

Spatial Modeling of Soil Lime Requirements with Uncertainty Assessment Using Geostatistical Sequential Indicator Simulation

Jussara de Oliveira Ortiz¹, Carlos Alberto Felgueiras¹, Eduardo Celso Gerbi Camargo¹, Camilo Daleles Rennó¹, Manoel Jimenez Ortiz²

¹INPE-Brazilian National Institute for Space Researches DPI-Divisão de Processamento de Imagens, São José dos Campos, Brazil

²Geopixel Soluções em Geotecnologias e TI, São José dos Campos, Brazil

Email: jussara@dpi.inpe.br, carlos@dpi.inpe.br, camilo@dpi.inpe.br, eduardo@dpi.inpe.br, manoel.ortiz@geopixel.com.br

How to cite this paper: de Ortiz, J.O., Felgueiras, C.A., Camargo, E.C.G., Rennó, C.D. and Ortiz, M.J. (2017) Spatial Modeling of Soil Lime Requirements with Uncertainty Assessment Using Geostatistical Sequential Indicator Simulation. *Open Journal of Soil Science*, 7, 133-148.

<https://doi.org/10.4236/ojss.2017.77011>

Received: April 10, 2017

Accepted: July 10, 2017

Published: July 13, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This work presents and analyses a geostatistical methodology for spatial modelling of Soil Lime Requirements (*SLR*) considering punctual samples of Cation Exchange Capacity (*CEC*) and Base Saturation (*BS*) soil properties. Geostatistical Sequential Indicator Simulation is used to draw realizations from the joint uncertainty distributions of the *CEC* and the *BS* input variables. The joint distributions are accomplished applying the Principal Component Analyses (*PCA*) approach. The Monte Carlo method for handling error propagations is used to obtain realization values of the *SLR* model which are considered to compute and store statistics from the output uncertainty model. From these statistics, it is obtained predictions and uncertainty maps that represent the spatial variation of the output variable and the propagated uncertainty respectively. Therefore, the prediction map of the output model is qualified with uncertainty information that should be used on decision making activities related to the planning and management of environmental phenomena. The proposed methodology for *SLR* modelling presented in this article is illustrated using *CEC* and *BS* input sample sets obtained in a farm located in Ponta Grossa city, Paraná state, Brazil.

Keywords

Spatial Modeling of Soil Attributes, Indicator Geostatistics, Joint Simulation, Principal Component Analyses, Spatial Uncertainty Analyses

1. Introduction

Soil acidity is one of the factors that limits crop yields in various places of the world. Brazilian soils are mostly acids, mainly for the tropical Savanna regions,

known in Brazil as Cerrado. Such soils are characterized by low concentrations of calcium and magnesium, elements directly involved in the development of the plant [1].

The appropriate correction of soil acidity, or liming, is considered an effective practice for the use of soil nutrients by plants [2], which promoting increased soil fertility and hence productivity. For this purpose, the limestone is an agricultural input with efficient response to correct soil acidity. It is relatively cheap in Brazil and it is of simple application. The literature shows a limestone relation to the increase in production, involving the raising of productivity of grain crops, mainly soybeans, wheat and corn, with proper fertilization, or replacement of nutrients in the soil [3] [4] [5] [6].

However, any application of high or insufficient doses of inputs in the soil will reflect in plant nutrition so that, if not corrected by cover fertilizations, will increase or decrease the productivity. Therefore, the recommendation of the amount of limestone, like any other fertilizer, must comply with a soil analysis, which should avoid unnecessary applications that would lead to super liming. According to some authors [7] [8], super liming would be as damaging as high acidity and it would be difficult to correct.

To obtain higher yields and to apply inputs in the soil with no waste, there is a tendency to use methods of input estimates based on spatial variability of soil properties. These estimates lead to application of inputs, by variable rates in the geographic space, with the purpose of optimizing profit, productivity and environmental sustainability. These practices, known as Precision Agriculture (PA), implement the process of agricultural automation, dosing fertilizers and pesticides for the soil of a geographical area of interest. This set of agriculture tools involves the use of Geodesic Positioning Systems (GPS), Geographic Information Systems (GIS) with integrated statistics methods, and instruments with sensors for measurement or detection of parameters on targets of interest in the agroecosystem (soil, plant, insects and diseases).

Several authors have adopted the geostatistical procedures in PA to estimate and evaluate the spatial variability of soil properties [9] [10] [11] [12] [13]. Specifically, indicator geostatistical procedures have been widely used because they are non-parametric methods, *i.e.*, they do not require the definition of a priori probability model [14] [15] [16]. In addition, predictions maps (mean, median or mode value), and uncertainty maps (based on deviation values or probability of default quantiles) are extracted by inference and simulation methods. So, the predictions are accompanied with their respective uncertainties, which are also spatially distributed in the study area. Take into account the uncertainty information during the modeling process is important, because it allows to qualify the result of the used model, which should be considering for planning and decision making related to the investigated properties [17] [18] [19].

In this context, this work applies and analyses indicator geostatistical procedures for spatial modelling of Soil Lime Requirements (*SLR*), derived from a model proposed by [20]. The *SLR* model is based on the relationship between

Cation Exchange Capacity (*CEC*) and Base Saturation (*BS*) properties. These soil chemical properties affect soil acidity, and subsequently soil fertility and yield. So, the idea of the *SLR* modelling is to show the soil regions that need correction, raising the saturation of the bases to a value that provides maximum economic efficiency of the limestone application.

This work is organized in 6 sections. Besides the Section 1, presenting the introduction, the Section 2 addresses the main necessary theoretical concepts. Section 3 synthesizes the proposed methodology. The Section 4 explores a case study of *SLR* modeling with real data. Section 5 presents the results, followed of analyses and discussions. Finally, the Section 6 reports important conclusions with suggestions for future works involving research aspects of this work.

2. Concepts

2.1. The Spatial *SLR* Model

Liming depends on the decision to apply or not the limestone and the definition of its quantities, if it is required. This is done through some formulated mathematically methods and they are related to the soil characteristics of each geographic region.

Among the several methods for liming recommendation, the base saturation method, presented by [20] was defined as the most suitable for the study region. It is based on the relationship between *BS* and *CEC*. The model parameters are considered for the soil, soil amendment and the crop. The method consists in raising the saturation of the bases to a value that provides maximum economic efficiency of the of limestone application.

The neutralizing power, the granulometry and reactivity of the limestone are important factors for the correct choice of it. The neutralizing power is determined by comparison with the power of neutralization of pure calcium carbonate (CaCO_3), with the maximum value equal 100. For this reason, it is called the Relative Power of Total Neutralization (*RPTN*) or calcium carbonate equivalent, and the knowledge of this parameter is relevant to determine the *SLR*. The model to determine *SLR*, in tons/hectare (t/ha), according to recommendations about soils presented in [6] is expressed as:

$$SLR(\mathbf{u}) = \frac{(BS_2 - BS_1(\mathbf{u})) \cdot CEC(\mathbf{u})}{RPTN} \quad (1)$$

where: \mathbf{u} is a vector of the geographic coordinates where the input variables are sampled inside of the study area; $BS_1(\mathbf{u})$ is the base saturation value of the original soil, given in percentage, before correction. They are sampled directly from the ground, analyzed in the laboratory and should be elevated to the level considered suitable for the crop and soil studied; BS_2 is the base saturation value that you want to achieve; there is a default value for each crop; $CEC(\mathbf{u})$ is the sum of the bases with the values of potential acidity; *RPTN*: is the Relative Power of the Total Neutralization in relation to the limestone adopted and it is, generally, less than 100 (the total neutralization).

According to [21] the BS_2 is variable for each State or region and, as defined before, at Parana State this value is 70% and $RPTN$ is the Relative Power of the Total Neutralization in relation to the limestone adopted and it is, generally, less than 100%, that is the total neutralization; The limestone that has been used at the study farm presents a $RPTN$ equivalent to 85%. So, the model to calculate Limestone Requirement is:

$$SLR(\mathbf{u}) = \frac{(70 - BS(\mathbf{u})) * CEC(\mathbf{u})}{850} \tag{2}$$

The spatial modeling applied to the investigation of a phenomenon of interest requires mathematical models that work on estimate values of the input variables and their uncertainties. The output estimates only depend on the input variables while the output uncertainties are propagated from the uncertainties of the input variables. The output uncertainties are used to qualify the spatial model output. Knowing the quality of the model results is important, especially when they are used in decision making activities associated to the planning and management of the investigated phenomena [22].

When management planning requires local estimates and mathematical models are considered, the uncertainty propagated in the predictions might be evaluated. Knowing the quality of the model results is fundamental, especially when they are used in spatial decision making [22], with GIS operations.

Some uncertainty propagation techniques were presented by [23]. The first order Taylor series and the Monte Carlo simulation can be stand out.

The Monte Carlo simulations were considered in this work. Let $U(.)$ be the output of a GIS operation $g(.)$ on m input attributes

$A_i(.) : U(.) = g(A_1(.), \dots, A_m(.))$. The idea of the method is to compute the result on $g(a_1, \dots, a_m)$ repeatedly, with input values a_i that are randomly sampled from their joint distribution.

Application of the Monte Carlo method to uncertainty propagation with non-point operations requires the simultaneous generation of realizations from $A_i(.)$. This implies that spatial correlation will have to be accounted for. The technique adopted in this paper for stochastic spatial simulation, was the joint sequential Indicator simulation algorithm [14] with principal component analysis, as presented in the next item.

The idea of the Taylor series method is to approximate $g(.)$ by a truncated Taylor series centred at $\bar{b} = (b_1, \dots, b_m)$ [23].

The first order Taylor series of $g(.)$ around \bar{b} is given by:

$$U = g(\bar{b}) + \sum (A_i - b_i) g'_i(\bar{b}) \tag{3}$$

where $g'_i(.)$ is the first derivative of $g(.)$ with respect to its i -th argument. The variance σ^2 is given as:

$$\sigma^2 = \sum \sum \rho_{ij} \sigma_i \sigma_j g'_i(\bar{b}) g'_j(\bar{b}) \tag{4}$$

where: σ_i and σ_j are the standard deviation of the i and j variables respectively and ρ_{ij} is the correlation between them.

As the *SLR* model considers only two variables, the Equation (4) will be as bellow:

$$\sigma^2 = \sum_{i=1}^2 \sum_{j=1}^2 \rho_{ij} \sigma_i \sigma_j g'_i(\bar{b}) g'_j(\bar{b}) \tag{5}$$

i represents the variable *BS* and *j* represents the variable *CEC*.

And when Equation (5) is applied on the model as,

$$g = SLR = \frac{(70 - BS) * CEC}{850} \tag{6}$$

the derivatives are as bellow:

$$\frac{\partial g}{\partial BS} = \frac{-CEC}{850} \quad \text{and} \quad \frac{\partial g}{\partial CEC} = \frac{70 - BS}{850} \tag{7}$$

2.2. Principal Component Transformation

Many spatial models that make use of the several input variables required the determination of joint probability distributions of these variables for generation of the new information and for uncertainty evaluation, that usually are built via simulation methods.

When the input variables are correlated, their distributions should not be simulated independently. The direct approach is to use a joint simulation of the dependent variables, but it requires the inference and modeling of direct and cross covariance matrices that are computational costly to determine [14].

An alternative mode is first to decorrelate the input variables using Principal Component Analysis-*PCA* [24] [25] and, after apply the geostatistical simulation procedure. So, the *M* interdependent input variables, denoted by $Z_M(\mathbf{u}) = \{Z_1(\mathbf{u}), Z_2(\mathbf{u}), \dots, Z_m(\mathbf{u})\}$, $m = 1, \dots, M$, are transformed in *M* independent variables, denoted by $Y_M(\mathbf{u}) = \{Y_1(\mathbf{u}), Y_2(\mathbf{u}), \dots, Y_m(\mathbf{u})\}$, as follows:

$$Y_M(\mathbf{u}) = \varphi(Z_M(\mathbf{u})) \tag{8}$$

where φ is the transformation function of the *PCA*.

To recover the interdependent variables $Z_M(\mathbf{u})$ applies the inverse transform function, as follows:

$$Z_M(\mathbf{u}) = \varphi^{-1}(Y_M(\mathbf{u})) \tag{9}$$

where φ^{-1} is the inverse transformation of the *PCA*.

2.3. Indicator Geostatistical Approach

The indicator geostatistical approach for continuous variables, for example for the $Y_M(\mathbf{u})$ variables, allows to estimate a set of values at non sampled location, that represent a discretized approximation of the conditional cumulative distribution function (*ccdf*) [15] [26].

Initially the variables $Y_M(\mathbf{u})$ are transformed into indicator variables, considering y_M^c cutoffs values, $c = 1, \dots, n$ cutoffs. These transformations are defined by the relation:

$$I_M(\mathbf{u}; y_M^c) = \begin{cases} 1 & \text{se } Y_M(\mathbf{u}) \leq y_M^c \\ 0 & \text{se } Y_M(\mathbf{u}) > y_M^c \end{cases} \tag{10}$$

The transformation expressed in Equation (8) is equivalent to associate probability 1 (100%) for $Y_M(\mathbf{u})$ values which are smaller than or equal to the cutoff y_M^c and 0 otherwise. So, for each cutoff value indicator fields are generated, with 0 and 1 values, of the indicator variable $I_M(\mathbf{u}; y_M^c)$.

Next, empirical indicator semivariograms are determined for each one of the indicator fields to estimated their spatial correlation structures:

$$\hat{\gamma}_M(\mathbf{h}, y_M^c) = \frac{1}{2N(\mathbf{h})} \sum_{j=1}^{N(\mathbf{h})} [i(\mathbf{u}; y_M^c) - i(\mathbf{u} + \mathbf{h}; y_M^c)]^2 \tag{11}$$

where $i(\mathbf{u}; y_M^c)$ and $i(\mathbf{u} + \mathbf{h}; y_M^c)$ are values of the indicator variable $I_M(\mathbf{u}; y_M^c)$ separated by the distance vector \mathbf{h} , and $N(\mathbf{h})$ is the number of the pairs points that are separated by the distance vector \mathbf{h} .

The indicator fields associated with their respective cutoff values and theoretical semivariograms are used by Indicator Kriging (IK) and Conditional Sequential Indicator Simulation (SIS) procedures for estimating probabilities values at non-sampled location. Thus, this set of estimated probabilities, at non sampled location, is used to create a discretized approximation of the *ccdf* [15] [27]) and it represents the uncertainty model of the variable.

2.4. Conditional Sequential Indicator Simulation

A stochastic simulation is a process of drawing equally likely realizations of values that are obtained from the probability distribution of a Random Variable (RV).

Consider the M independent factors $Y_M(\mathbf{u})$. Then, a set of the realizations equally likely are generated by geostatistical simulation, denoted by $Y_M^L(\mathbf{u}_\alpha)$, where L is the simulation number; $\mathbf{u}_\alpha = 1, \dots$, gridsize (nlines x ncolums), are locations regularly distributed in the geographic space, determining a regular grid representation structure. For example, to the $Y_2(\mathbf{u})$ independent variable, $Y_2^L(\mathbf{u}_\alpha) = \{y_2^l(\mathbf{u}_\alpha)\}$, where $l = 1, \dots, L$; $y_2^l(\mathbf{u}_\alpha)$ is the l -th simulated value at \mathbf{u}_α ; and the set $\{y_2^l(\mathbf{u}_\alpha)\}$ represent the l -th random field generated by simulation of the $Y_2(\mathbf{u})$. In that case, the SIS procedures makes use of the *ccdf*, conditioned to the n nearest of RV $Y_2(\mathbf{u})$ and also to the pre-simulated values, inside the \mathbf{u}_α neighborhood, to get $y_2^l(\mathbf{u}_\alpha)$ values.

The SIS procedure has the following steps [15]:

1) Sets up randomly one \mathbf{u}_α location. Then simulates a value of the $y_2^l(\mathbf{u}_\alpha)$ from the univariate *ccdf* of $Y_2(\mathbf{u}_\alpha)$, $\text{Prob}[Y_2(\mathbf{u}_\alpha) \leq y_2(\mathbf{u}_\alpha) | n]$, conditioned to the n nearest sample data $Y_2(\mathbf{u})$;

2) Once simulated, $y_2^l(\mathbf{u}_\alpha)$ becomes a conditioning data for subsequent simulation steps. The conditioning data is updated to

$$(n+1) = (n) \cup \{Y_2(\mathbf{u}) = y_2^l(\mathbf{u}_\alpha)\};$$

3) Sets up randomly another location $\mathbf{u}'_\alpha \neq \mathbf{u}_\alpha$. Simulated a new value of the $y_2^l(\mathbf{u}'_\alpha)$ from the univariate *ccdf* of $Y_2(\mathbf{u}'_\alpha)$, $\text{Prob}[Y_2(\mathbf{u}'_\alpha) \leq y_2(\mathbf{u}'_\alpha) | (n+1)]$, conditioned to the information set $(n+1)$;

4) Update the information set $(n+1)$ to a new information set

$$(n+2) = (n+1) \cup \{Y_2(\mathbf{u}) = y_2(\mathbf{u}'_\alpha)\};$$

- 5) Repeat the two previous steps until all the locations \mathbf{u}_α of the spatial grid have been simulated. At this time was produced a random field of the $Y_2(\mathbf{u})$;
- 6) Repeat all steps for generating new random fields of the $Y_2(\mathbf{u})$. Stop when you reach the desired number of simulations.

3. Methodology

The proposed methodology for *SLR* modelling is illustrated in the **Figure 1**, using $BS(\mathbf{u})$ and $CEC(\mathbf{u})$ input sample sets obtained in a farm located in Ponta Grossa city of Paraná State, Brazil. The variables $BS(\mathbf{u})$ and $CEC(\mathbf{u})$, hereafter

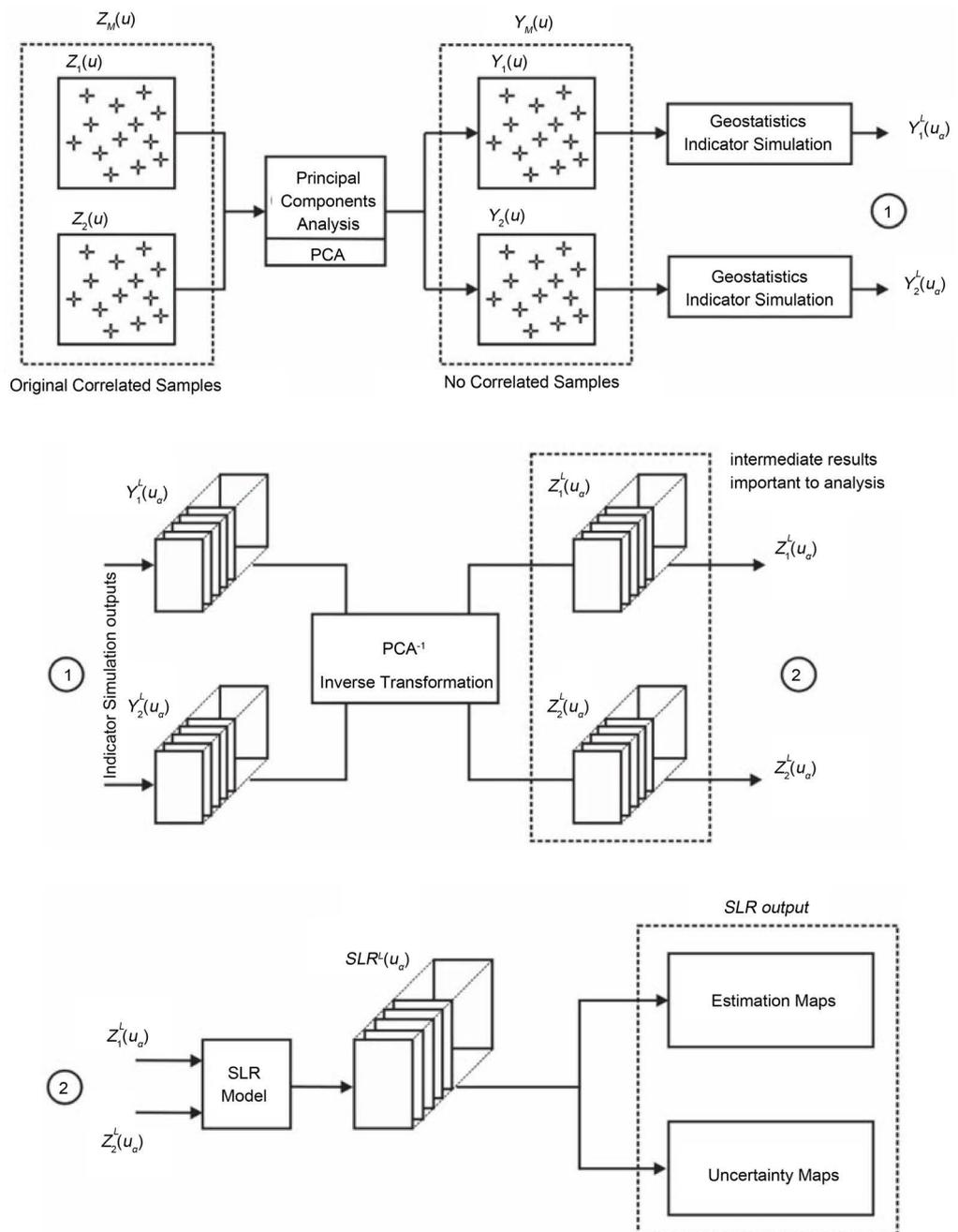


Figure 1. Methodological sequence applied to $CEC(\mathbf{u})$ and $BS(\mathbf{u})$ soil attributes to model *SLR*.

are named in this Section as $Z_1(\mathbf{u})$ and $Z_2(\mathbf{u})$, respectively.

Initially, the input variables $Z_1(\mathbf{u})$ and $Z_2(\mathbf{u})$ are decorrelated via *PCA* transformation in order to get two uncorrelated sample sets named $Y_1(\mathbf{u})$ and $Y_2(\mathbf{u})$ variables. Following, $Y_1(\mathbf{u})$ and $Y_2(\mathbf{u})$ are independently spatialized applying SIS procedure. The result is a set of the realizations equally likely, $Y_1^L(\mathbf{u}_\alpha)$ and $Y_2^L(\mathbf{u}_\alpha)$. Since the simulation values were obtained from the independent variables distributions [$Y_1(\mathbf{u})$ and $Y_2(\mathbf{u})$], it is necessary to apply to them PCA^{-1} , the inverse transformation of *PCA*, in order to obtain a set of dependent realizations, $Z_1^L(\mathbf{u}_\alpha)$ and $Z_2^L(\mathbf{u}_\alpha)$. From these realizations, prediction and uncertainty maps can be generated. These maps are intermediate results that help the analyst to better understand the spatial distribution of the variables involved in the *SLR* model and their respective uncertainties. So, the simulated fields $Z_1^L(\mathbf{u}_\alpha)$ and $Z_2^L(\mathbf{u}_\alpha)$ are selected randomly several times, via Monte Carlo method and after applied in the spatial *SLR* model. From the resulting *SLR* fields, prediction and uncertainty maps are generated. These maps represent the spatial variation of the output variable and the propagated uncertainty, respectively. Therefore, the prediction map generated by spatial *SRL* model is qualified with uncertainty information

Given a spatial region of interest, the methodology applied has the following steps:

- 1) In a set of spatial correlated sample points of the *CEC* and *BS* soil attributes, hereafter named in this text Z_1 and Z_2 variables, apply the *PCA* transformation in order to get two uncorrelated variables Y_1 and Y_2 ;
- 2) Apply the indicator sequential simulation on the Y_1 and Y_2 variables in order to get independent grids representing fields of draw values of these variables;
- 3) The draw values of Y_1 and Y_2 are back transformed using a PCA^{-1} approach resulting on a set of dependent grids representing fields of draw values of the Z_1 and Z_2 variables;
- 4) Extract statistic properties of the *CEC* and *BS* draw values: prediction maps of mean or median values, for example, and uncertainty maps based in confidence interval of standard deviations or quantiles. This is important to observe the individual spatial distribution of the estimated and uncertainty values of the input variables;
- 5) Obtain draw values, randomly chosen, from the *SLR* model applied to the dependent draw values of the Z_1 and Z_2 variables;
- 6) Extract statistic properties of the *SLR* draw values: prediction maps of mean or median values, for example, and uncertainty maps based in confidence interval of standard deviations or quantiles.

In this article the final resulting maps of predictions and uncertainties are analyzed and. **Figure 1** illustrates the applied methodology.

4. A Case Study

4.1. The Study Area and Input Data

The study region, Figueira farm, is located in the city of Carambeí, PR, Brazil,

and comprises an area of 392 ha. The farm has been adopted a precision agriculture system and the main cultures are soy, wheat and corn. In order to illustrate the methodology of this work, it was used as information a set of points of soil chemical properties (*BS* and *CEC*). These data were sampled at field by a Brazilian company that works with precision agriculture [28]. The geographical positioning of the samples was gathered by Global Positioning System, GPS, which ensures accurate georeferencing, as shown in the **Figure 2**.

5. Results and Discussion

The results presented below were obtained from simulated fields of input variables of *CEC* and *BS*. These fields were produced using specific functions from geostatistical software called GSLIB [14]. As mentioned before the soil attributes, *CEC* and *BS* are interdependent and the correct spatial modeling requires the joint simulation of these variables.

In this study an alternative to joint simulation was implemented, and it was possible to simulate separately a set of independent factors from which the original variables can be reconstituted [14] [15]. The factors were obtained from Principal Component Analysis and the interdependence was guaranteed by the common inverse transform. The indicator simulation was performed over the

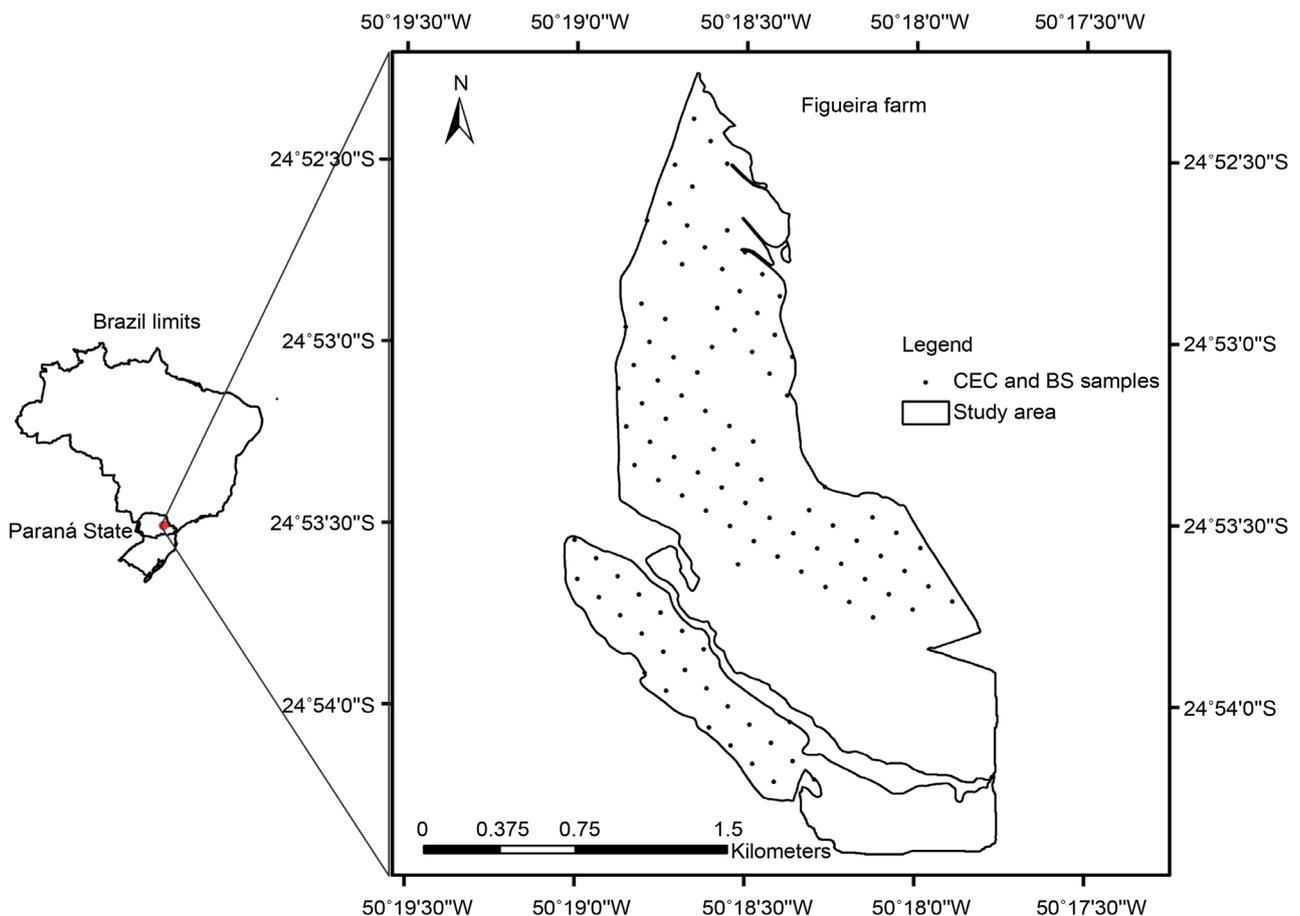


Figure 2. Study area Figueira farm in Paraná state.

independent sample set and simulated fields were generated. These independent simulated fields were back transformed into simulated values/fields for the original variables and they represent their stochastic uncertainty models.

In this case study, for each input variable, the number of realizations was fixed to 400 and they were represented in regular grid structures with 200 lines by 200 columns. That number of runs is considered sufficiently large to reach high accuracies. From those simulated values it could be generated the prediction and uncertainties maps. Thus, the mean estimates map was calculated by arithmetic average of the simulated values at each spatial position. Similarly, an uncertainty map was based on calculating the standard deviation of those same simulated values.

The resulting data were organized in a geographic database using the GIS SPRING [29]. This GIS enabled the visualization of spatial input information and the results of the spatial model explored.

Figure 3 shows the predicted values, grid of mean values, obtained from the uncertainty models of *CEC* and *BS* soil properties. From these maps that represent the outline of the study area, it can be observed the spatial distribution of each attribute, along with their minimum and maximum values.

Figure 4 depicts the uncertainty maps with confidence intervals based on standard deviation values of the uncertainty models of *CEC* and *BS* soil properties.

The maps of the **Figure 4** indicate that the highest uncertainties, red areas, appear in regions where higher local variations of the predicted values occur. This is an expected result since the uncertainties of the indicator simulation approach are proportional to the local density of the samples and to their local value variations. These maps represent the spatial distribution of the uncertainties where one can get, for instance, the general idea of the areas with the most

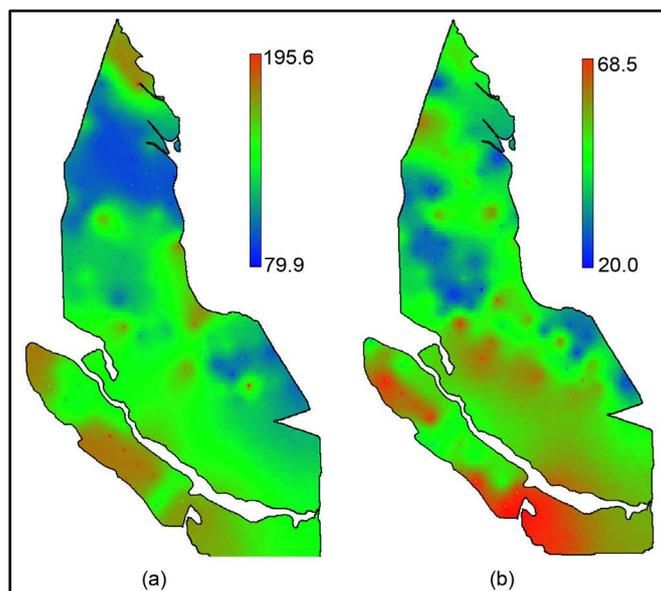


Figure 3. Map of predicted mean values from simulated fields of (a) *CEC* and (b) *BS*.

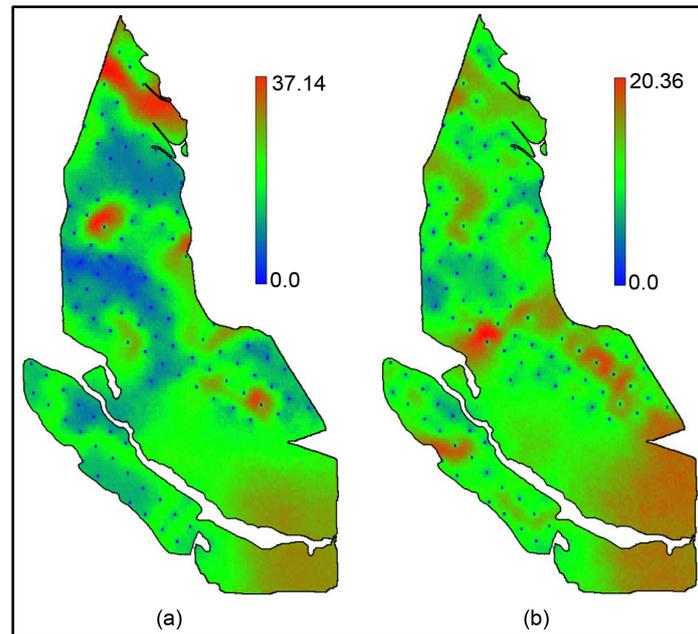


Figure 4. Map of uncertainty values, based on standard deviation confidence intervals, from simulated fields of (a) *CEC* and (b) *BS* soil properties.

reliable estimates. These uncertainty values are propagated to the resulting uncertainty maps of any spatial modelling that integrate these variables.

6. Modeling Results and Uncertainty Propagation

In the Figueira Farm the correction of soil acidity was modeled by estimating soil lime requirement, as a function of the *CEC* and *BS* that are the input variables of the model, *SLR*, presented in Equation (2). Those variables were simulated previously (Figure 3) and their uncertainty values (Figure 4) will propagate through these model, leading to uncertain response values of the *SLR*. Figure 5 presents the predicted and uncertain maps resulting of the *SLR* spatial modelling. These maps were generated from the application of the Monte Carlo approach on the back transformed fields, or the correlated simulated fields, of the *CEC* and *BS* variables. In this procedure 400 fields were generated from the final model obtained by applying Equation (2) in the simulated values of the input variables, randomly drawn. The simulated values from those fields, at each spatial location, were used to generate the maps presented in the Figure 5(a), mean of *SLR* and Figure 5(b) standard deviation of the *SLR*.

In the Figure 5(a) the map of means presents the spatial distribution of the mean values estimated of the *SLR*. That map shows that regions with high *SLR* values, in red, are those in which the *CEC* values are higher and *BS* is lower. The map of standard deviation, Figure 5(b), shows the spatial distribution of the uncertainty values propagated to final modeling result. As expected, at this uncertainty map are observed high uncertainty areas in red, where there is a greater spatial variability in the values inferred by the model.

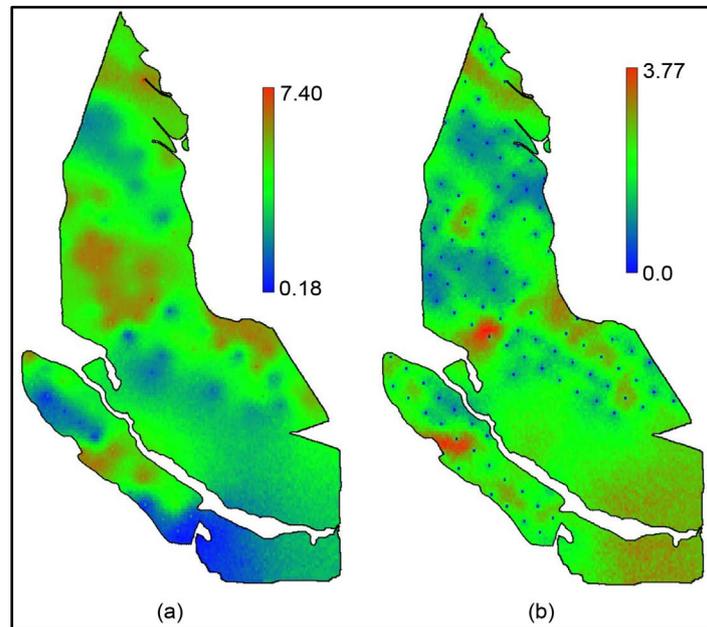


Figure 5. SLR results of (a) predicted mean (b) uncertainty values from joint simulation process.

7. Discussion

Figure 6(a) presents uncertainty map resulting from the *SLR* spatial modelling regarding to correlation between the variables *CEC* and *BS* to perform the joint simulation, and **Figure 6(b)** shows the uncertainty map considering independence between them. The maps are plotted in a same color scale (0 to 6.44) in order to be compared. Comparing the uncertainty maps of **Figure 6**, it can be observed that the uncertainty gets smaller for the *SLR* results when the correlation between the input variables is considered. In this case this correlation is positive and the final propagated result of the *SLR* model agrees with the concepts explained in the Section 2.5.

When the Taylor expression (Equation (6)) is applied on the model *SLR* (Equation (2)), replacing the first derivatives as presented in Equation (7), it is possible to evaluate that as higher correlation between the input variables as lower the uncertainty propagation in the final results. So this information can be associated with the results analysis.

Although most users are aware with uncertainty in GIS operations, they have not paid attention to the problem of correlated variables in spatial modelling. To emphasize this important issue we have considered three different correlations levels, including the measured, to apply the Taylor series: one of them is higher than the original measured correlation $\rho = 0.8$; the true correlation of the data, $\rho = 0.51$ and a lower correlation $\rho = 0.2$. The results, presented in the **Figure 7**, show that the propagated uncertainty decreased when the correlation increased.

The uncertainty maps are used to qualify the inferences. So, in decision-making processes, it can be considered just uncertainties values above or below of a settled threshold, giving priority to areas that are best suited to the problem that is

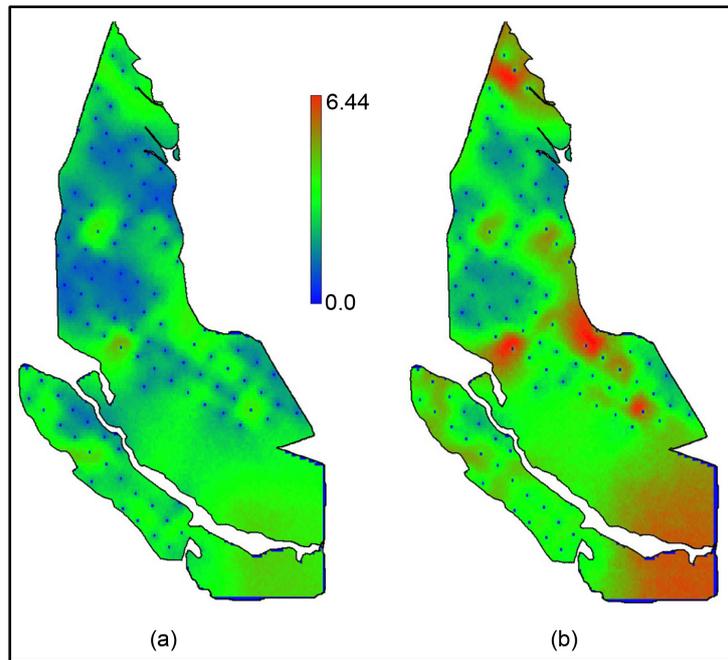


Figure 6. Uncertainty maps: (a) Joint simulation; (b) Independence between CEC and *BS*.

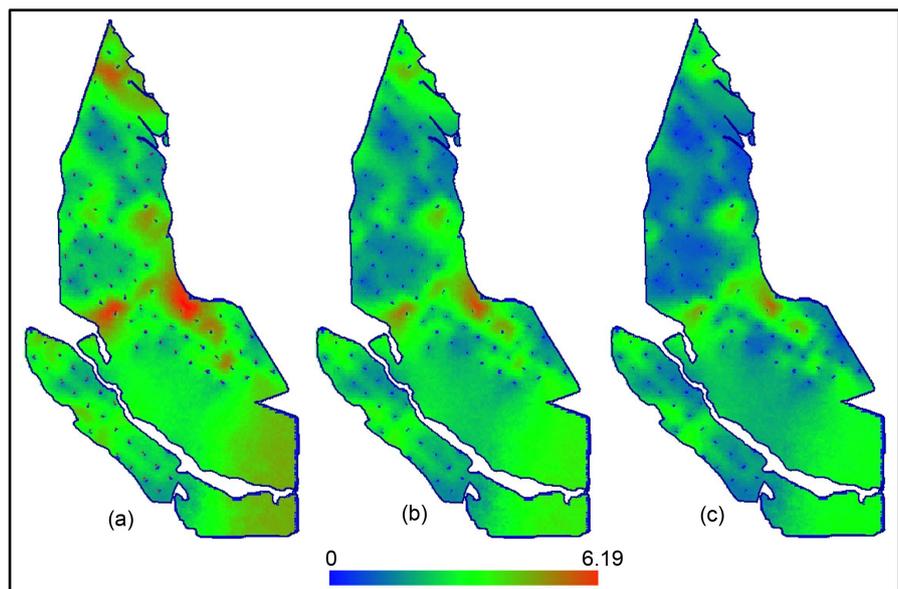


Figure 7. Uncertainty maps with Taylor series: (a) Correlation 0.8; (b) Correlation 0.5; and (c) Correlation 0.2.

being addressed. In this work, which is modeling the need for limestone, the decision maker can evaluate the available resources and define important thresholds, upper and lower, for the implementation of necessary agricultural inputs. It is important to consider that the excess of limestone can be as or more harmful than the lack of it for a particular crop. So, incorporate uncertainty maps to a precision farming system adds quality to the agricultural planning that requires information with spatial representation.

8. Conclusions

In order to make good decisions in precision agriculture, considering a field with acid soil, as this study case, the determination of the best liming practice depends on the uncertainty analysis. Spatial modelling as presented in this paper has to consider the level of uncertainty about the input parameters of the model, soil properties, and their propagation. Specifically, when those parameters are correlated their distributions cannot be sampled independently. So, this study was concerned about it and replaces the joint distribution to the principal component analysis.

In this work, we used the geostatistical simulation procedure for nomination to represent soil properties that are correlated. Applied an approach based on the analysis of principal components to obtain the achievements, with dependencies, these simulations. These achievements have been integrated via Monte Carlo method to obtain limestone need values according to a predefined spatial model.

The methodology for modeling the liming requirements was proved easy to implement and the results showed consistency with the spatial model and its input variables. In this modeling were obtained, besides the prediction maps, maps the uncertainties of the input variables and the final model that reflects the propagated uncertainty. He showed up at work, the propagated uncertainty is overrated if you do not consider different correlations of zero between model input variables. It is able to highlight areas with higher lime requirements and it is possible to identify areas with higher uncertainties in relation to the predictions. With those results the decision maker can, for instance, choose the first areas candidate for liming if there are financial problems. To reduce uncertainties is a very important point to policy makers who, many times, have to decide and to assess the consequences of different decisions, even decide for no action if the uncertainties are big.

Acknowledgements

The authors would like to thank FAPESP the São Paulo Research Foundation for the support obtained, via process number 2015/24676-9, to develop the researches that allows the realization of this article.

References

- [1] Pitta, G.V., Coelho, A.M., Alves, V.M.C., França, G.E. and Magalhães, J.V. (2006) Cultivo Do Milho. Embrapa Sorgo E Milho, versão eletrônica, Dez. <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/487015/1/Com53.pdf>
- [2] Caires, E.F., Garbuio, F.J., Churka, S., Barth, G. and Correa, J.C.L. (2008) Effects of Soil Acidity Amelioration by Surface Liming on No Till Corn, Soybean And Wheat Root Growth and Yield. *European Journal of Agronomy*, **28**, 57-64. <https://doi.org/10.1016/j.eja.2007.05.002>
- [3] Anderson, N.P., Hart, J.M., Sullivan, D.M., Christensen, N.W., Hornek, D.A. and Pirelli, G.J. (2013) Applying Lime to Raise Soil pH for Crop Production, Oregon State University. Extension Publication FG52, Corvallis, Oregon.

- <https://catalog.extension.orenstate.edu/em9057>
- [4] Mallarino, A., Pagani, A. and Sawyer, J.E. (2011) Corn and Soybean Response to Soil pH Level and Liming. *Integrated Crop Management Conference*, Ames, Iowa, 30 November-1 December 2011, 93-102.
- [5] Korndorfer, G.H., Pereira, H.S. and Camargo, M.S. (2004) Silicatos De Cálcio E Magnésio Na Agricultura. 3rd Edition, GPSi/ICIAG/UFU, Technical Report, Uberlândia, 28 p.
- [6] Fukuda, C. and Otsubo, A.A. (2003) Cultivos na Região Centro Sul do Brasil. Embrapa Solos, Sistema de Produção, Eletronic Version. <http://www.embrapa.br>
- [7] Padua, T., de Padua, P., Silva, C. and de Oliveira Dias, B. (2008) Nutrição E Crescimento Do Algodoeiro Em Latossolo Sob Diferentes Coberturas Vegetais E Manejo De Calagem. *Ciência e Agrotecnologia*, **32**, 1481-1490.
<https://doi.org/10.1590/S1413-70542008000500019>
- [8] Silva, P.A., Rigato, L.I., Jales, de Lima Jales, L., Polessa, M.F., Sather, M., Neto, R.B., Passos, J.L., de Souza, C.M., dos Santos, J.N.B. and do Amaral, J.A.T. (2013) Estudo Mineral De Uma Super Calagem No Milho. CCA-UFES/Fitotecnia, Alto universitário, Alegre, ES.
<http://biblioteca.univap.br/dados/INIC/cd/inic/IC5%20anais/IC5-13.PDF>
- [9] Vettorazzi, C.A. and Ferraz, S.F.B. (2000) Silvicultura De Precisão: Uma Nova Perspectiva Para O Gerenciamento De Atividades Florestais. In: Borém, A., Giudice, M.P., de Queiróz, D.M., *et al.*, Eds., *Agricultura De Precisão*, Os Autores, Viçosa, 65-75.
- [10] Ortiz, J.L. (2003) Emprego Do Geoprocessamento No Estudo Da Relação Entre Potencial Produtivo De Um Povoamento De Eucalipto E Atributos Do Solo E Do Relevo. Dissertação (Mestre em Recursos Florestais), ESALQ/USP, Piracicaba, SP.
- [11] Oliver, M.A. (2013) Precision Agriculture and Geostatistics: How to Manage Agriculture More Exactly. *The Royal Statistical Society*, **10**, 17-22.
- [12] Krasilnikov, P., Carré, F. and Montanarella, L. (2008) Soil Geography and Geostatistics Concepts and Applications. JRC, Institute for Environment and Sustainability, Luxembourg.
- [13] Qu, M., Li, W. and Zhang, C. (2013) Assessing the Spatial Uncertainty in Soil Nitrogen Mapping through Stochastic Simulations with Categorical Land Use Information. *Ecological Informatics*, **16**, 1-9. <https://doi.org/10.1016/j.ecoinf.2013.04.001>
- [14] Deutsch, C.V. and Journel, A.G. (1998) GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York.
- [15] Goovaerts, P. (1997) Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.
- [16] Felgueiras, C.A. (1999) Modelagem Ambiental Com Tratamento De Incertezas Em Sistemas De Informação Geográfica: O Paradigma Geoestatístico Por Indicação. 165p. Tese (Doutorado em Computação Aplicada), Instituto Nacional de Pesquisas Espaciais, São José dos Campos. <http://urlib.net/sid.inpe.br/deise/2001/08.03.12.35>
- [17] Viscarra-Rossel, A.R., Goovaerts, P. and McBratney, A.B. (2001) Assessment of the Production and Economic Risks of Site-Specific Liming Using Geostatistical Uncertainty Modeling. *Environmetrics*, **12**, 699-711. <https://doi.org/10.1002/env.471>
- [18] Searcy, S.W. (1997) Precision Farming: A new Approach to Crop Management. *Texas Agricultural Extension Service*. The Texas A&M University System, College station, Texas, 4. http://lubbock.tamu.edu/files/2011/10/precisionfarm_1.pdf
- [19] Ortiz, J.O (2008) Análise De Risco Na Fase De Palnejamento Em Sistemas De Produção Agrícola Por Meio De Simulação Geosestatística Condicionada. Tese de

Doutorado, INPE, São José dos Campos, SP, Brazil.

<http://urlib.net/8JMKD3MGP8W/33JR7HB>

- [20] Rajj, B.V. (1991) Fertilidade Do Solo E Adubação. Ceres, São Paulo, 343 p.
- [21] Quaggio, J.A. (2004) Manejo Da Acidez Do Solo E A Prática Da Calagem: Acidez E Calagem Em Solos Tropicais. Instituto Agronômico, Campinas-SP, Artigo técnico da Calpar-calcário agrícola, Paraná.
<http://www.calpar.com.br/calpar.com.br/calpar-site2015/artigos-boletim2.html>
- [22] Buttafuoco, G., Conforti, M., Robustelli, G. and Scariciglia, F. (2011) Assessing Spatial Uncertainty in Mapping Soil Erodibility Factor Using Geostatistical Stochastic Simulation. *Environmental Earth Science*, **66**, 1111-1125.
<https://doi.org/10.1007/s12665-011-1317-0>
- [23] Heuvelink, G.B.M. (1998) Error Propagation in Environmental Modeling with GIS. Taylor and Francis Inc., Bristol, 127 p.
- [24] Suro Perez, V.S. and Journel, A.G. (1991) Indicator Principal Component Kriging. *Mathematical Geology*, **23**, 759-793. <https://doi.org/10.1007/BF02082535>
- [25] Richards, J.A. (1996) Remote Sensing Digital Image Analysis: An Introduction, Chapter 6. Springer Verlag, Berlin, 280 p.
- [26] Journel, A.G. (1983) Nonparametric Estimation of Spatial Distributions. *Mathematical Geology*, **15**, 445-468. <https://doi.org/10.1007/BF01031292>
- [27] Goovaerts, P. (2001) Geostatistical Modelling of Uncertainty in Soil Science. *Geoderma*, **103**, 3-26. [https://doi.org/10.1016/S0016-7061\(01\)00067-2](https://doi.org/10.1016/S0016-7061(01)00067-2)
- [28] Impar-Consultoria e Assessoria Agrícola (2005). <http://www.imparag.com.br/>
- [29] Camara, G., Souza, R.C.M., Freitas, U.M and Garrido, J.G. (1996) SPRING: Integrating Remote Sensing and GIS by Object-Oriented Data Modelling. *Computers & Graphics*, **20**, 395-403. [https://doi.org/10.1016/0097-8493\(96\)00008-8](https://doi.org/10.1016/0097-8493(96)00008-8)



Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact ojss@scirp.org