COMPARISON OF MAPS OF SPATIAL VARIABILITY OF SOIL RESISTANCE TO PENETRATION CONSTRUCTED WITH AND WITHOUT COVARIABLES USING A SPATIAL LINEAR MODEL

FERNANDA DE BASTIANI[1], MIGUEL A. URIBE-OPAZO[2], GUSTAVO H. DALPOSSO[3]

ABSTRACT: A study about the spatial variability of data of soil resistance to penetration (RSP) was conducted at layers 0.0-0.1 m, 0.1-0.2 m and 0.2-0.3 m depth, using the statistical methods in univariate forms, i.e., using traditional geostatistics, forming thematic maps by ordinary kriging for each layer of the study. It was analyzed the RSP in layer 0.2-0.3 m depth through a spatial linear model (SLM), which considered the layers 0.0-0.1 m and 0.1-0.2 m in depth as covariable, obtaining an estimation model and a thematic map by universal kriging. The thematic maps of the RSP at layer 0.2-0.3 m depth, constructed by both methods, were compared using measures of accuracy obtained from the construction of the matrix of errors and confusion matrix. There are similarities between the thematic maps. All maps showed that the RSP is higher in the north region.

KEYWORDS: geostatistics, maximum likelihood, error matrix.

COMPARAÇÃO DE MAPAS DE VARIABILIDADE ESPACIAL DA RESISTÊNCIA DO SOLO À PENETRAÇÃO CONSTRUÍDOS COM E SEM COVARIÁVEIS USANDO UM MODELO ESPACIAL LINEAR

RESUMO: Realizou-se um estudo sobre a variabilidade espacial de dados de resistência do solo à penetração (RSP), nas camadas de 0,0-0,1 m, 0,1-0,2 m e 0,2-0,3 m de profundidade, utilizando métodos estatísticos em forma univariada, isto é, utilizando a geoestatística tradicional, construindo os mapas temáticos por krigagem ordinária para cada camada em estudo. Foi analisada a RSP na camada de 0,2-0,3 m de profundidade por meio de um modelo espacial linear (SLM), em que se consideraram as camadas de 0,0-0,1 m e 0,1-0,2 m como covariáveis, obtendo um modelo de estimação e um mapa temático por krigagem universal. Os mapas temáticos da RSP da camada de 0,2-0,3 m de profundidade, construídos por ambos os métodos, foram comparados por meio de medidas de acurácia obtidas a partir da construção da matriz de erros e da matriz de confusão. Verificou-se semelhança entre os mapas temáticos. Todos os mapas mostraram que a RSP é mais alta na região norte.

PALAVRAS-CHAVE: geoestatística, máxima verossimilhança, matriz de erro.

INTRODUCTION

The intensive cultivation of the soil and the use of machinery and heavy equipment lead to degradation of the physical conditions and, consequently, increased soil compaction. WEIRICH NETO et al. (2006) and LIMA et al. (2010) indicate the moto mechanization as one of the main causes of soil compaction, which may hinder the penetration of roots. In this context the soil resistance to penetration (RSP) is a good indicator of soil compaction for being property directly related to plant growth and quick and easy determination. A compacted soil adversely affects the growth of roots, and consequently, may decrease productivity in the area. COLLARES et al. (2008) found that the presence of compacted areas on the soil restricts root growth at that point and the

---

[1] Matemática, Mestre em Estatística, Doutoranda em Estatística, UFPE, Recife-PE, Grupo de Pesquisa GGEA, Fone: (0xx81) 2126-7692. fernandadebastiani@yahoo.com.br.
[2] Estatístico, Dr. Estatística, Grupo de Pesquisa GGEA, Pesquisador de Produtividade do CNPq, Prof. Associado da PGEAGRI, UNIOESTE, Cascavel- PR, Fone: (0xx45) 3220-3228. miguel.opazo@unioeste.br.
[3] Matemático, Mestre em Engenharia Agrícola, Grupo de Pesquisa GGEA, Prof. Assistente, UTFPR, Toledo-PR, Fone: (0xx45) 3379-6800. gustavodalposso@utfpr.edu.br.

roots are induced to grow in another direction. OLIVEIRA et al. (2011) studied the relation between soybean yield and soil chemical properties for the selection of statistical models and found that there is autocorrelation and cross-correlation between soil resistance to penetration in the first layers and soybean yield measured over the transversal space, and that in the scenario considered by the authors, the soil bulk density is not autocorrelated or have cross-correlation with soybean yield.

Although the literature presents critical values of RSP in which occurs mechanical impedance to root development of plants, there has been disagreements to these values, considering the soil type and species studied. In this context, CANARACHE (1994) proposed a model to show that the penetration resistance values vary under different conditions, such as combinations of textures, densities, and moisture, and from this model generally concluded that the values of RSP in the interval [1.10; 2.59] MPa have few limitations to root growth, while values in the range [2.6; 5.0] MPa indicate that there are some limitations.

The objective of this reseach was to study the spatial variability of soil resistance to penetration retreat in the layers of 0.0–0.1 m, 0.1–0.2 m and 0.2–0.3 m deep, using geostatistical methods in univariate analysis (traditional) and using a spatial linear model (SLM) of the RSP in the layer 0.2–0.3 m depth as covariables with the RSP of the layers of 0.0–0.1 m, 0.1–0.2 m. To compare the thematic maps of the RSP layer of 0.2–0.3 m of both methods, it was calculated measures of accuracy obtained from the error matrix and confusion matrix.

MATERIAL AND METHODS

Data collection was accomplished in the year 2007/2008 in a commercial producer of grains with 101.47 hectares located in the city of Cascavel, state of Paraná, Brazil, with approximate geographic coordinates of Latitude 24º57'21"S and Longitude 53º34'32"W, and an average altitude of 650 m. The climate presents as mild mesothermal and super humid, Cfa - Köeppen type climate, with moderate temperatures, well-distributed rainfall and hot summer. In winter, the average temperature is below 16°C, subject to frost, and in summer the maximum exceeds 30ºC and the average annual temperature is 21°C.

The planting system used was no-till and the soybean cultivar VMax NK412113 was planted in the north, and in the rest of the area the soybean cultivar CD213 was planted. In soil, which is classified as Hapludox, clayey, and presents historical succession of oat yield in the winter and soybean in the summer, it was performed a systematic sampling centered on pairs of nearby points (*lattice plus close pairs*), with maximum distance of 141 m between points and in some random places the sampling was performed with distances of 75 and 50 m between points. 63 sample elements were used in each layer of the study (0.0–0.1 m, 0.1–0.2 m and 0.2–0.3 m deep) to 101.47 ha, and the location of sampling points was performed with a GPS GEOEXPLORE 3, a Universal Transverse Mercatur (UTM) spatial coordinate system, with an accuracy of five meters.

The RSP in layers studied was measured with a penetrometer SC-60 SOILCONTROL brand, with shaft of 600 mm and 9.53 mm in diameter, equipped with a cone at the tip of 129.3 mm^2 base area, 12.83 mm in diameter and 30 degrees of angle vertex, four replications per location were made, taken at random in each plot, setting an average value of soil mechanical resistance to penetration. The layers 0.0–0.1 m, 0.1–0.2 m and 0.2–0.3 m depth had mean humidity of 12.74%, 14.70% and 15.17%, respectively, considering that for the three layers the coefficient of variation did not exceed 13%, characterizing homogeneity of observations of humidity in the region (GOMES, 2009).

To model the spatial structure with RSP, it was considered a Gaussian stochastic process $\{Z(s), s \in S\}$, $S \subset \Re^2$, $\Re^2$ a two-dimensional Euclidean space. Suppose that the data, $Z(s_1),...,Z(s_n)$, from this process are recorded in known spatial locations, s_i $(i = 1,...,n)$, and generated by the model $Z(s_i) = \mu(s_i) + e(s_i)$. In this model, the deterministic $\mu(s_i)$ and stochastic $e(s_i)$ terms may depend upon the spatial location where $Z(s_i)$ was obtained. It is assumed that the stochastic $e(.)$ error

has zero mean, $E[e(s_i)] = 0$, and the variation between points in the space is determined by a function covariance $COV[e(s_i), e(s_u)] = C(s_i, s_u)$, and for some known functions of s, such as $x_1(s),...,x_p(s)$, the mean of the process is $\mu(s_i) = \sum_{u=1}^{p} x_u(s_i)\beta_u$, and is called spatial linear models (SLM), in which $\beta_1,...,\beta_p$ are unknown parameters and to be estimated.

Equivalently, in matrix notation, it is understood that the SLM is:

$$Z = X\beta + \varepsilon, \tag{01}$$

In which the vector of random errors ε has $E(\varepsilon) = \mathbf{0}$ (zero vector) and covariance matrix $\Sigma = [(\sigma_{iu})]$, with $\sigma_{iu} = C(s_i, s_u)$. It is assumed that the covariance matrix Σ is not singular, that X is a nxp matrix of full rank, $\beta = (\beta_1,...,\beta_p)^T$ and Z follows a normal distribution n-variate with mean vector $X\beta$ and covariance matrix Σ, i.e., $Z \sim N_n(X\beta, \Sigma)$.

Given the parametric form of the covariance matrix,

$$\Sigma = \varphi_1 I_n + \varphi_2 R(\varphi_3), \tag{02}$$

In which, φ_1 is the nugget effect or error variance; φ_2 is the contribution or dispersion variance (sill); $R(\varphi_3)$ is a nxn matrix which is a function of φ_3, $R(\varphi_3)=[(r_{iu})]$ is symmetrical with its diagonal elements $r_{ii} = 1$, to $i=1,...,n$; φ_3 is a function of range (a) of the model and I_n is a nxn identity matrix. The parametric form of the covariance matrix is isotropic for various processes (GUEDES et al., 2008), where the covariance $C(s_i,s_u)$ is defined according to the covariance function $C(h_{iu})=\varphi_2 r_{iu}$, in which $h_{iu} =||s_i - s_u||$ is the Euclidean distance between the points s_i and s_u. In the covariance functions $C(h_{iu})$, the variance of Z is $C(0) = \varphi_1 + \varphi_2$, and semi variance may be defined as $\gamma(h)= C(0) - C(h)$ (DIGGLE & RIBEIRO JR, 2007).

To identify the structure of spatial dependency between the sampling elements, it was used the classic Matheron semivariogram, as defined in Equation (3)

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)]^2. \tag{03}$$

Choosing an appropriate model is to obtain estimators of the parameter vector $\delta =(\beta^T, \theta^T)^T$, in which $\beta =(\beta_1,...,\beta_p)^T$ and $\theta =(\varphi_1, \varphi_2, \varphi_3)^T$ by the methods of parameter estimation for maximum likelihood (ML) and maximum restricted likelihood (MRL). To estimate the parameter vector δ, it was chosen a vector $\hat{\delta} =(\hat{\beta}^T, \hat{\theta}^T)^T$ which maximizes the likelihood function in the Θ field, in which Θ is the parameter space. Considering the stochastic process $Z =(Z(s_1),...,Z(s_n))^T$, in which $Z \sim N_n(X\beta, \Sigma)$, the method of ML estimation of δ consists in maximizing the logarithm of the likelihood function

$$l(\delta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(Z - X\beta)^T \Sigma^{-1}(Z - X\beta)$$

The MRL method used to estimate the parameters of the covariance matrix Σ, consists in obtaining less biased estimators than the estimators obtained by ML. The method of MRL estimation of θ which consists in maximizing the logarithm of the restricted likelihood function (BORSSOI et al., 2009),

$$l(\theta) = -\frac{(n-p)}{2}\log(2\pi) - \frac{1}{2}\log|X^TX| - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\log|X^T\Sigma^{-1}X| - \frac{1}{2}Z^TPZ,$$

in which, $P = \Sigma^{-1}(I - A)$, with $A = X(X^T \Sigma^{-1}X)^{-1} X^T \Sigma^{-1}$.

To choose the model that best fits the data; it was used the cross-validation technique and the maximum value of the logarithm of the likelihood function (FARACO et al., 2008, JOHANN et al., 2010; BORSSOI et al., 2011). After the selection, it was used the vector of parameters $\hat{\delta}$ for the universal kriging interpolation. The comparison between the univariate map (traditional method), called the reference map, and the map built by SLM, called model map, of the RSP in the layer 0.2–0.3 m depth was performed using measurements obtained from the error matrix (Table 1) which has as measurement unit the *pixel*. Each matrix element represents the total number of *pixels* or the area belonging to the class *i* of the model map and to the class *j* of the reference map, being *N* the total number of *pixels* or the total area.

TABLE 1. Error matrix of the reference map in relation to the model map.

Model Map	Reference Map				
	1	2	…	r	Total
1	x_{11}	x_{12}	…	x_{1r}	$\sum_{j=1}^{r} x_{1j}$
⋮	⋮	⋮	⋮	⋱	⋮	⋮
r	x_{r1}	x_{r2}	…	x_{rr}	$\sum_{j=1}^{r} x_{rj}$
Total	$\sum_{i=1}^{r} x_{i1}$	$\sum_{i=1}^{r} x_{i2}$	…	$\sum_{i=1}^{r} x_{ir}$	N

The main diagonal is the case in which the *pixels* (or areas) had the same classification in the two maps, while the elements outside the main diagonal represent the classifications that do not coincide. The global accuracy (*GA*), $GA = \sum_{i=1}^{r} x_{ii} / N$, is a measurement used to measure the similarity between the reference map and the model map and according to ANDERSON et al. (2001), the minimum level of accuracy is 0.85. It is possible to build a range of (1-α) % confidence for the form *GA*, $IC[GA,(1-\alpha)\%] = GA \pm z_{\alpha/2}\sqrt{GA(1-GA)/N}$ for *N>50*, in which $z_{\alpha/2}$ is the nominal value of a random variable with standard normal distribution (FOODY, 2009).

The Kappa index (*K*) (BAZZI et al., 2008, FOODY, 2009; FOODY, 2010) has been used to measure the accuracy of thematic classifications, recommended as an appropriate measure of accuracy for using all elements of the error matrix. Provides a measure of agreement between the values of the reference map and the values from the model map, being defined by Equation (4),

$$\hat{K} = \frac{[N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+}x_{+i})]}{[N^2 - \sum_{i=1}^{r}(x_{i+}x_{+i})]} \tag{04}$$

in which, $x_{i+} = \sum_{j=1}^{r} x_{ij}$ is the number of *pixels* (or area) of the row *i*, $x_{+i} = \sum_{u=1}^{r} x_{ui}$ is the number of *pixels* (or area) of the column *i*, x_{ii} is the number of pixels (or area) of the row *i* and column *i*. According to the classification of KRIPENDORFF (1980), *K* is classified with low accuracy if *K* < 0.67, average accuracy if 0.67 ≤ *K* < 0.80, and high accuracy if *K*≥ 0.80. The variance of the Kappa index is obtained using Equation (5),

$$\sigma^2(\hat{K}) = \frac{1}{N}\left(\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2)^2}{(1-\theta_2)^4}\right) \tag{05}$$

In which, $\theta_1 = \frac{1}{N}\sum_{i=1}^{r} x_{ii}$; $\theta_2 = \frac{1}{N^2}\sum_{i=1}^{r} x_{i+}x_{+i}$; $\theta_3 = \frac{1}{N^2}\sum_{i=1}^{r} x_{ii}(x_{i+} + x_{+i})$; $\theta_4 = \frac{1}{N^3}\sum_{i=1}^{r}\sum_{j=1}^{r} x_{ij}(x_{i+} + x_{+j})^2$.

The range of (1-α)% confidence for K is from the form $IC[K,(1-\alpha)\%]= \hat{K} \pm z_{\alpha/2} \sqrt{\sigma^2(\hat{K})}$.

MA & REDMOND (1995), proposed an index that shows less subjective and easier to understand and use, the concordance index Tau (T), also known as Kappa modified, defined by $T=(\theta_1-p_i)/(1-p_i)$, being pi the *a priori* probability for each class i. When the *a priori* probabilities for the classes are the same, there is $p_i =1/r$, in which r represents the number of classes of the error matrix. The index Tau can follow the same classification of Kappa. The variance of Tau index is calculated with the same formula of variation of Kappa, and the difference is the inclusion of *a priori* probability. Another way to compare thematic maps is by setting up the confusion matrix by class, which is obtained by the elements of the error matrix, presented in Table 2.

TABLE 2. Confusion matrix of the i-th class (i=1,...,r).

	Reference Map	
	Present (+)	Absent (-)
Model Map Present (+)	a_i	b_i
Absent (-)	c_i	d_i

In Table 2, $a_i = x_{ii}$ is the amount of *pixels* (or area) of a class i of the reference map which was correctly classified as belonging to class i on model map (true positive - TP_i); $b_i = x_{i+} - x_{ii}$ is the amount of *pixels* (or area) that does not belong to class i of the reference map and was classified as belonging to class i in the model map (false positive - FP_i); $c_i = x_{+i} - x_{ii}$ is the number of pixels (or area) of a class i of the reference map and which belongs to a different class in the model map (false negative - FN_i) and $d_i = N - (a_i+b_i+c_i)$ is the amount of *pixels* (or area) that does not belong to class i in the reference map and was classified as not belonging to class i on model map (true negative - TN_i). TP_i and TN_i values are correct predictions in each class. The values FP_i and FN_i are considered estimation errors, FP_i known as errors of commission or overestimation, and FN_i known as errors of omission, for $i = 1, ..., r$.

Table 3 shows measurements of the confusion matrix per class. The sensitivity index (S_i) is a measure indicating the probability that a unit area on the model map is classified as belonging to class i ($i = 1,..., r$) if it really belongs to the class i of the reference map. The specificity index (E_i) indicates the probability that a unit area does not belong to the class i of the reference map or to the class i of the model map. The errors of commission (FPR_i) indicate the proportion of area units that do not belong to the class i of the reference map and that belong to class i of the model map. Errors of omission are considered serious as it indicates the proportion of area that belongs to the class i on reference map and has been classified in another class in the model map.

TABLE 3. Measures obtained from the confusion matrix for the i-th class ($i =1,..., r$).

Measures	Equatio
Sensitivity index (S_i)	$a_i/(a_i + c_i)$
Specificity index (E_i)	$d_i/(b_i + d_i)$
False positive rate (FPR_i) (error of commission) $(1 - E_i)$	$b_i/(b_i + d_i)$
False negative rate (FNR_i) (error of omission) $(1 - S_i)$	$c_i/(a_i + c_i)$
Positive estimation power (PEP_i)	$a_i/(a_i + b_i)$
Negative estimation power (NEP_i)	$d_i/(c_i + d_i)$
Conditional Kappa User (CKU_i)	$(UA_i- p_{+i})/(1- p_{+i})$
Conditional Kappa Producer (CKP_i)	$(PA_i- p_{i+})/(1- p_{i+})$
Conditional Tau User (CTU_i)	$(UA_i- 1/r)/(1 - 1/r)$
Conditional Tau Producer (CTP_i)	$(PA_i- 1/r)/(1 - 1/r)$

$p_{+i=} x_{+i}/N; \ p_{i+=} x_{i+}/N; \ AU_i= x_{ii}/x_{i+}; \ AP_i= x_{ii}/x_{+i}$.

The negative estimation power (NEP_i) indicates the probability of a unit area not to be classified by both reference map and model map since it was classified as absent by the model map. Positive estimation power (PEP_i) indicates the probability of a unit area to be classified by both reference map and model map since it was classified as present by the model map. The conditional Kappa index (CKU_i and CKP_i) is presented as an additional means of incorporation of hit by hypothetical chance in the evaluation of the accuracy per class, and the conditional Tau index (CTU_i and CTP_i) is an alternative to this, for the Kappa may provide wrong conditional estimates for not knowing the distribution of reference data in advance. The values of UA_i and PA_i, called user accuracy and producer accuracy, are given respectively by $\dfrac{x_{ii}}{x_{i+}}$ and $\dfrac{x_{ii}}{x_{+i}}$. Based on measurements in Table 3, it is possible to compare for each class the map reference with the model map. For a global comparison of the maps, JENNESS & WYNNE (2005) present the global confusion matrix, shown in Table 4.

TABLE 4. Global confusion matrix.

		Reference Map	
		Presente(+)	Absent (–)
Model Map	Present (+)	$a = \sum\limits_{i=1}^{r} x_{ii}$	$b = \sum\limits_{i=1}^{r}\sum\limits_{j\neq i}^{r} x_{ij}$
	Absent (–)	$c = \sum\limits_{j=1}^{r}\sum\limits_{i\neq j}^{r} x_{ij}$	$d = \sum\limits_{i=1}^{r}\sum\limits_{i\neq j}^{r}\sum\limits_{j\neq i}^{r} x_{ij}$

Table 5 presents measurements of global confusion matrix. The interpretation for the global indexes of total sensitivity (S) and total specificity (E) is analogous to the versions by class. The values of Matthews correlation coefficient (MCC) are within the range [-1, 1], and the value 1 corresponds to a perfect prediction, the value 0 corresponds to a random prediction and value -1 corresponds to an inverse prediction. The average mutual information AMI (FINN, 1993) calculates the conditional probability that an area on a map belongs to a particular class given the class of this area on the second map, thereby providing means to measure the similarity between maps with different themes.

TABLE 5. Measures derived from the global confusion matrix.

Measurements	Equation
S	$a /(a + c)$
E	$d /(b + d)$
MCC	$[(a{\cdot}d) - (b{\cdot}c)] /[(a + b){\cdot}(a + c){\cdot}(d + b){\cdot}(d + c)]^{1/2}$
AMI	$\sum\limits_{i=1}^{r}\sum\limits_{j=1}^{r} p_{ij} \log(p_{ij} / p_{i+} p_{+i})$

$p_{+i=} x_{+i}/N$; $p_{i+=} x_{i+}/N$; $p_{ij=} x_{ij}/N$

For data analysis, it was used the software R (R DEVELOPMENT CORE TEAM, 2009) and its modules geoR (DIGGLE & RIBEIRO JR, 2007).

RESULTS AND DISCUSSION

Table 6 presents descriptive statistics of the data of RSP [MPa] in the different layers. According to the average values, it is verified that there is little restriction in the growth of roots (CANARACHE, 1994). According to the coefficient of variation (CV) there is an average data homogeneity of RSP [MPa] in three layers, for 10% < CV < 20% (GOMES, 2009). According to the coefficient of asymmetry and kurtosis, the RSP in the three layers have characteristics of normal

distribution, according to criteria shown in JONES (1969). In this study no evidence sample was discarded.

TABLE 6. Descriptive statistics of the RSP [Mpa] in the three layers of depth.

Statistics	RSP		
	0.0-0.1 m	0.1-0.2 m	0.2-0.3 m
Sample size	63	63	63
Average	2.554	2.255	1.979
Median	2.410	2.220	1.910
Q1	2.215	2.035	1.770
Q3	2.935	2.430	2.090
Minimum	1.580	1.670	1.470
Maximum	3.600	3.240	3.090
Standard Deviation	0.507	0.336	0.332
CV%	19.84	14.91	16.77
Coefic. Asymmetric	0.281	0.797	1.598
Coefic. Kurtosis	- 0.788	0.671	3.093

Q1: first quartile, Q3: third quartile, CV: coefficient of variation.

Table 7 presents the results of univariate geostatistical analysis with the models adjusted and the estimated parameters of the RSP for the layers under study. According to the cross-validation criteria and the maximum value of the logarithm of the likelihood function (*MLL*), for the layer of 0.0 – 0.1 m depth the model which best fit was the Gaussian with parameters estimated by ML method. To the layers 0.1-0.2m and 0.2 – 0.3 m, the model which best fit was the Gaussian with parameters estimated by MRL. The ranges (*a*) obtained represent the distance at which the sampling points are correlated, ranging from 1274.35 m to 1381.60 m, not exceeding the maximum distance of 1710.26 m. Analyzing the coefficient of relative nugget effect (*RNE*), according to the scale of CAMBARDELLA et. al (1994), we found a moderate spatial dependence (25% < *RNE* ≤75%) for the RSP in the layers of 0.0 – 0.1 m and 0.1 – 0.2 m deep and weak spatial dependence (*RNE* > 75%) to the RSP in the layer of 0.2 – 0.3 m deep.

TABLE 7. Adjusted spatial models and the estimated parameters of RSP[MPa].

RSP	Method	Model	$\hat{\beta}$	$\hat{\varphi}_1$	$\hat{\varphi}_2$	$\hat{\varphi}_1+\hat{\varphi}_2$	$\hat{\varphi}_3$	$\hat{a}=g(\hat{\varphi}_3)$	*RNE*	*LML*
0,0-0,1 m	ML	Gaus	2.63	0.162	0.158	0.320	776.40	1344.77	50.51	-38.66
0,1-0,2 m	MRL	Gaus	2.29	0.087	0.077	0.164	797.67	1381.60	53.04	-16.03
0,2-0,3 m	MRL	Gaus	1.98	0.095	0.024	0.119	735.75	1274.35	79.71	-17.45

$\hat{\beta}$: mean; $\hat{\varphi}_1$: nugget effect; $\hat{\varphi}_2$: contribution; ($\hat{\varphi}_1 + \hat{\varphi}_2$): parameter; $\hat{\varphi}_3$: parameter associated with range; $\hat{a}$: range; $RNE = (\hat{\varphi}_1/\hat{\varphi}_1 + \hat{\varphi}_2$))x100 relative nugget effect; *LML*: maximum value of the logarithm of the likelihood function; ML: maximum likelihood; MRL: maximum restricted likelihood; Gaus: Gaussian model.

Table 8 presents the results of cross validation for the adjusted models for the three layers in the study. Note that the values of the mean error (*ME*) and reduced error (*RE*) are close to zero. The standard deviation of the mean error (S_{ME}) and absolute error (*AE*) are the smallest among the models studied. And yet, the standard deviation of the reduced error (S_{RE}) is very close to one, as expected (FARACO et al, 2008).

TABLE 8. Cross-validation criterion for the adjusted models.

Variables	Method	Model	*ME*	*RE*	S_{ME}	S_{RE}	*AE*
RSP 0.0-0.1 m	ML	Gaus	0.0014	0.0016	0.4253	1.0061	21.23
RSP 0.1-0.2 m	RML	Gaus	0.0002	0.0004	0.3090	1.0052	15.23
RSP 0.2-0.3 m	RML	Gaus	-0.0004	-0.0007	0.3198	1.0049	14.19

ME: mean error; *RE*: reduced error; S_{ME}: standard deviation of mean error; S_{RE}: standard deviation of reduced error; *AE*: absolute error

Figure 1 presents the theoretical semivariograms which were best adapted to the experimental semivariograms for each layer, according to the criteria for validation of models studied.
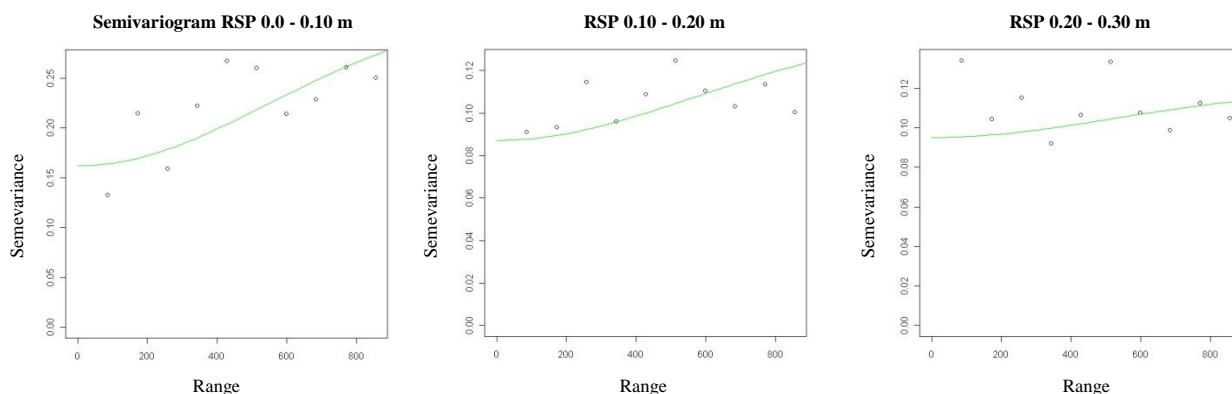

FIGURE 1. Theoretical semivariogram adapted to experimental semivariograms a) RSP 0.0 – 0.10 m b) RSP 0.10 – 020 m c) RSP 0.20 – 0.30 m.

Figure 2 presents the thematic maps of RSP built by ordinary kriging. There was a decrease in soil compaction in accordance with increasing depth. Figure (2a) shows the thematic map of the RSP in the layer of 0.0 – 0.1 m deep, where 17.86% of the area has RSP greater than 2.88 MPa, 18.36% between 2.61 and 2.88 MPa, 42.36% between 2.34 and 2.60 MPa, and 21.42% between 2.07 and 2.33 MPa. Figure (2b) shows the map of the RSP in the layer of 0.1 – 0.2 m deep, in which 30.07% of the area has RSP between 2.34 and 2.60 MPa, 65.09% between 2.07 and 2.33 MPa, and 4.84% less than 2.07 MPa. Figure (2c) shows the map of RSP in the layer of 0.2 – 0.3 m deep, and 26.69% of the area has RSP between 2.07 and 2.33 MPa, and the remaining 73.31% presents RSP less than 2.07 MPa.
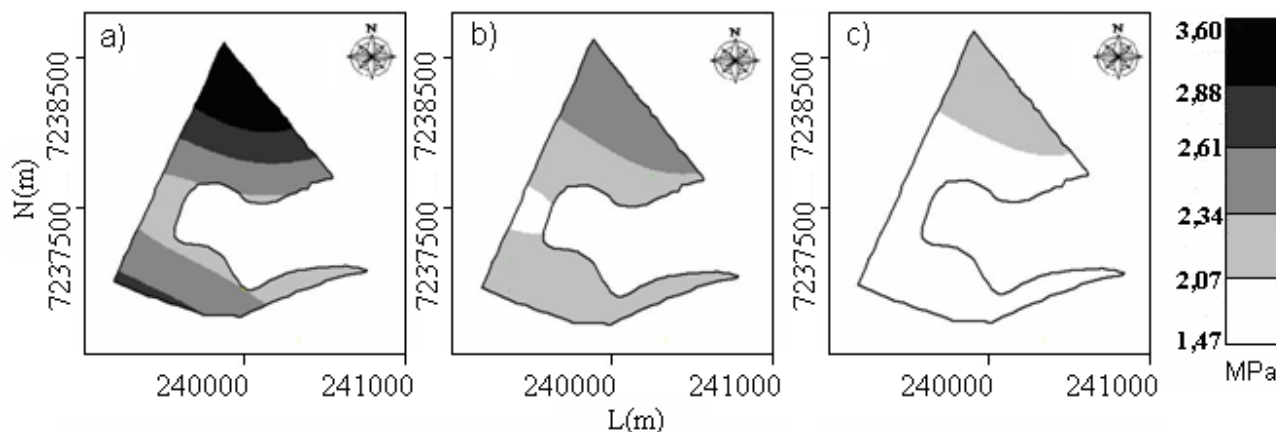

FIGURE 2. Thematic maps of the RSP [MPa] (UTM coordinates, zone 22 and datum WGS 84) in the layers: a) 0.0 – 0.1 m, b) 0.1 – 0.2 m, and c) 0.2 – 0.3 m deep.

Although there are questions about values of soil resistance to penetration considered critical to crop development, CARVALHO et al. (2006) reported that RSP values ranging between 1.3 and 2.9 MPa did not limit grain yield of the bean in an Oxisol. Considering the conditions of this study and the criteria for CANARACHE (1994), it appears that for the three layers, the RSP is higher in the North region, and 36.22% of the soil has limited the roots growth up to 0.1 m deep, featuring a compacted soil. It should be noted that the layers samples showed little change in water content of the soil.

In the analysis using spatial linear model (SLM) the interest was to study the spatial variability of the mean $\mu_{23}(s_i)$ of the RSP in the layer 0.2 – 0.3 m deep as a function of covariables

X_2 (RSP in the layers from 0 – 0.1 m) and X_3 (RSP in the layers 0.1 – 0.2 m). The model that best fit was the Gaussian by the ML according to the cross-validation criteria and the maximum value of the logarithm of the likelihood function, LML = 16.66. The SLM obtained for the mean of the RSP in the layer of 0.2 – 0.3 m deep was: $\hat{\mu}_{23}(s) = 0{,}1607 - 0{,}0613X_2(s) + 0{,}8757X_3(s)$, with spatial dependence parameters obtained of the covariance matrix Σ, in which $\hat{\varphi}_1$ =0.0326; $\hat{\varphi}_2$ =0.0019 e $\hat{\varphi}_3$ =12.694, with a range estimative of 21.99 on relative nugget effect (RNE) equal to 94.5%, characterizing weak spatial dependence. Table 9 shows the values obtained for the cross-validation criteria for the adjusted model with covariables, which followed the same criteria for analyzing the values shown in Table 8.

TABLE 9. Cross-validation criterion for the adjusted models with covariables.

Variable	Method	Model	*ME*	*RE*	S_{ME}	S_{RE}	*AE*
RSP 0.0 – 0.1 m	ML	Gaus	0.0021	0.0055	0.1981	1.0364	10.28

ME: mean error*; RE:* reduced error*; S_{ME}:* standard deviation of mean error*; S_{RE}:* standard deviation of reduced error*; AE:* absolute error

The thematic map of Figure 3 shows the RSP to be higher in the North region, but the soil has no limitation to root growth.
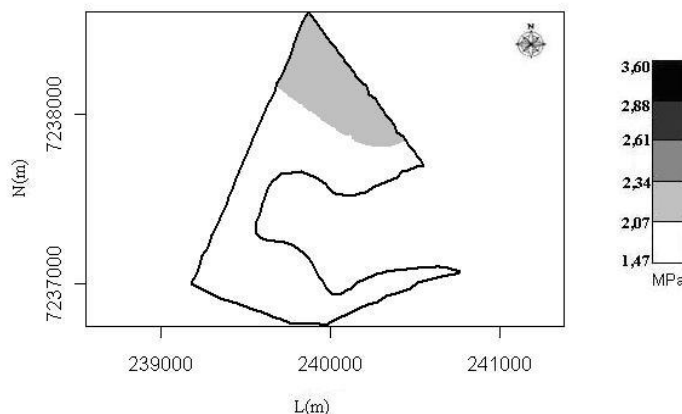

FIGURE 3. Thematic map of RSP [MPa] at 0.2 – 0.3 m according to the preceding two layers.

Table 10 presents the error matrix associated with the RSP maps in the layer 0.2 – 0.3 m deep, in which it was considered as a reference map the univariate map obtained by ordinary kriging and as a model map the map obtained by the SLM with universal kriging.

TABLE 10. RSP error matrix (0.2 – 0.3 m) of the number of pixels and area (ha) by class.

	RSP [MPa] Class	Referrence Map (univariant) [1.47–2.06]	[2.07–2.34]	Total
Model Map (SLM)	[1.47– 2.06]	11864 (74.29)	31 (0.19)	11895 (74.48)
	[2.07– 2.34]	330 (2.07)	3979 (24.92)	4309 (26.99)
	Total	12194 (76.36)	4010 (25.11)	16204 (101.47)

In parenthesis is the value of area in ha.

It is noteworthy that both maps were generated with 16,204 pixels, which means that each pixel represents an area of approximately 62.62 m^2.

Table 11 presents the estimates of the accuracy measures and their corresponding 95% confidence interval for the comparison of the maps. The index of global accuracy (GA) obtained

was 0.98, indicating an acceptable precision according to ANDERSON et al (2001). According to the classification of KRIPENDORFF (1980), Kappa (0.94) and Tau (0.95) indexes are regarded as high accuracy. Thus, it can be seen that there are few differences in the classification of the thematic maps studied.

TABLE 11. RSP accuracy measures in the layer 0.2-0.3 m deep.

Measures	Results	Confidence Interval	Values
Global Accuracy –GA	0.978	IC[EG.95%]	[0.975; 0.982]
Kappa -K	0.942	IC[K. 95%]	[0.933; 0.950]
Tau - T	0.955	IC[T. 95%]	[0.952; 0.958]

Table 12 presents the confusion matrix by class of RSP in layer 0.2-0.3 m.

TABLE 12. Confusion matrix by class of RSP in layer of 0.2-0.3 m deep.

Classes	[1.47– 2.06]	Present Map (+)	Reference Absent (-)	[2.07 – 2.33]	Present Map (+)	Reference Absent (-)
Model Map (SLM)	Present (+)	11864	31	Present (+)	3979	330
	Absent (-)	330	3979	Absent (-)	31	11864

Table 13 presents measures of accuracy calculated by class of RSP in the layer of 0.2 – 0.3 m deep. In class [1.47 – 2.06] MPa, the sensitivity index S_1= 0.973 is the probability that a pixel belonging to this class in the two kinds of maps constructed, the specificity index E_1 = 0.992 is the probability a pixel does not belong to this class in both types of maps constructed. The omission error (FNR_1) indicates 2.7% of the pixels belonging to this class of the reference map, belongs to a class different from the model map. The commission error (FPR_1) indicates 0.8% of the pixels that do not belong to this class in the reference map, belongs to this class in the model map. The positive estimation power (PEP_1) indicates 0.997 and is the probability of a pixel to be classified by both the reference map and the model map since it was classified as present by the model map. The negative estimation power (NEP_1) is 0.923 indicating the probability that a pixel not to be classified by neither the reference map nor the model map since it has been classified as absent by the model map.

In class [2.07 – 2.34] MPa, the sensitivity index S_2 = 0.992, is the probability of a pixel to belong to this class of both types of maps constructed, the specificity index E_2 = 0.973 is the probability a pixel to not belong to this class in both types of maps constructed. The omission error (FNR_2) indicates 0.8% of the pixels belonging to this class of the reference map, belongs to another class different from the model map. The commission error (FPR_2) indicates 2.7% of pixels that do not belong to this class in the reference map, belong to this class in the model map. The index PEP_2 indicates 0.923 is the probability that a pixel is classified by both the reference map and the model map since it was classified as present by the model map. The index PEN_2 indicates that 0.997 is the probability that a pixel not to be classified by both the reference map and the model map since it has been classified as absent by the model map.

TABLE 13. Accuracy measures by class of RSP [MPa] in layer 0.2-0.3 m.

Measures by class	C1	C2		C1	C2
Sensitivity index (S_i)	0.973	0.992	Negative estimation power (NEP_i)	0.923	0.997
Specificity index (E_i)	0.992	0.973	Conditional Kappa Producer (CKP_i)	0.898	0.989
Omission error (FNR_i)	0.027	0.008	Conditional Kappa User (CKU_i)	0.989	0.898
Comission error (FPR_i)	0.008	0.027	Conditional Tau Producer (CTP_i)	0.946	0.989
Positive estimation power (PEP_i)	0.997	0.923	Conditional Tau User (CTU_i)	0.995	0.847

C1: RSP [1.47 – 2.06];   C2: RSP [2.07 – 2.34].

The global confusion matrix for the RSP in the layer of 0.2-0.3 m deep presented 15,843 pixels classified as (+) (+), 15,843 as (-) (-) and 361 as (+) (-) and (-) (+). The global indexes of sensitivity (*S*), specificity (*E*), Matthews correlation coefficient (M*CC*) and Average Mutual Information (*AMI*) were 0.978, 0.978, 0.955 and 0.455, respectively. Thus, the sensitivity (*S*), which is the conditional probability that a pixel in the reference map belongs to a model map is 0.978. The specificity index (*E*) of 0.978 indicates the probability of not belonging to the reference map and to the model map, the Matthews correlation coefficient (M*CC* = 0.955) shows that there is a direct proportional correlation in the classification of thematic maps. The index of average mutual information *AMI* = 0.455 indicates the similarity between the reference map and the model map.

CONCLUSIONS

By geostatistical methods, it was possible to detect that the RSP in three layers showed higher degrees of compression in the North region. Through the thematic maps constructed, it was possible to determine the areas where it is necessary to apply techniques to reduce the problem of soil compaction. The modeling by SLM of RSP in the layer 0.2-0.3 m deep is very important to estimate the RSP in terms of previous layers in sectors not sampled in the experimental area. The accuracies measures presented allowed realization of the comparison between the univariate thematic map using ordinary kriging and the thematic map using the SLM with universal kriging. The index values showed the existence of similarity between the constructions of both maps.

REFERENCES

ANDERSON, J.F.; HARDY, E.E.; ROACH J.T.; WITMER R.E. A land use and land cover classification system for use with remote sensor data. U.S. Geological Survey Professional Paper 964, Washington: U.S. Geologic Survey, 2001. 41p. (Paper 964).

BAZZI, C.L.; SOUSA, E.G.; URIBE-OPAZO, M. A.; NÓBREGA, L.H.P.; PINHEIRO R. Influência da distância entre passadas de colhedora equipada com monitor de colheita na precisão dos mapas de produtividade na cultura do milho. *Engenharia Agrícola*, Botucatu, v.28, n.2, p. 355-363, 2008.

BORSSOI, J. A.; URIBE-OPAZO, M. A.; GALEA, M. Diagnostics tecniques applied in geoestatistics for agricultural data analysis. *Revista Brasileira de Ciência do Solo*, Viçosa-MG, v.6, n.6, p. 1-16, 2009.

BORSSOI, J. A.; URIBE-OPAZO, M. A.; GALEA, M. Técnicas de Diagnóstico de Influência Local na Análise Espacial da Produtividade da Soja. *Engenharia Agrícola*, Jaboticabal, v.1, n.2, p. 376-387, 2011.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, I.M.; PARKIN, T.B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa Soils. *Soil Science Society of America Journal*, Madison, v.58, n.2, p. 1501-1511, 1994.

CANARACHE, A. PENETR – a generalized semi-empirical model estimating soil resistance to penetration. *Soil and Tillage Research*, Amsterdam, v.16, n.1, p. 51-70, 1994.

CARVALHO, G.J.; CARVALHO, M. de P.; FREDDI, O. da S.; MARTINS, M.V. Correlação da produtividade do feijão com a resistência à penetração do solo sob plantio direto. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, v.10, n.3, p.765-771, 2006.

COLLARES, G. L.; REINERT, D. J.; REICHERT, J. M.; KAISER, D. Compactação de um Latossolo induzida pelo tráfego de máquinas e sua relação com o crescimento e produtividade de feijão e trigo. *Revista Brasileira de Ciência do Solo*, Viçosa-MG, v.32, n.3, p. 933-942, 2008.

DIGGLE, P. J.; RIBEIRO JUNIOR, P. J. *Model based geoestatistics*. New York :Springer Series in Statistics, 2007. 228p.

FARACO, M.A.; URIBE-OPAZO, M. A.; SILVA, E. A.; JOHANN, J. A.; BORSSOI, J. A. Seleção de modelos de variabilidade espacial para elaboração de mapas temáticos de atributos físicos do solo e produtividade da soja. *Revista Brasileira de Ciência do Solo*, Viçosa-MG, v. 32, n.2, p. 463-476, 2008.

FINN, J.T. Use of the average mutual information index in evaluating classification error and consistency. *International Journal of Geographical Information Society,* London, v.7, p. 349–366, 1993.

FOODY, G.M. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sensing of Environment*, New York, v.113, p. 1658-1663, 2009.

FOODY, G.M. Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment,* New York, v. 114, p. 2271-2285, 2010.

GOMES, P. F. *Curso de estatística experimental*. 15 ed. Piracicaba: Degaspari, 2009. 451p.

GUEDES, L. P. C.; URIBE-OPAZO, M. A.; JOHANN, J. A.; SOUZA, E. G. Anisotropia no estudo da variabilidade espacial de algumas variáveis químicas do solo. *Revista Brasileira de Ciência do Solo*, Viçosa-MG, v.32, n.6, p. 2217-2226, 2008.

JENNESS, J.; WYNNE, J. J. *Cohen's Kappa and classification table metrics 2.0*: an ArcView 3x extension for accuracy assessment of spatially explicit models: U.S. Geological Survey Open-File Report OF 2005-1363. Flagstaff, Southwest Biological Science Center, 2005.

JOHANN, J.A.; SILVA, M. C. A.; URIBE-OPAZO, M. A.; DALPOSSO, G.H. Variabilidade Espacial da rentabilidade, perdas na colheita e produtividade do Feijoeiro. *Engenharia Agrícola*, Jaboticabal, v.30, n.4, p.700-714, 2010.

JONES, T.A. Skewness and kurtosis as criteria of normality in observed frequency distributions. *Journal Sedimentary Petrology*, Salida, v.39, n.4, p. 1622-1627, 1969.

KRIPPENDORFF, K. Content analysis: an introduction to its methodology. Beverly Hills: Sage Publications. 1980. 189p.

LIMA, C. L. R.; REINERT, D. J.; REICHERT, J. M.; SUZUKI, L. E. A. S. Produtividade de culturas e resistência à penetração de Argissolo Vermelho sob diferentes manejos. *Pesquisa Agropecuária Brasileira*, Brasília, v.45, n.1, p. 89-98, 2010.

MA, Z.; REDMOND, R. L. Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*, Bethesda, v. 61, n.4, p.453-439. 1995.

OLIVEIRA, M. P.; TAVARES, M. H. F.; URIBE-OPAZO M. A.; TIMM, L. C. Selecting statistical models to study the relationship between soybean yield and soil physical properties. *Revista Brasileira de Ciência do Solo,* Viçosa-MG, v.35, n.1, p. 97-104, 2011.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Disponível em: <http://www.R-project.org>. 2009.

WEIRICH NETO, P. H.; BORGHI, E.; SVERZUT, C. B.; MANTOVANI, E. C.; GOMIDE, R. L.; NEWES, W. L. C. Análise multivariada da resistência do solo à penetração sob plantio direto. *Ciência Rural*, Santa Maria, v.36, n.4, p. 1186-1192, 2006.