly from the anisotropic ratio ($F_{1\alpha}$) equal to 2, the similarity measures of the studied sampling configurations differed significantly at 5% probability. The random sampling had the lowest values of these measures, whereas the systematic sampling had the greatest values.

Table 2 shows the percentage of rejections of the nonparametric tests of isotropy at the 5% nominal level from 100 simulations. The method from Maity & Sherman (2012) considering the random sampling and lattice plus close pairs was applied. The method from Guan et al. (2004) considering the systematic sampling was also applied. For all sampling designs, an inflated type I error (case isotropic with factor anisotropy equal to 1) was observed.

The power of tests using random and systematic sampling is relatively larger than the lattice plus close pairs. In addition, for all configurations examined, the empirical power increased as the anisotropic ratio increased. However, according to Guan et al. (2004) and Maity & Sherman (2012), empirical powers tend to be low when the sample size is small and/or the anisotropy is weak.

The real data grid was also used in our simulations. The results not presented in this article showed that for the irregular area, a higher inflating of type I error (case isotropic), compared with the rectangular area, was observed.

TABLE 2. Percentage of rejections at 5% nominal level of the nonparametric tests of isotropy.

Anisotropic ratio	MS test		GSC test
	Random	L. Plus Close Pairs	Lattice
1 ^(a)	36.70	42.20	33.50
2	49.90	56.00	75.00
3	60.10	72.40	93.60
4	68.00	80.00	98.60
5	74.20	84.60	99.30

L.: lattice; ^(a) Isotropic case; MS: method from Maity & Sherman (2012); GSC: method from Guan et al. (2004).

ANALYSIS OF SOIL CHEMICAL PROPERTIES

Fig. 3-(a) and Fig. 3-(b) illustrate the values of variables C and $H + Al^{+3}$ for the sampled points in the study area, using quartile classification. Groups of points with the same classification extended in the direction of $\beta = 90^\circ$ (azimuth system), indicating the presence of geometric anisotropy in that direction.

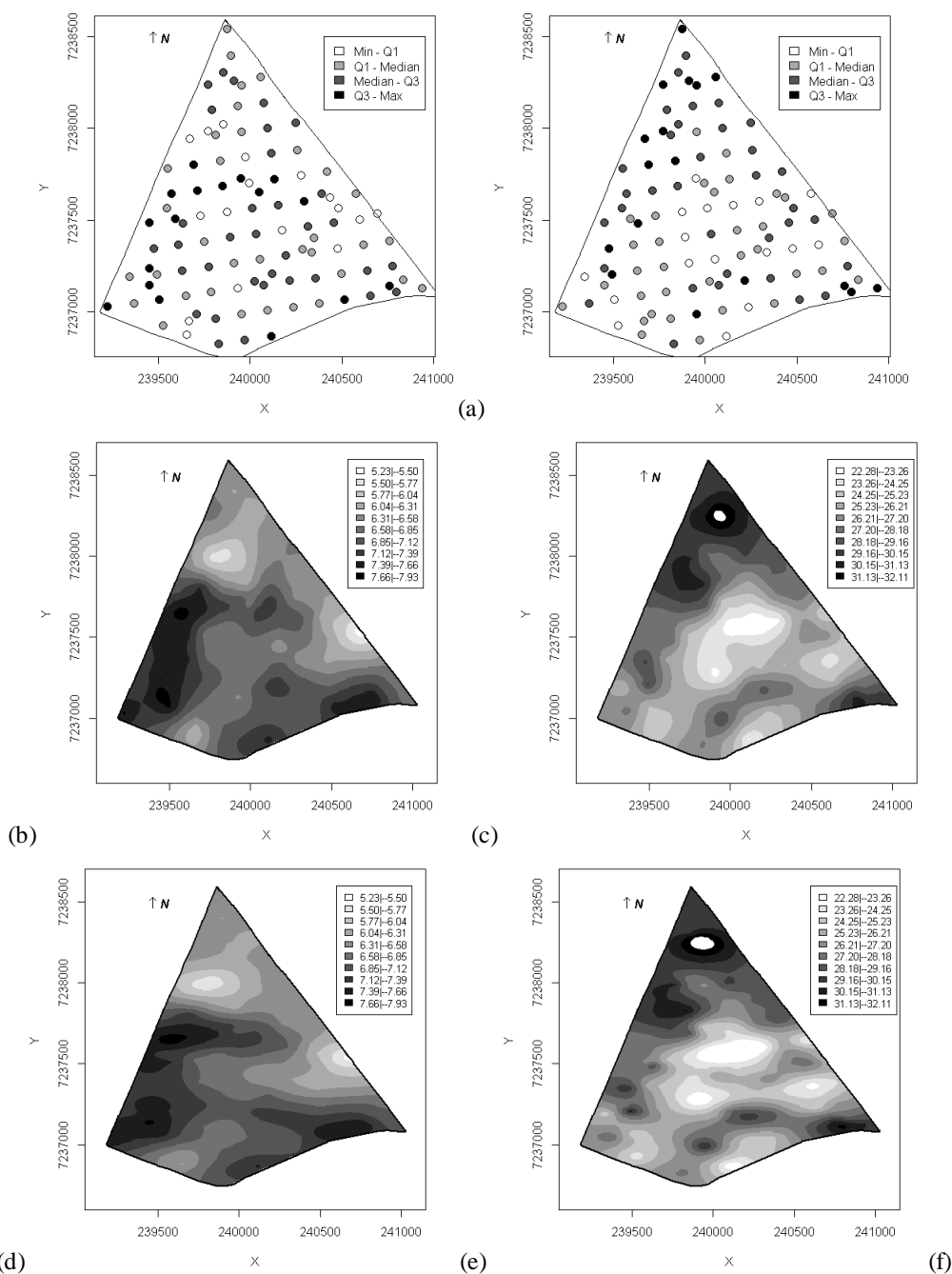


FIGURE 3. Spatial representation of the locations sampled in the study area, classified into equal-amplitude ranges for the variables (a) $H+Al^{+3}$ and (b) C and thematic maps for (c) $H + Al^{+3}$ considering the isotropic model, (d) C considering the isotropic model, (e) $H + Al^{+3}$ considering the anisotropic model, and (f) C considering the anisotropic model. Min: minimum, Q1: first quartile, Q3: third quartile, Max: maximum.

Table 3 shows the results of univariate geostatistical analysis with the estimated spatial parameters for the variables C and $H + Al^{+3}$ for isotropic model and geometric anisotropic model. According to the cross-validation criteria (Monego et al., 2015) (Table 4), the exponential model was the best fit for C and $H + Al^{+3}$ by the maximum likelihood method, and the best estimates for the measure cross-validation were obtained for the anisotropic models.

Higher spatial continuity directions were established using the conventional directions adopted by Guedes et al. (2013) considering the azimuth system. In

most models, there is a moderate (Cambardella et al., 1994; $25\% < RNE \leq 75\%$) spatial dependence, where $RNE = 100\varphi_1 / (\varphi_1 + \varphi_2)$. Only the isotropic model estimated for the variable C exhibited a strong spatial dependence (Cambardella et al., 1994; $RNE \leq 25\%$). The anisotropic ratio is near 2, and the estimated values of the range and its standard error were higher in the isotropic model. Moreover, applying the MS nonparametric test (Table 3), the null hypothesis of isotropy evidently can be rejected ($p\text{-value} < 0.05$).

TABLE 3. Adjusted spatial models and estimated values of parameters by ML. The standard deviations of the estimates are in parentheses.

Variable	Model	$\hat{\mu}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_1 + \hat{\varphi}_2$	\hat{a}	\widehat{RNE}	$\hat{\beta}$	\widehat{Fa}
H+AI ³	Iso Expo.	6.784 (0.260)	1.619 (0.632)	0.700 (0.645)	2.319	512.97 (171.46)	69.81		
	Ani Expo.	6.783 (0.290)	1.639 (0.475)	0.655 (0.495)	2.294	462.53 (156.77)	71.45	90°	2.22* (2.31)
C	Iso Expo.	27.284 (0.882)	3.487 (1.937)	7.239 (2.867)	10.726	684.79 (134.23)	32.51		
	Ani Expo.	27.026 (0.729)	1.890 (1.963)	7.721 (2.673)	9.611	355.33 (61.64)	19.66	90°	2.09* (0.89)

$\hat{\mu}$: mean; \hat{a} : range; $\hat{\varphi}_1$: nugget effect; ($\hat{\varphi}_2$): sill; $\widehat{RNE} = 100\hat{\varphi}_1 / (\hat{\varphi}_1 + \hat{\varphi}_2)$; (%): relative nugget effect; $\hat{\beta}$: higher spatial continuity angle; \widehat{Fa} : Ani ratio; Iso: isotropic; Ani: anisotropic, and Expo.: exponential model; * significant at the 5 percent level, by MS nonparametric test fort isotropy.

The results of the comparison of spatial estimates obtained by the isotropic and anisotropic models (Table 4) revealed that the anisotropic model resulted in lower values of the average Kriging variance for both C and H + AI³. According to Yamamoto (2000), the kriging variance is used as a quality indicator of spatial estimation. This result indicates that the anisotropic model produced the best efficiency and spatial estimation.

In addition, the values of the similarity measures indicate a relevant difference in the spatial estimates of the maps generated by the isotropic and anisotropic models (Kripendorff, 1980; De Bastiani & Uribe-Opazo, 2012). The results for the similarity measures in this work are similar to conclusions presented by Guedes et al. (2013)

considering simulated and real data from a systematic sampling.

The differences in the spatial estimation of the variables under study upon the incorporation of geometric anisotropy in the spatial dependence structure can also be displayed in thematic maps (Fig. 3-(c) to 3-(f)). In thematic maps constructed from the anisotropic models (Fig. 3-d and 3-f), a greater spatial continuity of subregions was observed in the higher spatial continuity directions compared to the thematic maps generated by isotropic models (Fig. 3-c and 3-e). Similar results were described by Guedes et al. (2013), who estimated chemical soil properties for unsampled locations, considering stratified systematic unaligned sampling.

TABLE 4. Adjusted spatial models with their respective measurements obtained by cross-validation and associated measures the spatial estimation and the similarity in the comparison between anisotropic and isotropic models, its spatial estimation.

Variable	Model	ME	SME	S _{ME}	S _{SME}	AE	$\overline{\sigma^2}$	\widehat{OA}	\widehat{K}	\widehat{T}
H+AI ³	Iso	-4 · 10 ⁻⁵	-2 · 10 ⁻⁵	1.47	1.01	121.8	2.09	0.653	0.587	0.614
	Ani	-2 · 10 ⁻⁵	1 · 10 ⁻⁵	1.46	1.01	120.5	2.03			
C	Iso	-3 · 10 ⁻²	-5 · 10 ⁻³	2.66	1.00	219.9	6.35	0.651	0.596	0.613
	Ani	-3 · 10 ⁻²	-7 · 10 ⁻³	2.56	1.04	210.7	5.26			

\widehat{OA} : Overall accuracy; Concordance index of Kappa (\widehat{K}) and Tau (\widehat{T}); $\overline{\sigma^2}$: average kriging variance; ME: mean error; S_{ME}: standard deviation of the mean error; SME: mean standardized error; S_{SME}: standard deviation of the standardized error; AE: absolute error; Iso: isotropic; Ani: anisotropic.

CONCLUSIONS

The analyses of the simulations demonstrated that the systematic sampling exhibited the worst performance in relation to the quality of estimation of the geostatistical anisotropic model compared with the random sampling and lattice plus close pairs. This lower performance generally resulted in overestimated values of the nugget and anisotropic ratio, the highest estimated values of the nugget, range and anisotropic ratio, and the highest estimated values of the standard errors of the nugget, range and anisotropic ratio.

Moreover, the results of the simulations demonstrated that the random sampling and lattice plus close pairs are more appropriate for the incorporation of geometric anisotropy with respect to spatial estimation for georeferenced variables at unsampled locations. The similarity measures and the nonparametric tests demonstrated that random sampling, lattice plus close pairs

and lattice sampling have a relevant influence on the incorporation of geometric anisotropy in most simulations from an isotropic ratio values of 2.

Most of the similarity measures differed significantly for the different sampling designs. These confidence intervals did not overlap at 95% confidence when the anisotropic ratio was changed.

The similarity measures for the chemical attributes of soil sampled in an irregular area with a lattice plus close pairs sampling design indicated a relevant difference in the spatial estimation depending on the incorporation or omission of geometric anisotropy in the geostatistical model. Moreover, the smallest values of the average kriging variance and improved detail of the subregions in the thematic maps were observed in the spatial estimation that incorporated geometric anisotropy. These results demonstrate that the incorporation of geometric anisotropy

produces more reliable estimates that best indicate the spatial continuity of the studied attributes.

ACKNOWLEDGEMENTS

Authors acknowledge CAPES, CNPq and Araucaria Foundation for financial supports.

REFERENCES

- Boisvert JB, Manchuk JG, Deutsch CV (2009) Kriging in the presence of locally varying anisotropy using non-euclidean distances. *Mathematical. Geoscience* 41(5):585-601. DOI: <https://doi.org/10.1007/s11004-009-9229-1>
- Bowman AW, Azzalini A (2015) R package *sm*: nonparametric smoothing methods (version 2.2-5.4). University of Glasgow, UK and Università di Padova.
- Cambardella CA, Moorman TB, Novak IM, Parkin TB, Karlen DL, Turco RF, Konopka AE (1994) Field-scale variability of soil properties in Central Iowa Soils. *Soil Science Society of American Journal* 58(2):1501-1511. DOI: 10.2136/sssaj1994.03615995005800050033x
- Cressie NAC (2015) *Statistics for spatial data*. New York, John Wiley & Sons. 928p.
- Dalposso GH, Uribe-Opazo MA, Johann JA (2016) Soybean yield modeling using bootstrap methods for small samples. *Spanish Journal of Agricultural Research* 14(3):1-15. DOI: <http://dx.doi.org/10.5424/sjar/2016143-8635>
- De Bastiani F, Uribe-Opazo MA (2012) Comparison of maps of spatial variability of soil resistance to penetration constructed with and without covariables using a spatial linear model. *Engenharia Agrícola* 32(2):394-404. DOI: <http://dx.doi.org/10.1590/S0100-69162012000200019>
- De Bastiani F, Mariz de Aquino Cysneiros AH, Uribe-Opazo MA, Galea M (2015) Influence diagnostics in elliptical spatial linear models. *Test* 24(2):322-340. DOI: 10.1007/s11749-014-0409-z
- Efron B, Tibshirani RJ (1993) *An introduction to the bootstrap*. Boca Raton, Chapman & Hall. 456p.
- Facas NW, Mooney MA, Asce M, Furrer R (2010) Anisotropy in the spatial distribution of roller-measured soil stiffness. *International Journal of Geomechanics* 10(4):129-135. DOI: [https://doi.org/10.1061/\(ASCE\)GM.1943-5622.0000053](https://doi.org/10.1061/(ASCE)GM.1943-5622.0000053)
- Fang Y, Wang J (2012) Selection of the number of clusters via the bootstrap method. *Computational Statistics and Data Analysis* 56(3):468-477. DOI: 10.1016/j.csda.2011.09.003
- Guan Y, Sherman M, Calvin JA (2004) A nonparametric test for spatial isotropy using subsampling. *Journal of American Statistical Association* 99(467):810-821. DOI: 10.1198/016214504000001150
- Guedes LPC, Uribe-Opazo MA, Johann JA, Souza EG (2008) Anisotropy to analyze spatial variability of some spatially referenced soil chemical variables. *Revista Brasileira de Ciência do Solo* 32(6):2217-2226. DOI: <http://dx.doi.org/10.1590/S0100-06832008000600001>
- Guedes LPC, Uribe-Opazo MA, Ribeiro JR. PJ (2013) Influence of incorporating geometric anisotropy on the construction of thematic maps of simulated data and chemical attributes of soil. *Chilean Journal of Agricultural Research* 73(4):414-423. DOI: <http://dx.doi.org/10.4067/S0718-58392013000400013>
- Kempen B, Brus DJ, Heuvelink GBM (2012) Soil type mapping using the generalized linear geostatistical model: a case study in a Dutch cultivated peatland. *Geoderma* 189-190:540-553. DOI: <https://doi.org/10.1016/j.geoderma.2012.05.028>
- Krippendorff K (1980) *Content analysis: an introduction to its methodology*. Beverly Hills, Sage Publications. 456p.
- Maity A, Sherman M (2012) Testing for spatial isotropy under general designs. *Journal of Statistical Planning and Inference* 142(5):1081-1091. DOI: 10.1016/j.jspi.2011.11.013
- Monego MD, Ribeiro Jr. PJ, Ramos, P (2015) Comparing the performance of geostatistical models with additional information from covariates for sewage plume characterization. *Environmental Science and Pollution Research* 22(8):5850-5863. DOI: 10.1007/s11356-014-3709-7
- Mucha J, Blaszczyk MW (2012) Variability anisotropy of mineral deposits parameters and its impact on resources estimation - a geostatistical approach. *Gospodarka Surowcami Mineralnymi – Mineral Resources Management* 28(4):113-135. DOI 10.2478/v10269-012-0037-8
- R Development Core Team (2016) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. Available in: <http://www.R-project.org>. Accessed: May 19, 2016.
- Ribeiro Jr. PJ, Diggle PJ (2016) *geoR: A package for geostatistical analysis*. *R-NEWS* 1:15-18.
- Rossoni DF, Lima RR, Oliveira MS (2014) Proposal and validation of bootstrap tests for anisotropy detection in continuous-spatial phenomena. *Revista de Estatística UFOP* 3(2):210-227.
- Soares A (2014) *Geoestatística para as ciências da terra e do ambiente*. Lisboa, Press. 214p.
- Uribe-Opazo MA, Borssoi JA, Galea M (2012) Influence diagnostics in Gaussian spatial linear models. *Journal of Applied Statistics* 39(3):615-630. DOI: <http://dx.doi.org/10.1080/02664763.2011.607802>
- Yamamoto JK (2000) An alternative measure of the reliability of ordinary kriging estimates. *Mathematical Geology* 32(4):489-509. DOI: <https://doi.org/10.1023/A:1007577916868>