Revista Brasileira de Fruticultura

Soil sampling optimization using spatial analysis in irrigated mango fields under brazilian semi-arid conditions

Kátia Araújo da Silva[1], Marcos Sales Rodrigues[2], Filipe Bernard Ramos Moreira[3],
Aíris Layanne Ferreira Lira[3], Augusto Miguel Nascimento Lima[2], Ítalo Herbert Lucena Cavalcante[2]

Abstract - Soil sampling is a fundamental procedure in the decision making regarding the management of the soil, thus, a sampling plan should represent as accurately as possible the evaluated crop field. Therefore, the objectives of this study were to suggest a soil sampling approach and soil sampling point allocation using spatial analyses and compare to the classic statistic method in irrigated mango orchards in the Brazilian semi-arid region. The experiment was carried out in three commercial mango orchards located in the region of the São Francisco Valley, Brazil. Soil samples were collected in 0-0.2 m and 0.2-0.4 m depths following regular grids where the number of samples varied from 50 to 56. Soil texture, soil bulk density, soil total porosity, microporosity, macroporosity, pH, Ca, Mg, Na, K, Al, P, potential acidity, and the sum of basis were evaluated. Classical and geostatistical statistics were used to determine the ideal number of soil samples. Fuzzy c-means clustering technique was used to separate the areas into homogeneous zones and to allocate the sampling points. The wide method of 20 individual soil samples proved to be inefficient. On the other hand, the use of geostatistics proved to be efficient and is required for each crop field. The c-means clustering was adequate to separate the areas into homogeneous zones and, thus, to assist the sampling point allocation.

Index Terms: Geostatistics. *Mangifera indica* L. Semivariogram. Spatial dependence.

Otimização da amostragem de solo usando análise espacial em área de mangueira irrigada nas condições semiáridas brasileiras

Corresponding author:
katia_a.s@outlook.com

Resumo - A amostragem do solo é um procedimento fundamental na tomada de decisão sobre o manejo do solo; portanto, um plano de amostragem deve representar o mais acuradamente possível o campo de cultivo avaliado. Os objetivos do estudo foram avaliar uma abordagem de amostragem de solo e a alocação de pontos amostrais usando análise espacial e comparar com o método da estatística clássica em pomares de manga irrigada na região do semiárido brasileiro. O experimento foi realizado em três pomares comerciais de manga, no Vale do São Francisco, Brasil. As amostras de solo foram coletadas nas camadas de 0-0,2 e 0,2-0,4 m, seguindo grades regulares, onde o número de amostras variou de 50 a 56. A textura do solo, sua densidade, porosidade total, microporosidade, macroporosidade, bem com seu pH, Ca, Mg, Na, K, Al, P, acidez potencial e a soma de bases foram avaliados. Estatística clássica e geoestatística foram usadas para determinar o número ideal de amostras de solo. A técnica de agrupamento *fuzzy c-means* foi usada para separar as áreas em zonas homogêneas e alocar os pontos de amostragem. O método de coleta de 20 amostras simples de solo mostrou-se ineficiente. Por outro lado, o uso da geoestatística provou ser eficiente e é necessário para cada campo de cultivo. O agrupamento *fuzzy c-means* foi adequado para separar as áreas em zonas homogêneas e, assim, auxiliar na alocação dos pontos amostrais.

Termos para indexação: Geoestatística. *Mangifera indica* L. Semivariograma. Dependência espacial.

[1]Agronomy Engineer, Master in Agronomia, Universidade Federal do Vale do São Francisco, Petrolina city, Pernambuco state, Brazil. E-mail: katia_a.s@outlook.com (ORCID 0000-0001-6618-8753)

[2]Doctor in Agronomy, Teacher, Universidade Federal do Vale do São Francisco, Petrolina city, Pernambuco state, Brazil, marcos.rodrigues@univasf.edu.br (ORCID 0000-0002-6567-1273); augusto.lima@univasf.edu.br (ORCID 0000-0002-8567-9600); italo.cavalcante@univasf.edu.br (ORCID 0000-0003-1610-1546)

[3]Undergraduate student in Agronomy, Universidade Federal do Vale do São Francisco, Petrolina city, Pernambuco state, Brazil, fe_brnd@hotmail.com (ORCID 0000-0001-9464-9334); layannevina@hotmail.com (ORCID 0000-0001-9281-1350)

Introduction

The irrigated land clusters of the São Francisco valley, located in the semi-arid region, have emerged as the most important fruit-growing center in Brazil. In addition, it is considered the largest mango producing region in Brazil, accounting for 85% of the total mangoes exported from the country (CARVALHO et al., 2017). Considering the importance of this region, it is necessary to improve mango cropping techniques to ensure sustainability and reduce production costs. Among the possible techniques, soil sampling may be highlighted, as successful cropping depends on the success of fertility management and irrigation in crop fields (OLIVEIRA et al., 2007).

The condition of the agricultural field should be reproduced as accurately as possible using a minimum number of soil samples to estimate the mean value of the soil properties with some precision (LIMA et al., 2010). In this context, there is a limitation to the use of classical statistics, as suggested in the study by Catani et al. (1954) wherein the collection of 20 individual soil samples is recommended to obtain a composite sample in areas of up to 10 ha that are considered homogeneous. This method has been used indiscriminately in Brazilian agricultural fields regardless of the management system (conventional or no-tillage), crops (annual or perennial), and soil variability. However, the presence of spatial dependence, verified in most agricultural variables such as soil properties (LEMOS FILHO et al.. 2017; RODRIGUES et al., 2012, RODRIGUES et al., 2015), does not satisfy the requirement of randomness of errors, which is one of the premises of classical statistics (WANG et al., 2008). Thus, geostatistics has been shown as a feasible alternative as it considers spatial dependence (RODRIGUES et al., 2013) for the optimization of the sample number.

The semivariogram is the main tool used in geostatistics for the description of spatial correlation (SEIDEL; OLIVEIRA, 2016). From this, the range parameter is obtained, which is defined as the distance to where the sample points have spatial dependence (OLIVER; WEBSTER, 2014). Therefore, distances greater than the range value guarantee the spatial independence of the sampling points, and is useful in determining the sampling density (CARVALHO et al., 2002).

Some studies on annual and semi-perennial crops (LIMA et al., 2010; MONTANARI et al., 2012; OLIVEIRA et al., 2015; SÓRIA et al., 2018; SOUZA et al., 2006) have shown the efficiency of the use of geostatistics relative to classical statistics to determine the ideal number of soil samples. However, these studies do not indicate the location of sample collection within the crop field; they only define the final number of samples. Therefore, the use of management zone delineation techniques, such as the fuzzy c-means clustering, allows the division of crop fields into smaller and more homogeneous areas (RODRIGUES JR. et al., 2011; TAGARAKIS et al., 2013; VALENTE et al., 2012), which can be useful for assisting the sampling point allocation within the field such that the mean can be even more representative of the crop field and is estimated more accurately.

Additionally, there is no study defining the ideal number of samples in fruit growing areas, especially in the Brazilian semi-arid region. These conditions differ greatly from those found in annual crops, mainly in relation to the management of fertilization, which is carried out by fertigation concentrated under the canopy region of the plant, and commonly makes use of foliar fertilization in complementing the nutritional needs of the plant.

The hypothesis of this study is that the soil sampling strategy using spatial analysis may represent the soil properties of commercial mango areas in the semi-arid region of Brazil with greater accuracy.

The objectives of this study were to suggest a soil sampling approach and soil sampling point allocation using spatial analyses and compare to the classic statistic method in irrigated mango orchards in the Brazilian semi-arid region.

Materials and methods

- Site description

The experiment was carried out during 2017 and 2018, in three commercial mango orchards considered homogeneous by the farmers (the same type of soil, slope, plant age and management), in the São Francisco valley in the Brazilian semiarid region. The cultivation was under irrigation with the cultivar Tommy Atkins in distinct soils, focusing on the texture and the textural gradient of the three fields, the main characteristics of which are described in Table 1. According to Köppen's classification, the local climate is BSh, which is semi-arid, with an annual precipitation of less than 500 mm concentrated only in three to four months of the year and annual averages of temperature varying between 18.7 and 33.6 °C (ALVARES et al., 2013).

Table 1. Location, slope, soil type, soil texture, area size, age and spacing of the crop, irrigation and fertilization of the Barreiro de Santa Fé, Mandacaru and Sempre verde crop areas cultivated with irrigated mango cv. Tommy Atkins in the São Francisco valley region, Brazil

	Fazenda Barreiro de Santa Fé	Fazenda Mandacaru	Fazenda Sempre Verde
Location	Petrolina, Pernambuco (9°23'39.37" S e 40°44'32.91" W)	Petrolina, Pernambuco (9°20'50.58" S e 40°33'04.51" W)	Juazeiro, Bahia (9°14'59.38" S e 40°16'58.40" W)
Slope	Very gentle	Very gentle	Very gentle
Soil type	Oxisol	Ultisol	Ultisol
Sand (g kg^{-1})	767 A; 705 B	869 A; 797 B	758 A; 470 B
Clay (g kg^{-1})	160 A; 210 B	101 A; 165 B	95 A; 362 B
Size	9 ha (797 x 114 m)	4.5 ha (353 x 125 m)	9.5 ha (411 x 232 m)
Crop age	25 years old	26 years old	25 years old
Crop spacing	8 x 5 m	8 x 5 m	7 x 6 m
Irrigation	sprinkler (two sprinkles /plant)	sprinkler (two sprinkles /plant)	sprinkler (two sprinkles /plant)
Fertilization	fertigation for N, P and K and manual application for Ca and Mg	fertigation for N, P and K and manual application for Ca and Mg	fertigation for N, P and K and manual application for Ca and Mg

A: A horizont; B: B horizont.

- Data collection and soil analyses

Soil samples were collected under the canopy region after harvest (before application of any product), at depths of 0-0.2 and 0.2-0.4 m, following regular grids containing 56 georeferenced points (56 × 30 m) in the Barreiro de Santa Fé area (Figure 1a), 50 points (32 × 25 m) in the Mandacaru area (Figure 1b), and 53 points (42 × 35 m) in the Sempre Verde area (Figure 1c). The number of samples was defined according to two criteria: 1) obtaining at least 30 pairs of points for calculation of the semivariance in the first lag (ARÉTOUYAP et al., 2016) following Yamamoto and Landim's (2013) recommendation of at least 30 to 40 points, and 2) the geometric shape of each area.
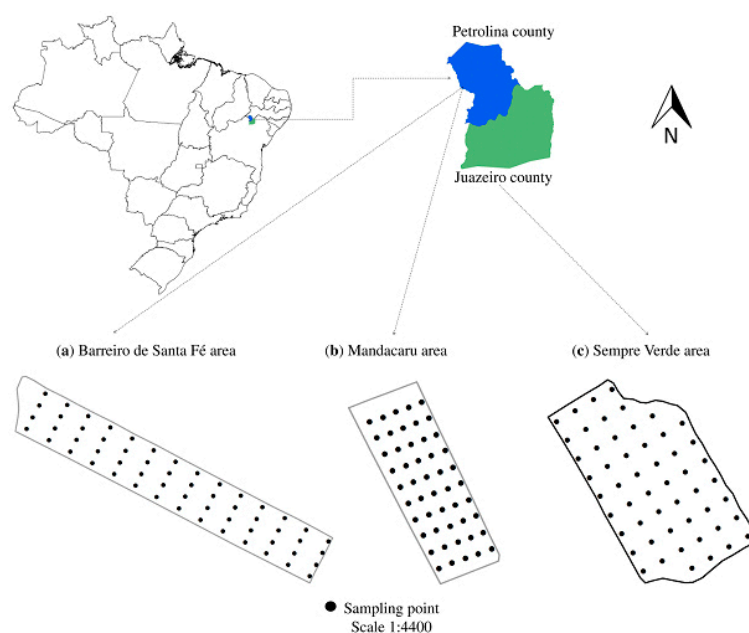

Figure 1. Sampling design in the mango fields located in the São Francisco valley region, semi-arid, Brazil. **a** 56 georreferenced points in the Barreiro de Santa Fé area, **b** 50 points in the Mandacaru area, and **c** 53 points in the

Sempre Verde area

The disturbed soil samples were obtained using a Dutch auger. Each soil sample was analyzed for particle size (pipette method), pH (1:2 soil/water mixture), Ca (KCl 1.0 mol L^{-1} solution extract), Mg (KCl 1.0 mol L^{-1} solution extract), Na (Mehlich-1 extract), K (Mehlich-1 extract), Al (KCl 1.0 mol L^{-1} solution extract), P (Mehlich-1 extract), and H + Al (calcium acetate and alkalimetric titration of the extract) content according to Teixeira et al. (2017a). From the analytical determinations, the sum of bases was calculated.

The undisturbed samples were collected using a double-cylinder, hammer-driven core sampler to determine soil bulk density (BD) and total porosity (TP), according to Teixeira et al. (2017b). Macroporosity (Ma) and microporosity (Mi) were estimated by a mathematical model proposed by Stolf et al. (2011) using data from BD and sand content.

- Descriptive statistics and obtaining the number of samples by the classical method

Descriptive analyses of the data (mean, minimum, and maximum values, and coefficient of variation) was performed. Data normality was verified by the Shapiro-Wilk's test at 5% probability using R statistic software (version 3.2.2) (R Core Team, 2015). The coefficient of variation (CV) was classified according to Pimentel-Gomes and Garcia (2002) who define CV ≤ 10% as low, 10% < CV ≤ 20% as medium, 20% < CV ≤ 30% as high, and CV > 30% as very high variability.

To calculate the ideal number of individual soil samples that is representative of the area, disregarding the spatial dependence of the samples, the Cline method (1944) obtained from Equation 1 was used, which was the same procedure adopted in the study that defined the conventional 20 individual samples recommendation

$$n = \left(\frac{t_\alpha x CV}{d} \right)^2 \qquad \text{Eq. 1}$$

(CATANI et al., 1954),

where n is the minimum number of samples, tα is the Student's *t*-test value α = 0.95, CV is the coefficient of variation, and d is the percentage variation around the mean (5%).

The soil properties were divided into three groups of variables, where each group was created based on two criteria: similarity of CV value and number of samples based on the classical approach. The chemistry group comprised the chemical properties of the soil, the physics group included the physical properties that are easily influenced by management (BD, TP, Ma, and Mi), and the texture group comprised the soil texture fractions (sand and clay content) that are considered stable and more related to soil formation than to anthropogenic practices.

- Geostatistical analyses in obtaining the number of samples

To calculate the ideal number of individual soil samples representative of the crop field, considering the spatial distribution of the samples, geostatistics was used. Semivariogram models were utilized to estimate the spatial dependence between the samples and to identify whether the variations were systematic or random. The semivariogram is a basic tool that can quantitatively represent the variation in a regionalized phenomenon in space and was estimated by Equation

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[Z(x_i) - Z(x_i + h) \right]^2 \qquad \text{Eq. 2}$$

2 (OLIVER; WEBSTER, 2014),

where N(h) is the number of experimental pairs of data separated by a distance h, $Z(x_i)$ is the value determined at each sampled point, and $Z(x_i + h)$ is the value measured at a point plus a distance h.

For variables that showed a trend, that is, there was no stabilization of the sill values and thus, the intrinsic hypothesis was not satisfied, the trend surface method proposed by Vieira et al. (2010) was performed. This method involves fitting a trend surface by the least squares method and subtracting it from the original data generating a new variable called residuals. The semivariogram is then fitted using the residuals.

After the preparation of the experimental semivariogram, the fitting of the models that represent the spatial behavior of the variable under study was performed. The three most commonly used models for each semivariogram are: spherical, exponential, and Gaussian models, as tested by Zůvala et al. (2016). Details of the mathematical expressions of the models can be found in Oliver and Webster (2014).

After obtaining the parameters of the experimental semivariograms for each variable, the scaled semivariograms were performed by dividing the semivariance by the statistical variance, with the objective of reducing them to the same scale, to make the comparison between results of different variables easier (CASTIONI et al., 2019) and allowing the determination of an ideal number of samples that satisfy all the studied variables simultaneously. Three scaled semivariograms were prepared for each area, related to the three groups of variables: chemistry, physics, and soil texture in the two evaluated layers.

The number of individual samples, considering the spatial dependence, was determined using the range value obtained from the scaled semivariogram. Although distance values greater than the range value already showed spatial independence, and consequently, randomness of the samples (TREVISAN et al., 2017;

YAMAMOTO; LANDIM, 2013), the collection of samples was simulated at a distance equivalent to one and a half times the range value as a precaution. In addition, the shape of the crop field should take into account to determine the number of samples in each studied field, thus, the following equation was used to

$$\frac{X}{1,5 \times A} \times \frac{Y}{1,5 \times A} \qquad \text{Eq. 3}$$

calculate the number of samples (Equation 3),

where X is the vertical distance from the field; A is the value of range; and Y is the horizontal distance of the area, all variables in meters. When the first and / or second part of the equation obtained a decimal number, the value was rounded to the next integer.

- Obtaining homogeneous zones and allocating samples

After the preparation of the experimental semivariograms and fitting of the theoretical models, the data that presented spatial dependence were interpolated using ordinary kriging. This technique is a generic term for a range of least squares methods applied to obtain the best linear unbiased predictions in terms of the minimum variance. It requires only knowledge of the variogram function and data for its implementation (OLIVER; WEBSTER, 2014).

Aiming for the delineation of areas into homogeneous zones for each of the groups of variables, the fuzzy c-means clustering technique (FRIDGEN et al., 2004) was performed with the interpolation maps using Management Zone Analyst software (MZA, version 1.0.1). As the soil collection was carried out in one location for both layers, each group included the variables of the two layers. The configurations in the MZA were: fuzziness exponent = 1.3, maximum number of iterations = 300, convergence criterion = 0.0001, minimum number of zones = 2, and maximum number of zones = 6. The determination of the ideal number of zones was performed using two indexes proposed by Odeh et al. (1992): NCE (Normalized Classification Entropy) and FPI (Fuzziness Performance Index), in which the lowest number for both indexes was selected.

To confirm that the number of zones suggested by the indexes was adequate, the Student's t-test ($p < 0.05$) was used to compare the mean values of the variables between the zones. It was considered adequate when at least one of the studied variables showed means considered statistically different between the indicated zones.

For the zone maps, the number of samples per zone was defined, which was one sample per zone when the number of zones was equal to the number of samples recommended based on the scaled semivariogram range value or established according to the percentage of the area that each zone represented when the recommended number of samples was higher than the number of zones.

Results and Discussion

The descriptive statistics of the chemical and physical properties of soil in the layers at depths of 0-0.2 and 0.2-0.4 m in the Barreiro de Santa Fé, Mandacaru, and Sempre Verde areas are described in tables 2 and 3, respectively. The highest variability in the soil properties common to all three areas in both layers (Tables 2 and 3), based on the CV was observed for K, P, Ma, and clay content. As for the chemical properties, other studies also observed very high variability for K and P (LIMA et al., 2010; OLIVEIRA et al., 2015; RODRIGUES et al., 2012). These results can be justified, according to Rodrigues et al. (2012), as caused by the continuous applications of fertilizers that increase the variability in the crop field.

Additionally, in the semi-arid region, this variability may be increased even more, owing to the use of micro-sprinkler irrigation systems, where the soil is irrigated daily and most of the fertilizers are applied by fertigation. Moreover, the soils from the semi-arid region generally have a sandy texture, and its adsorption sites (colloids) are easily saturated, which favors the increase in the mobility and leaching of K, as it is a monovalent cation (SANTOS et al., 2015). Silva et al. (2014) confirmed this hypothesis in a study of soil fertility changes in mango orchards in the São Francisco valley region, demonstrating that the movement of the nutrients is favored by daily irrigation practices and the sandy soil texture, similar to observations in the three areas of the present study, contributing to its high variability.

As for the physical attributes, the Ma is also influenced greatly by agricultural management, and its values can be increased by the addition of organic matter (REGELINK et al., 2015), and at the same time present reduction due to human trampling in manual harvest, and the cultivation itself (irrigation, fertilization, etc.), by inducing the fractionation of larger aggregates in smaller units, with consequent reduction of macropores (CUNHA et al., 2011), justifying this high variability.

The high CV values for clay content can be explained by the large range in variation in the area considering the large amplitude (minimum and maximum values, Tables 2 and 3), and in addition, consistent with the results obtained by Ceddia et al. (2009), owing to the greater error associated with clay suction during the pipette method. Similar results were obtained by Rodrigues et al. (2015) in an Ultisol under guava cultivation, which also presented high variability for the clay content.

Analyzing the maximum and minimum values (Tables 2 and 3), it was observed that there is a considerably large amplitude for many attributes, as seen by the high variability. This can result in an error of recommendation due to the underestimation or overestimation of the real value as the average is considered the best estimator of soil attributes in the sample unit in classical statistics

Table 2. Descriptive analysis of the soil chemical and physical properties at depth of 0-0.2 m cultivated with irrigated mango in the Barreiro de Santa Fé, Mandacaru and Sempre Verde areas, São Francisco valley region, Brazil

	pH	Na	K	H+Al	Al	Ca	Mg	SB	P	BD	Ma	Mi	TP	Sand	Clay
		---------------- $cmol_c\ dm^{-3}$ --------------------							$mg\ dm^{-3}$	------ $kg\ dm^{-3}$ ------			---- % ----		
Barreiro de Santa Fé															
Mean	7.09	0.20	0.29	0.83	0.10	4.51	1.84	6.77	60.02	1.29	0.27	0.25	42.91	83.24	8.28
Maximum	7.76	0.31	0.44	1.16	0.20	6.25	2.30	9.31	249.14	1.42	0.37	0.29	50.94	89.65	14.50
Minimum	6.49	0.09	0.15	0.66	0.10	3.15	1.47	4.97	10.45	1.12	0.20	0.21	36.23	78.04	3.60
CV (%)	4	32	25	15	35	16	11	14	76	6	15	7	8	4	33
W	0.13^{ns}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.01^{*}	0.00^{*}	0.59^{ns}	0.67^{ns}	0.08^{ns}	0.00^{*}	0.69^{ns}	0.43^{ns}
Mandacaru															
Mean	6.81	0.06	0.43	1.06	0.13	2.38	0.82	3.78	14.81	1.12	0.37	0.20	48.28	90.50	4.04
Maximum	7.74	0.13	0.92	1.82	0.20	3.78	1.43	6.16	68.75	1.29	0.48	0.23	56.46	92.50	6.70
Minimum	5.94	0.04	0.21	0.33	0.10	1.54	0.50	2.65	3.27	0.86	0.29	0.16	40.29	87.78	1.60
CV (%)	7	41	42	37	36	22	27	22	86	11	16	9	10	2	35
W	0.56^{ns}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.59^{ns}	0.00^{*}	0.00^{*}	0.34^{ns}	0.07^{ns}	0.13^{ns}
Sempre Verde															
Mean	6.51	0.26	0.67	1.19	0.05	6.00	2.09	8.91	69.84	1.41	0.18	0.31	40.06	67.80	17.07
Maximum	6.87	0.43	1.21	2.15	0.05	10.22	2.56	11.77	159.75	1.64	0.33	0.34	56.65	77.81	24.10
Minimum	6.20	0.13	0.38	0.50	0.05	3.00	1.50	5.72	22.22	1.05	0.09	0.27	29.86	57.63	9.40
CV (%)	3	28	31	44	0	31	14	20	46	10	35	7	17	7	21
W	0.94^{ns}	0.00^{*}	0.00^{*}	0.01^{*}	0.00^{*}	0.45^{ns}	0.30^{ns}	0.99^{ns}	0.00^{*}	0.01^{*}	0.01^{*}	0.00^{*}	0.03^{*}	0.00^{*}	0.65^{ns}

H+Al: potential acidity; Al: exchangeable acidity; SB: sum of basis; BD: bulk density; Ma: macroporosity; Mi: microporosity; TP: total porosity; CV: coefficient of variation; W: Shapiro-Wilk normality test; ns: not significant at 5% of probability, indicating data normality.

Table 3. Descriptive analysis of the soil chemical and physical properties at depth of 0.2-0.4 m cultivated with irrigated mango in the Barreiro, Mandacaru and Sempre Verde areas, São Francisco valley region, Brazil

	pH	Na	K	H+Al	Al	Ca	Mg	SB	P	BD	Ma	Mi	TP	Sand	Clay
		-------------------- $cmol_c\ dm^{-3}$ -----------------							$mg\ dm^{-3}$	----- $kg\ dm^{-3}$ -----			--------- % ---------		
Barreiro de Santa Fé															
Mean	6.79	0.26	0.53	1.31	0.09	3.46	1.56	5.79	49.33	1.54	0.15	0.29	35.70	80.21	13.13
Maximum	7.37	0.30	0.92	1.65	0.10	4.63	2.10	7.74	224.73	1.72	0.30	0.31	50.52	88.36	20.50
Minimum	6.35	0.22	0.26	1.16	0.05	2.36	1.08	4.41	11.59	1.29	0.07	0.23	27.10	74.44	7.40
CV (%)	4	14	31	16	25	17	13	12	96	7	37	7	17	5	25
W	0.47^{ns}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.22^{ns}	0.14^{ns}
Mandacaru															
Mean	6.76	0.05	0.33	0.72	0.19	1.19	0.65	2.33	37.55	1.45	5.23	0.21	37.55	90.27	4.43
Maximum	8.00	0.09	0.62	1.32	0.3	1.93	0.87	3.71	49.09	1.59	0.29	0.27	49.10	93.17	6.70
Minimum	5.87	0.04	0.13	0.33	0.05	0.60	0.42	1.6	26.86	1.28	0.12	0.22	26.86	86.07	1.70
CV (%)	8	35	38	42	38	27	19	23	14	5	17	5	14	2	29
W	0.38^{ns}	0.00^{*}	0.00^{*}	0.02^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.00^{*}	0.41^{ns}	0.26^{ns}	0.00^{*}	0.00^{*}	0.00^{*}	0.76^{ns}
Sempre Verde															
Mean	6.44	0.29	0.72	1.1	0.06	5.98	1.89	8.82	38.59	1.61	0.08	0.34	31.78	63.94	22.64
Maximum	6.74	0.43	1.56	1.82	0.10	10.61	2.73	13.47	91.98	1.74	0.20	0.38	44.55	71.98	32.70
Minimum	6.05	0.13	0.26	0.50	0.05	3.00	0.73	5.01	2.71	1.36	0.007	0.30	22.32	54.18	14.10
CV (%)	3	27	45	36	35	33	27	24	62	6	59	6	16	9	25
W	0.03^{*}	0.12^{ns}	0.00^{*}	0.04^{*}	0.00^{*}	0.43^{ns}	0.13^{ns}	0.79^{ns}	0.00^{*}	0.00^{ns}	0.00^{*}	0.03^{*}	012^{ns}	0.40^{ns}	0.37^{ns}

H+Al: potential acidity; Al: exchangeable acidity; SB: sum of basis; BD: bulk density; Ma: macroporosity; Mi: microporosity; TP: total porosity; CV: coefficient of variation; W: Shapiro-Wilk normality test; ns: not significant at 5% of probability, indicating data normality

(FERRAZ et al., 2012).

The minimum number of individual samples for estimating soil properties based on classical statistics using the Cline method (1944) with a 5% variation around representativeness of all samples.

the mean for the three studied areas can be found in Table 4. For the number of samples to be representative, it is necessary to choose the variable that needs the greatest number of samples, thus guaranteeing the

Table 4 The minimum number of single samples to estimate the soil properties based on the Cline's method (1944) with 5% variation around the mean at the layers of 0-0.2 m and 0.2-0.4 m depths cultivated with irrigated mango in the Barreiro de Santa Fé, Mandacaru and Sempre Verde areas, São Francisco valley region, Brazil

Crop field	pH	Na	K	H+Al	Al	Ca	Mg	SB	P	BD	Ma	Mi	TP	Sand	Clay
		---------------- $cmol_c$ dm^{-3} ---------------							mg dm^{-3}	-- kg dm^{-3} --			------- % -------		
						0-0.2 m									
Barreiro de Santa Fé	4	167	104	35	192	42	20	31	933	6	39	8	12	3	172
Mandacaru	9	271	287	216	204	77	121	79	1620	21	42	14	15	1	191
Sempre Verde	2	128	152	319	1	154	33	64	340	18	197	7	45	9	74
						0.2-0.4 m									
Barreiro de Santa Fé	3	34	154	43	97	48	27	24	1494	8	215	8	47	4	102
Mandacaru	11	192	227	277	226	113	57	82	1190	4	48	4	33	1	135
Sempre Verde	2	121	334	213	199	173	118	93	627	7	563	6	44	12	98

H+Al: potential acidity; Al: exchangeable acidity; SB: sum of basis; BD: bulk density; Ma: macroporosity; Mi: microporosity; TP: total porosity

It can be observed that the attributes which required a greater number of individual samples by classical statistics (Table 4) were the same ones that presented greater variability (Tables 2 and 3), which is expected from Cline's equation (1944), which accounts for the CV value. Thus, in the three studied areas, P was the determinant soil property to represent the chemical group, Ma was the determinant for the physical group, and clay was the determinant for the texture group, as they showed greater variability. In addition, it was verified that all groups in both layers and in all studied areas would require a greater number of individual samples than the current recommendation of 20 individual samples to obtain a composite sample.

The experiment carried out by Catani et al. (1954), where the current recommendation (20 individual samples) was determined, occurred in two areas under Oxisols, one cultivated with intercropped coffee and maize, and the other with a maize monoculture, both areas of up to 6 ha. Those conditions differ greatly from the conditions of the present study in terms of soil texture (mostly sandy), area (up to 10 ha), crop type (fruit crop), and soil management (daily irrigations and fertilization concentrated in the region under the plant canopy).

Thus, studying each crop field to determine the ideal sampling density is necessary as different natural characteristics and management are encountered. Schlindwein and Anghinoni (2000), estimating the horizontal variability of soil fertility and number of soil samples for no-tillage systems demonstrated that the current recommendation is inadequate to represent soil fertility in this system, and useful only for the conventional

tillage systems.

It was also observed that lower the clay content (Tables 2 and 3) between the three areas (Mandacaru < Barreiro de Santa Fé < Sempre Verde), greater was the variability of soil properties in the chemical and soil texture groups and, consequently, higher the number of samples in each group. This is due to the lower buffering capacity of cations and anions in low clay content soils, indicating the lower resistance of the soil to changes in the concentration of a given nutrient in solution, contributing to higher leaching and subsequently increased variability (MALUF et al., 2015). However, the physical group showed the opposite behavior; higher the clay content (Sempre Verde > Barreiro de Santa Fé > Mandacaru), greater was the Ma variability, which is the soil property that determined the number of samples for this group.

The CV values indicate the degree of variability in the soil properties, with the lowest values showing lower heterogeneity in the studied area (WANG et al., 2008). Although the statistical measure CV allows the comparison of the variability between samples of variables that present different units, it does not allow the analysis of the spatial pattern of the soil properties (OLIVEIRA et al., 2015), making the use of geostatistics necessary.

In addition, it can be shown that using CV as a parameter to define the number of soil samples results in an impractical number of samples. Some authors have observed that soil sampling may have its efficiency increased with the addition of the spatial component (CASTIONI et al., 2019; LIMA et al., 2010; SOUZA et al., 2006), which can be determined using the range value of the semivariogram. Smaller range values indicate greater variability of the data, and consequently, a greater number of samples are required to represent the area (OLIVEIRA et al., 2015).

It was observed that most of the soil properties showed spatial dependence. Only Na and TP in the 0-0.2 m layer and Na, K, and Mg in the 0.2-0.4 m layer in the Barreiro de Santa Fé area, and Na, BD, clay, and Ma in the 0.2-0.4 m layer in the Mandacaru area showed a pure nugget effect, that is, they did not present spatial dependence. The pure nugget effect may have occurred owing to the high spatial variability of these soil properties and use of insufficient distance in the sample grid to model their spatial dependence. The pH in both layers and Mi in the 0-0.2 m layer in the Barreiro de Santa Fé area, and Mg in the 0.2-0.4 m layer in the Sempre Verde area showed a trend in the data, thus, the semivariogram was fitted to the data residues for these variables.

The calculation of the number of individual samples for each variable as a function of the range of the experimental semivariogram allows verification the reduction in the number of samples when geostatistics was used compared to that of classical statistics, but it does not allow the definition of a single ideal number of representative individual samples for all the variables. This is because the range is inversely proportional to the number of samples and to guarantee the spatial independence of the samples should be collect in distance greater than the range value.

Therefore, if sample numbers were obtained for each variable, theoretically the best choice would be the one with the largest number of samples to be representative of the others. However, this variable would have a lower range value, requiring a smaller sampling distance, which would not guarantee the spatial independence of the variables that showed the need for a smaller number of samples (greater range).

On the other hand, the determination of the sample density as a function of the largest range would guarantee the randomness of all variables, however the minimum representative individual number would be not reached for some variable, since the larger the range, the smaller the number of individual samples. Therefore, the variable with shortest range, which would require a larger number of individual samples, would not be satisfactorily represented.

These results are consistent with the experiments conducted by Souza et al. (2006) who studied a very clayey Oxisol under sugar cane in the São Paulo state and Lima et al. (2010) who evaluated a clayey Ultisol under a native forest (Atlantic Forest) in Espírito Santo state, where the use of the semivariogram range reduced the number of individual samples to be collected in relation to classical statistics for most of the soil chemical properties; but, did not allow the determination of a representative number of individual samples for all variables simultaneously.

In addition, Souza et al. (2006) and Lima et al. (2010) observed that if the classic recommendation of 20 individual samples was used, the minimum number

obtained in their studies would not be met for some plant nutrients. It is also possible to verify that the number of individual samples for each variable differed between the two studies, reinforcing the need for the application of this technique to each area owing to the different conditions. Similar statement was made by Yu et al. (2011) who observed that the sample density may varies according to the type of soil, and patterns of land use, thus it is not appropriate to extrapolate results from smaller areas to larger areas, due to the different complexities of the involved factors.

In order to solve the problem of recommending an ideal individual sample number to represent each studied soil property individually, scaled semivariograms were performed (Figures 2, 3, and 4), which allow the fitting of a single model for all the variables of each group simultaneously, and therefore, a single range and a representative number of individual samples for all attributes (CASTIONI et al., 2019; OLIVEIRA et al., 2015; SÓRIA et al., 2018).

Considering that the distance between the individual samples collected should be greater than the range so that the randomness of the samples can be guaranteed (OLIVER; WEBSTER, 2014); a distance of one and a half times the value of the range of the scaled semivariogram was adopted as a precaution. Therefore, in the Barreiro de Santa Fé area five individual samples spaced at 180 and 172.5 m for the chemical (Figure 2a) and physical (Figure 2b) groups, respectively, and three individual samples spaced at 300 m for the texture group (Figure 2c) should be collected.

In the Mandacaru area, based on the range values, eight individual samples spaced at 105 m are required to estimate the mean value of the chemical group (Figure
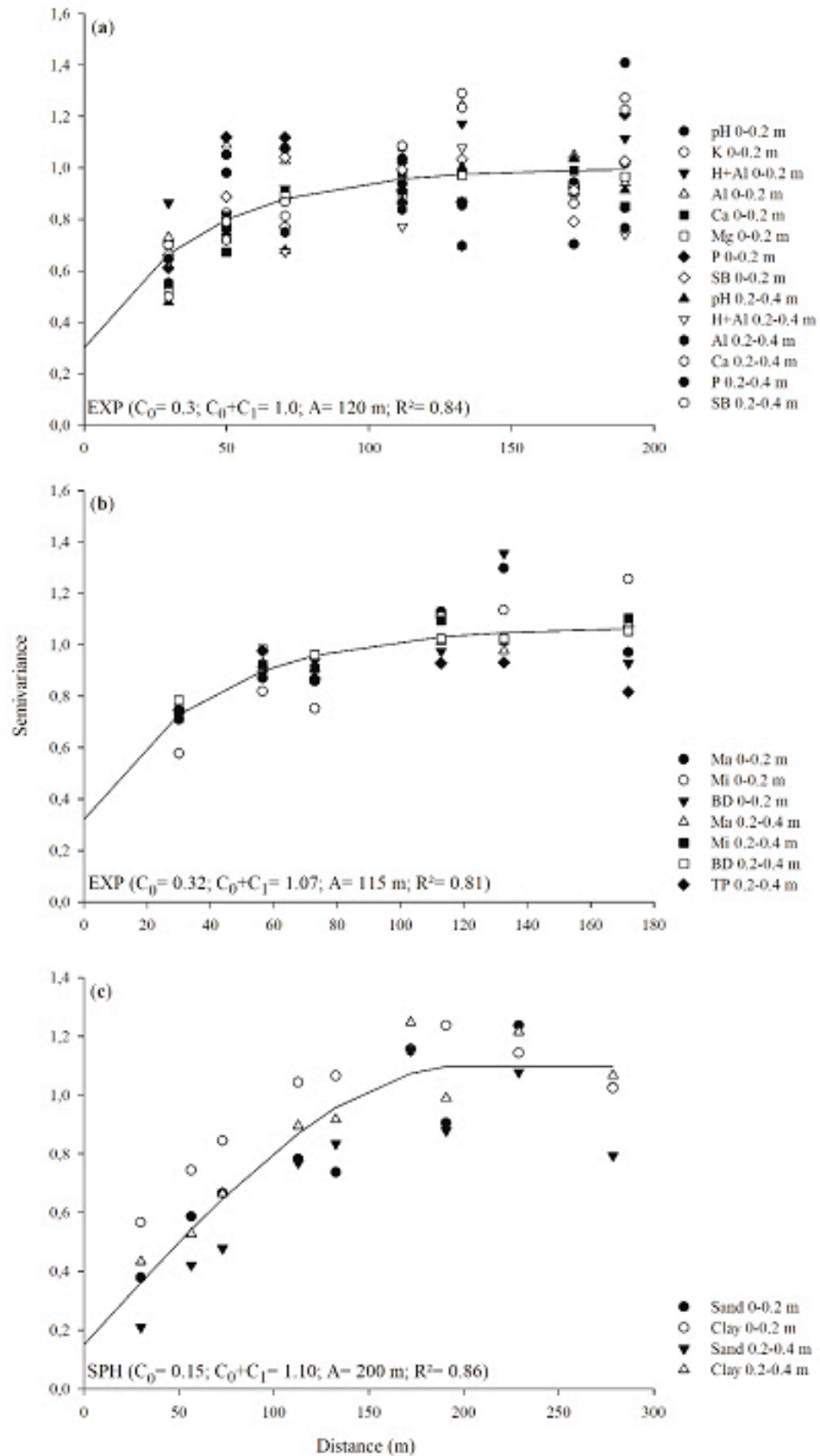
group **(c)** cultivated with irrigated mango, Mandacaru field, São Francisco valley region, Brazil [model (nugget effect, sill, range, the coefficient of determination)]. GAU: Gaussian; EXP: exponential
Figure 2. Scaled semivariograms fitted to the data of soil chemical group **(a)**, soil physical group **(b)** and soil texture
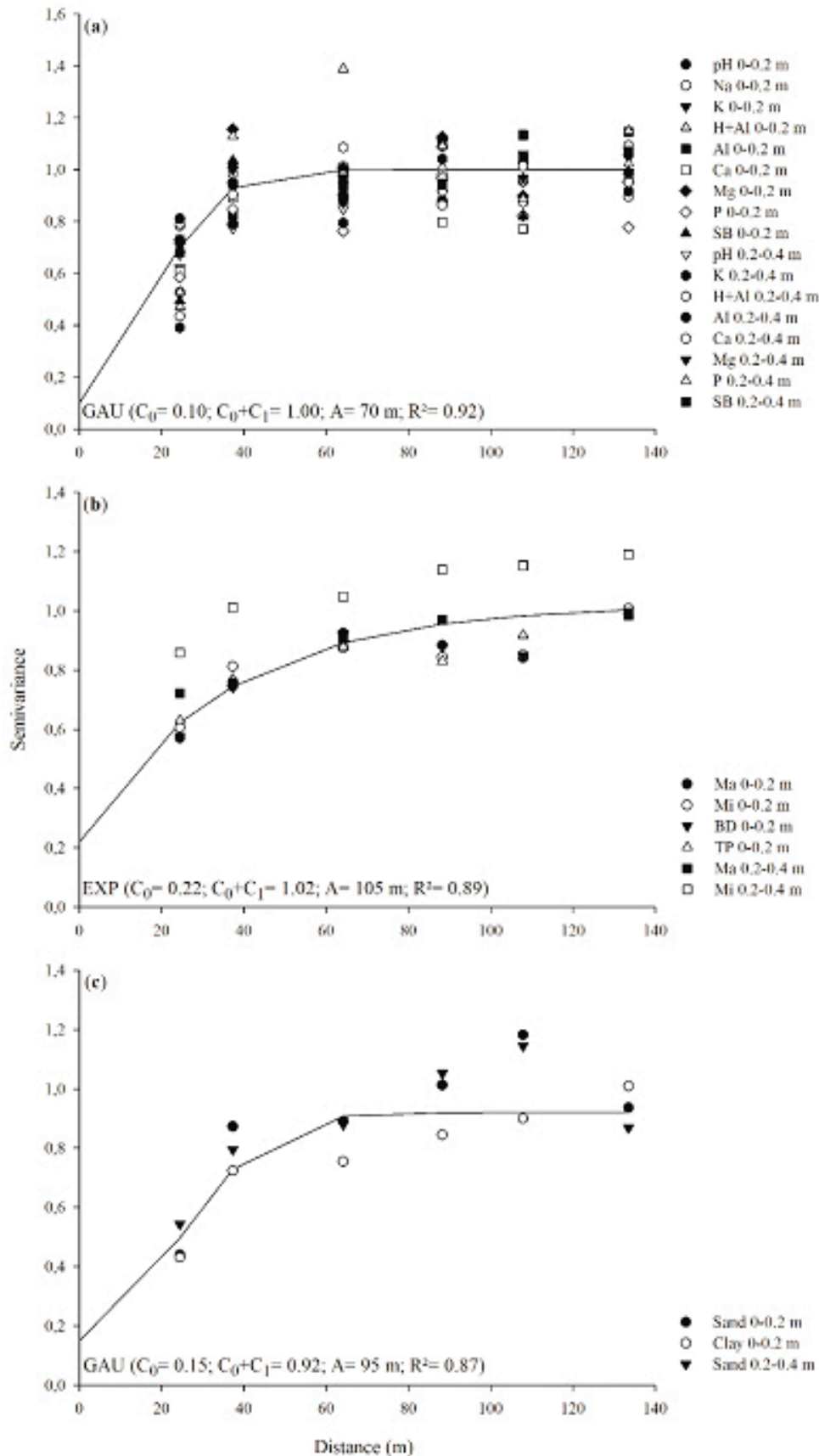
3a), and three individual samples to represent both the physical (Figure 3b) and the texture (Figure 3c) groups, spaced at 157.5 and 142.5 m, respectively.

Figure 3. Scaled semivariograms fitted to the data of soil chemical group **(a)**, soil physical group **(b)** and soil texture

group **(c)** cultivated with irrigated mango, Barreiro de Santa Fé field, São Francisco valley region, Brazil [model (nugget effect, sill, range, the coefficient of determination)]. SPH: spherical; EXP: exponential

Considering the range values of the Sempre Verde area, six individual samples are required to estimate both
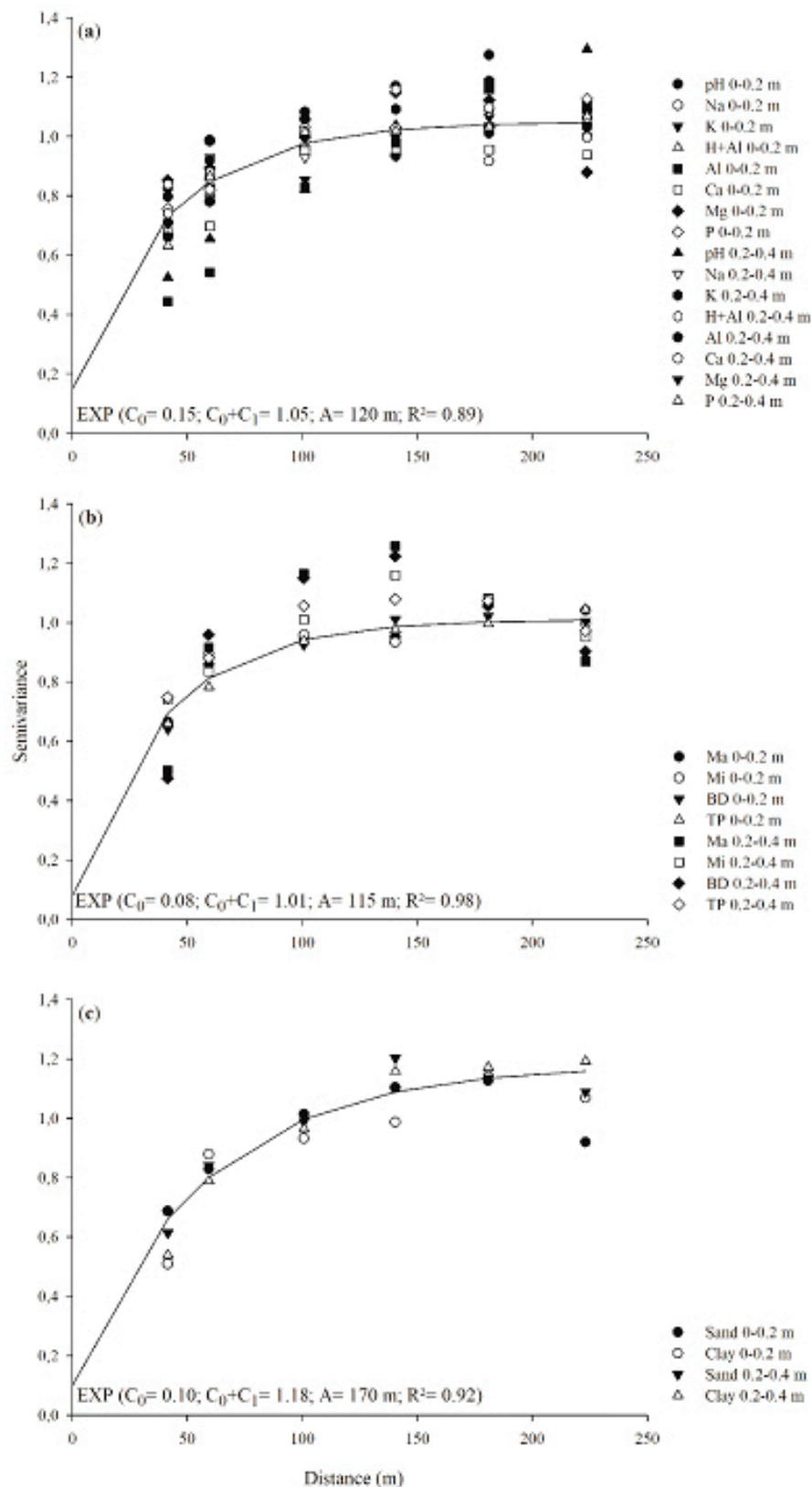
the mean values of the chemical (Figure 4a) and the physical (Figure 4b) groups spaced at 180 and 172.5 m, respectively, and two individual samples for the texture group (Figure 4c), spaced at 255 m.

Figure 4. Scaled semivariograms fitted to the data of soil chemical group **(a)**, soil physical group **(b)** and soil texture

group **(c)** cultivated with irrigated mango, Sempre Verde field, São Francisco valley region, Brazil [model (nugget effect, sill, range, the coefficient of determination)]. EXP: exponential

It can be observed that for all variables groups and studied areas the number of individual samples obtained from the scaled semivariogram range values was much lower than those obtained by classical statistics determined by the Cline (1944) method. Additionally, it was shown that the number of individual samples was lower also in relation to the current recommendation, possibly owing to the addition of the spatial component, reflecting the increase in the spatial dependence range and, consequently, in the reduction of the number of samples.

Similar results were obtained by Oliveira et al. (2015) in an Ultisol soil with a anthropic horizon under natural vegetation (Amazon forest) and pasture in the Amazonas state and Sória et al. (2018) studying an area with Oxisol and Ultisol under sugar cane in São Paulo state, with both studies demonstrating the efficiency of geostatistics in obtaining the sampling density through the range of the scaled semivariogram and change in the sampling intensity as the soil and cultivation conditions changed.
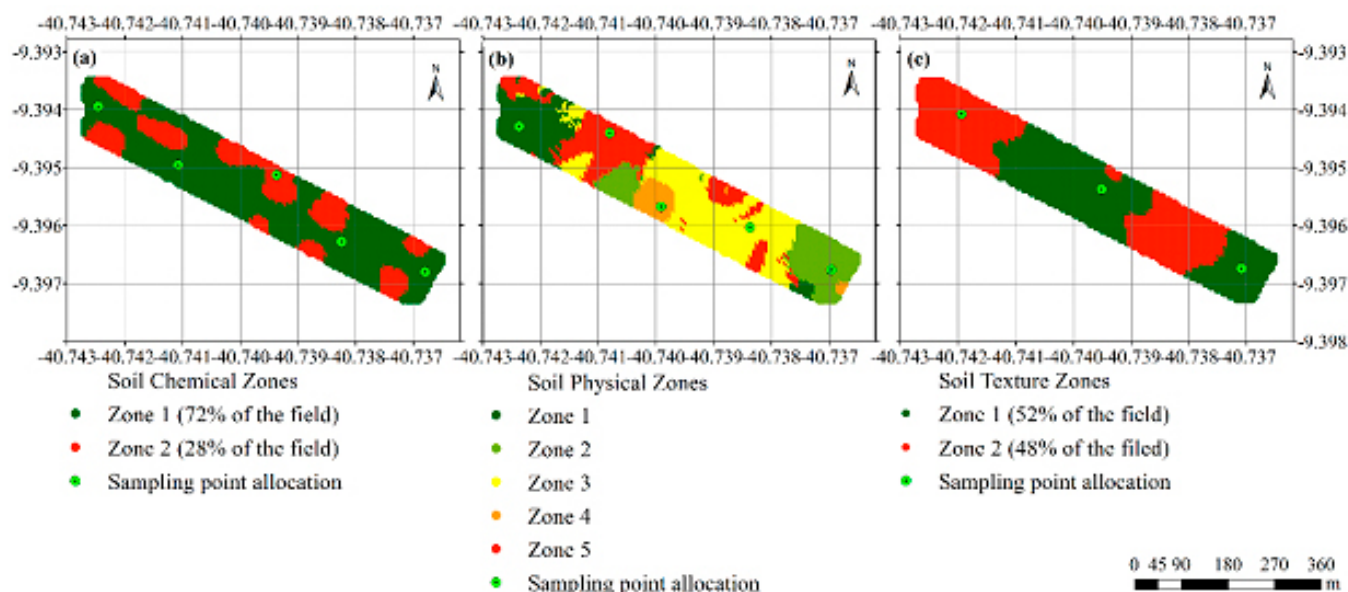
The experiments of Oliveira et al. (2015) and Sória et al. (2018) are consistent with the results of the present study and emphasize the need to specifically study each area to determine sampling density. Since, even in similar conditions of climate, soil type and crop (Table 1), a lot of others variables can be related to soil spatial variability

(fertilization management, fertilizer sources, irrigation management, etc.) (TESFAHUNEGN et al., 2011) and different soil sampling density could be found, which makes each crop field unique.

Despite the recommendation of the distances at which the individual samples should be collected, random collection in the area could result in errors in estimation of the mean, as there are several ways to allocate the samples in a field following the determined distances. To solve this problem, it was necessary to separate the area into zones to minimize the error and, consequently, increase the accuracy (Figures 5, 6, and 7).

In addition, the separation of the area into zones showed that areas initially identified as homogeneous by the farmers without considering spatial analysis, were considered heterogeneous. The significant statistical difference in the average values between the zones proved this heterogeneity of the area (BAZZI et al., 2013; MALLARINO; WITTRY, 2004), justifying the need to collect individual samples in each zone so that the errors in the estimation of the mean are minimized, as already mentioned, even if it is not possible to apply site-specific management in each zone.
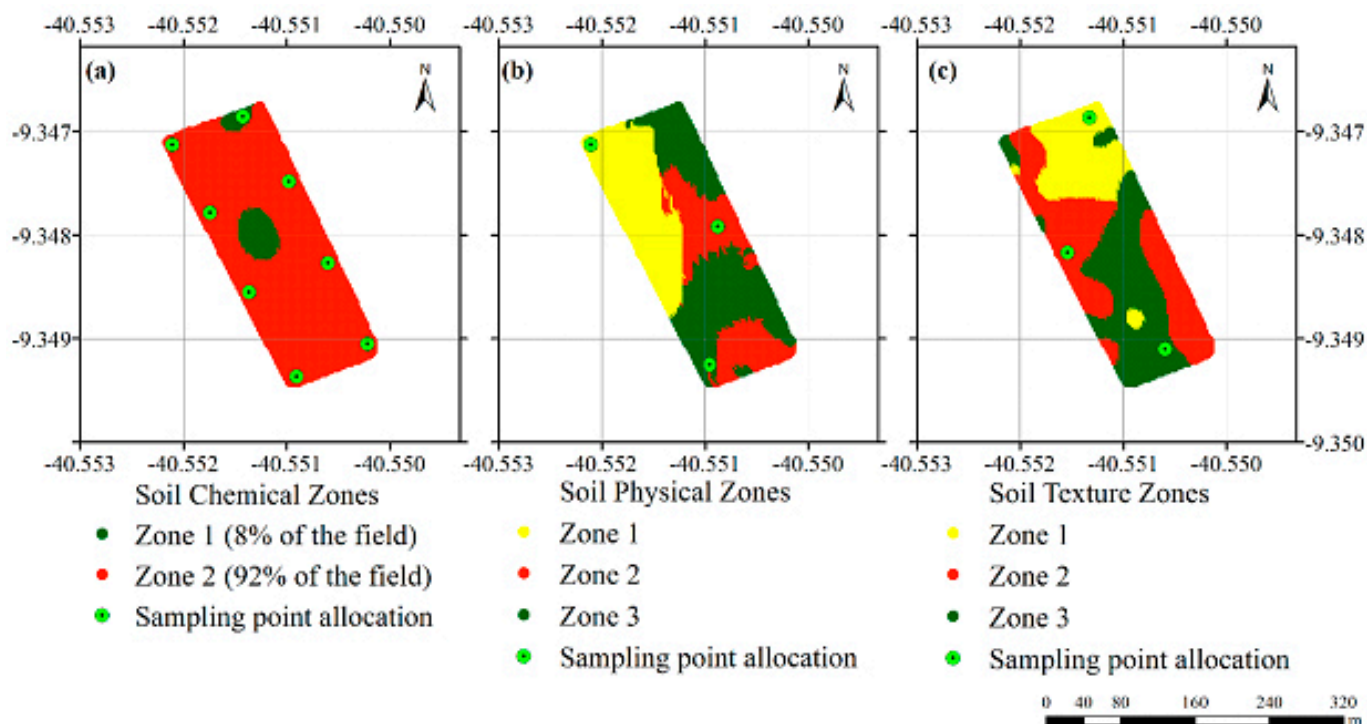
The ideal number of zones for the Barreiro de Santa Fé area was two, five, and two for the chemical (Figure 5a), physical (Figure 5b), and texture (Figure 5c) groups, respectively, considering the lowest NCE and FPI. As the number of individual samples required was five for the chemical group and three for the texture group, the number of individual samples for each zone took into account the percentage of the area of the zones. Therefore, four and two individual samples from zone 1 and one individual sample each from zone 2 for chemistry and

Soil Chemical Zones
- Zone 1 (72% of the field)
- Zone 2 (28% of the field)
- Sampling point allocation

Soil Physical Zones
- Zone 1
- Zone 2
- Zone 3
- Zone 4
- Zone 5
- Sampling point allocation

Soil Texture Zones
- Zone 1 (52% of the field)
- Zone 2 (48% of the filed)
- Sampling point allocation

0 45 90　180　270　360
m

texture, respectively, should be collected. For the physics group, the number of zones coincided with the number of individual samples required; so, the recommendation is one individual sample per zone.
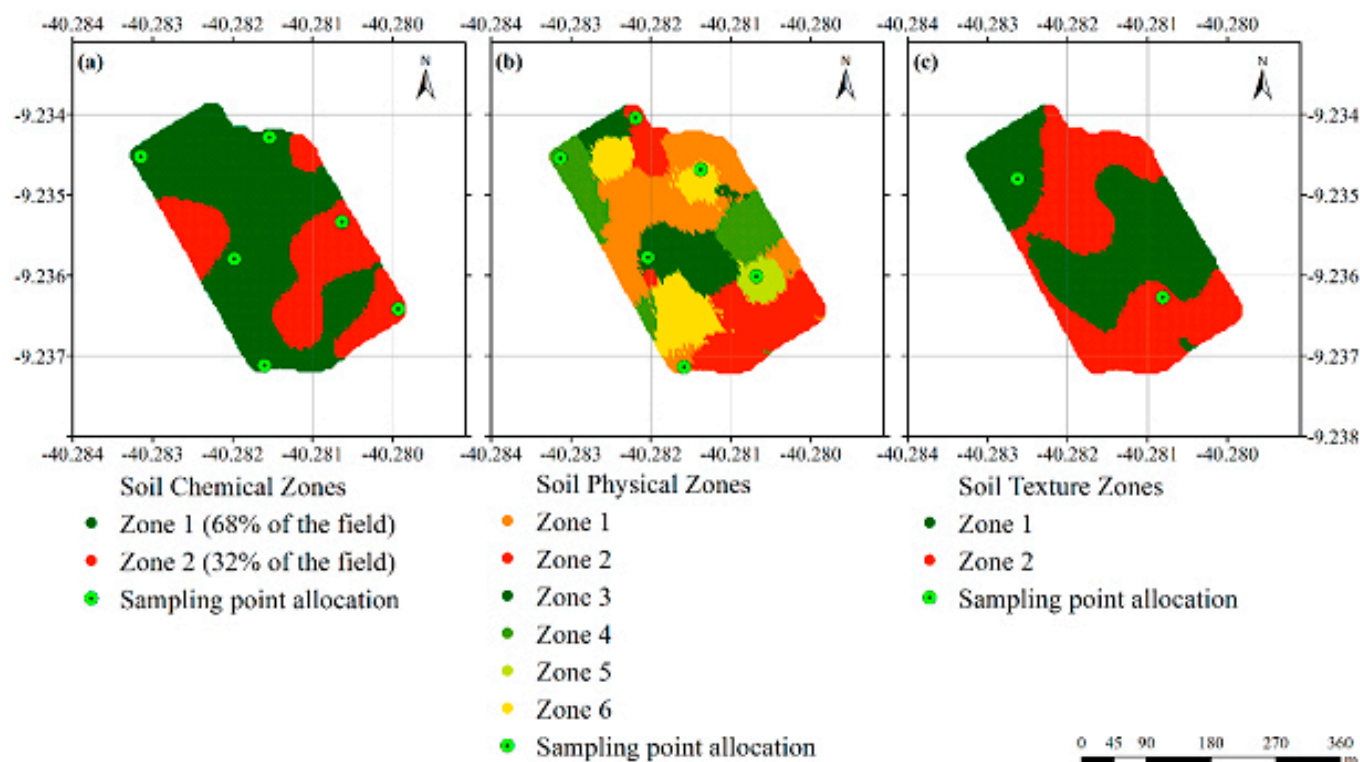
Figure 5. Maps of the zones to allocate the single samples for the soil properties chemical **(a)**, physical **(b)** and texture **(c)** group cultivated with irrigated mango, Barreiro de Santa Fé field, São Francisco valley region, Brazil

The ideal number of zones for the Mandacaru area was two, three, and three for the chemical (Figure 6a), physical (Figure 6b), and texture (Figure 6c) groups, respectively. As the number of individual samples required was eight for chemistry, the number of individual samples from each zone took into account the percentage of the area of the zones.

Figure 6. Maps of the zones to allocate the single samples for the soil properties chemical (**a**), physical (**b**) and texture (**c**) group cultivated with irrigated mango, Mandacaru field, São Francisco valley region, Brazil

Therefore, one individual sample should be collected in zone 1 and seven individual samples in zone 2. For the physical group, the number of zones coincided with the number of individual samples needed, so a single sample should be collected per zone. The ideal number of zones for this area considered the lowest NCE and FPI index, except for the texture group.

The appropriate number of individual samples for texture was three samples and the ideal number of zones was four, making the collection impossible. Therefore, it was decided to consider the number of zones as three, and that collection should be done by taking one individual sample per zone. Increasing the number of individual samples to four would not meet the condition of independence of the samples, because the distance between them would be smaller than the range.

The ideal number of zones for the Sempre Verde area was two, six, and two for the chemical (Figure 7a), physical (Figure 7b), and texture (Figure 7c) groups, respectively. As the number of individual samples required was six for chemistry, the number of individual samples from each zone took into account the percentage of the area of the zones. Therefore, four individual samples should be collected in zone 1 and two individual samples

Figure 7. Maps of the zones to allocate the single samples for the soil properties chemical **(a)**, physical **(b)** and texture **(c)** group cultivated with irrigated mango, Sempre Verde field, São Francisco valley region, Brazil

in zone 2. For the physics group, the number of zones coincided with the number of individual samples needed, so a single sample should be collected per zone.

The adequate number of individual samples for the texture group was two individual samples and the ideal number of zones was four, making it impossible to collect. Therefore, it was decided to consider the number of zones as two, and that collection should be done by taking one individual sample per zone. As in the Mandacaru area, increasing the number of samples to four would not meet the independence condition of the samples because the distance between the samples would be less than the range.

- Practical implications

The technique suggested in this paper provide to farmers two options for soil management which are: 1) the individual samples may be mixed to obtain a composite sample and therefore obtain an average value for each of the attributes for the total area with more accuracy recommendation or 2) may be individually assessed in each zone for site-specific management.

To demonstrate the differences that may exist between the commonly used sampling technique and the technique recommended in the present study, a sampling and analysis procedure was simulated for the Barreiro de Santa Fé area. Twenty individual samples were collected (simulated) in zigzag to obtain a composite sample, which is based on classical statistics. Also, five individual samples were collected (simulated) spaced 180 m in the zones as suggested by the present study (Figure 5).

To exemplify the possible differences in fertilizer recommendation the P content was used. The mean value of P content using the 20 individual samples was 29.53 mg dm^{-3}, thus, the nutrient level was classified as medium (20.1 - 30 mg dm^{-3}) as suggested by Ribeiro et al. (1999). Therefore, the P recommendation would be 100 g plant^{-1}. On the other hand, the mean value of P content using the geostatistic approach was 70.65 mg dm^{-3} and the nutrient level being classified as very good (> 45 mg dm^{-3}) (RIBEIRO et al., 1999). In this case, the P recommendation would be 50 g plant^{-1}, that is, the half of the value obtained in the classic sampling method. Of course, the recommendation could be different in another simulation, but the technique addressed in this paper can guarantee the representativeness of the field heterogeneity whereas the classic method could not.

If the farmer opts by using site-specific management, the zone technique also would be useful. The K content was used to exemplify. The mean value of K content using the 20 individual samples was 0.42 cmol$_c$ dm^{-3}, thus, the nutrient level was classified as very good and the recommendation would be 120 g plant^{-1} (RIBEIRO et al., 1999). However, if the mean value from each zone would be considerate, the K content mean would be 0.33 cmol$_c$ dm^{-3} for zone 1 and 0.28 cmol$_c$ dm^{-3} for zone 2, and the nutrient level would be classified as good and medium, respectively. In this case, the K recommendation would

be 120 g plant^{-1} for zone 1 and 240 g plant^{-1} for zone 2 (RIBEIRO et al., 1999). As a result, the plant could not express its full production potential in approximately 30% of the area. Thus, it is evident the importance of indicating the number of representative individual samples, the distance at which these individual samples should be collected to ensure spatial independence and the place of collection, considering the heterogeneity of the sampling unit, to obtain a representative sampling.

In Brazil, most evaluations of chemical and physical properties of soil are still performed using various laboratory analyses of different levels of complexity. When carried out for a large number of samples, these analyses require more time, financial resources, and chemical reagents. Consequently, they generate large amounts of residues, and thus, environmental impacts (SILVA et al., 2017). Therefore, a significant increase in sampling costs may result with the use of geostatistics in the determination of the ideal number of individual samples, limiting its application, as the use of this technique would initially require analyses of at least 100 samples (50 at each layer) in areas of up to 10 hectares.

An alternative is the use of sensors that estimate soil attributes, increasing the accuracy and speed, and reducing the cost of the analyses, and consequently, enabling the application of this technique. Among these sensors, the portable X-ray fluorescence spectrometer can be highlighted, which is useful for determining the total content of chemical elements in soils, allowing the inference of some soil properties (SILVA et al., 2017; UDEIGWE et al., 2015). Also, Vis-NIR spectroscopy technique can be successfully used to estimate soil properties (KODAIRA; SHIBUSAWA, 2013; NOCITA et al., 2013). Another is the sensor that determines the soil magnetic susceptibility which can be used to plan, map, and estimate soil chemical elements (JAKŠÍK et al., 2016; TEIXEIRA et al., 2017) and the apparent electrical conductivity of soil, which has been used to estimate and map the spatial variability of soil properties in multiple cropping systems (GRUBBS et al., 2019; STADLER et al., 2015), among others sensors.

Conclusions

The number of individual samples needed to satisfactorily represent the mango fields, using the classical statistics, was much higher than the current recommendation, which would possibly make the collection impractical and/or prohibitively expensive. On the other hand, the addition of the spatial component, using geostatistics, allowed a reduction of about 70% in the recommended number (20 individual samples)

of soil individual samples to represent the three mango production fields;

The fuzzy c-means clustering technique was adequate to separate the areas into homogeneous zones in order to allocate the sampling points;

A specific geostatistical analysis is required for each area in order to determine the optimal sample density and allocate the samples, since the climate, soil type, texture, field size, cultivation and management influence soil variability;

Critical differences in fertilizer recommendation were observed when comparing the classic sampling method and the allocation method based on soil variability within the field.

Acknowledgment

References

ALVARES, C.A.; STAPE, J.L.; SENTELHAS, P.C.; MORAES, G.; LEONARDO, J.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, Stuttgard, v.22, n.6, p.711-728, 2013.

ARÉTOUYAP, Z.; NOUCK, P.M.; NOUAYOU, R.; KEMGANG, F.E.G.; TOKO, A.D.P.; ASFAHANI, J. Lessening the adverse effect of the semivariogram model selection on an interpolative survey using kriging technique. **SpringerPlus**, London, v.5, p.549-559, 2016.

BAZZI, C.L.; SOUZA, E.G.; URIBE-OPAZO, M.A.; NÓBREGA, L.H.P.; ROCHA, D.M. Management zones definition using soil chemical and physical attributes in a soybean área. **Engenharia Agrícola**, Jaboticabal, v.34, n.5, p.952-964, 2013.

CARVALHO, C.; KIST, B.B.; SANTOS, C.E.; TREICHEL, M.; FILTER, C.F. **Anuário brasileiro da fruticultura 2017**. Santa Cruz do Sul: Gazeta Santa Cruz, 2017. 88 p.

CARVALHO, J.R.P.D.; SILVEIRA, P.M.D.; VIEIRA, S.R. Geoestatística na determinação da variabilidade espacial de características químicas do solo sob diferentes preparos. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v.37, n.8, p.1151-1159, 2002.

CATANI, R.A.; GALLO, J.R.; GARGANTINI, H.; CONAGIN, A. Amostragem de solo para estudos de fertilidade. **Bragantia**, Campinas, v.14, p.19-26, 1954.

CASTIONI, G.A.F.; SOUZA, Z.M.; NAZARIO, A.A.; BORGES, B.M.N.N.; TORRES, J.L.R.; SOARES, M.D.R.; CAMPOS, M.C.C. Variability of physical attributes in tropical weathered soil cultivated with irrigated beans ('Phaseolus vulgaris' L.). **Australian Journal of Crop Science**, v.13, n.5, p.656-661, 2019.

CEDDIA, M.B.; VIEIRA, S.R.; VILLELA, A.L.O.; MOTA, L.S.; ANJOS, L.H.C.; CARVALHO, D.F. Topography and spatial variability of soil physical properties. **Scientia Agricola**, Piracicaba, v.66, p.338-352, 2009.

CLINE, M.G. Principles of soil sampling. **Soil Science**, Baltimore, v.58, n.2, p.275-288, 1944.

CUNHA, E.Q.; STONE, L.F.; MOREIRA, J.A.A.; FERREIRA, E.P.B.; DIDONET, A.D.; LEANDRO, W.M. Sistemas de preparo do solo e culturas de cobertura na produção orgânica de feijão e milho.I - Atributos físicos do solo. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.35, p.589-602, 2011.

FERRAZ, G.A.S; SILVA, F.M.; CARVALHO, L.C.C.; ALVES, M.C.; FRANCO, B.C. Variabilidade espacial e temporal do Fósforo, Potássio e da produtividade de uma lavoura cafeeira. **Engenharia Agrícola**, Jaboticabal, v.32, n.1, p.140-150, 2012.

FRIDGEN, J.J.; KITCHEN, N.R.; SUDDUTH, K.A.; DRUMMOND, S.T.; WIEBOLD, W.J.; FRAISSE, C.W. Management Zone Analyst (MZA): software for subfield management zone delineation. **Agronomy Journal**, Madison, v.96, n.1, p.100-108, 2004.

GRUBBS, R.A.; STRAW, C.M.; BOWLING, W.J.; RADCLIFFE, D.E.; TAYLOR, Z.; HENRY, G.M. Predicting spatial structure of soil physical and chemical properties of golf course fairways using an apparent electrical conductivity sensor. **Precision Agriculture**, Dordrecht, v.20, p.496-519, 2019.

JAKŠÍK, O.; KODEŠOVÁ, R.; KAPIČKA, A.; KLEMENT, A.; FÉR, M.; NIKODEM, A. Using magnetic susceptibility mapping for assessing soil degradation due to water erosion. **Soil & Water Research**, Praga, v.11, n.2, p.105-113, 2016.

KODAIRA, M.; SHIBUSAWA, S. Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping.**Geoderma**, Amsterdam, v.199, p.64-79, 2013.

LEMOS FILHO, L.C.A.; FERREIRA, L.L.N.; LYRA, D.L. Variabilidade espacial de atributos do solo indicadores de degradação ambiental em microbacia hidrográfica. **Revista Agroambiente**, Boa Vista, v.11, n.1, p.11-20, 2017.

LIMA, J.S.D.S.; SOUZA, G.S.D.; SILVA, S.A. Amostragem e variabilidade espacial de atributos químicos do solo em área de vegetação natural em regeneração. **Revista Árvore,** Viçosa, MG, v.34, p.127-136, 2010.

MALLARINO, A.P.; WITTRY, D.J. Efficacy of grid and zone soil sampling approaches for site-specific assessment of phosphorus, potassium, pH, and organic matter. **Precision Agriculture**, Dordrecht, v.5, p.131-144, 2004.

MALUF, H.J.G.M.; SOARES, E.M.B.; SILVA, I.R.; NEVES, J.C.L.; SILVA, M.F.O. Disponibilidade e recuperação de nutrientes de resíduos culturais em solo com diferentes texturas. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.39, p.1690-1702, 2015.

MONTANARI, R.; SOUZA, G.S.A.; PEREIRA, G.T.; MARQUES, J.; SIQUEIRA, D.S.; SIQUEIRA, G.M. The use of scaled semivariograms to plan soil sampling in sugarcane fields. **Precision Agriculture**, Dordrecht, v.13, p.542–552, 2012.

NOCITA, M.; STEVENS A.; NOON C.; WESEMAEL B. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. **Geoderma**, Amsterdam, v.199, p.37-42, 2013.

ODEH, I.O.A.; MCBRATNEY, A.B.; CHITTLEBOROUGH, D.J. Soil pattern recognition with fuzzy-c-means: application to classification and soil–landform interrelationships. **Soil Science Society of America Journal**, Madison, v.56, p.505–516, 1992.

OLIVEIRA, F.H.T.; ARRUDA, J.A.; SILVA, I.F.; ALVES, J.C. Amostragem para avaliação da fertilidade do solo em função do instrumento de coleta das amostras e de tipos de preparo do solo. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.31, p.973-983, 2007.

OLIVEIRA, I.A.D.; MARQUES JÚNIOR, J.; CAMPOS, M.C.C.; AQUINO, R.E.D.; FREITAS, L.D.; SIQUEIRA, D.S.; CUNHA, J.M.D. Variabilidade espacial e densidade amostral da suscetibilidade magnética e dos atributos de argissolos da região de Manicoré, AM. **Revista Brasileira de Ciência do Solo,** Viçosa, MG, v.39, p.668-681, 2015.

OLIVER, M.A.; WEBSTER, R.A tutorial guide to geostatistics: Computing and modelling variograms and kriging. **Catena**, Amsterdam, v.113, p.56-69, 2014.

PIMENTEL-GOMEZ, F.; GARCIA, C.H. **Estatística aplicada a experimentos agronômicos e florestais**: exposição com exemplos e orientações para uso de aplicativos. Piracicaba: FEALQ, 2002. 309 p.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2015.

REGELINK, I.C.; STOOF, C.R.; ROUSSEVA, S.; WENG, L.; LAIR, G.J.; KRAM, P.; NIKOLAIDIS, N.P.; KERCHEVA, M.; BANWART, S.; COMANS, R.N.J. Linkages between aggregate formation, porosity and soil chemical properties. **Geoderma**, Amsterdam, v.247-248, p.24-37, 2015.

RIBEIRO, A.C.; GUIMARÃES, P.T.G.; ALVAREZ, V.A.H. **Recomendação para uso de corretivos e fertilizantes em Minas Gerais - 5º aproximação**.Viçosa: Comissão de Fertilidade do Solo do Estado de Minas Gerais, 1999. 359 p.

RODRIGUES, M.S.; CORÁ, J.E.; CASTRIGNANÒ, A.; MUELLER, T.G.; RIENZI, E.A. Spatial and temporal prediction model of corn grain yield as a function of soil attributes. **Agronomy Journal**, Madison, v.105, p.1878-1887, 2013.

RODRIGUES, M.S.; CORÁ, J.E.; FERNANDES, C. Spatial relationships between soil attributes and corn yield in no-tillage system. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.36, p.599-609, 2012.

RODRIGUES, M.S.; SANTANA, M.C.; UCHÔA, A.L.P.; MENEZES, A.X.S.M.; CAVALCANTE, I.H.L.; LIMA, A.M.N. Delineation of management zones based on soil physical attributes in an irrigated guava field in the Semi-Arid region, Brazil. **African Journal of Agricultural Research**, Lagos, v.10, n.45, p.4185-4192, 2015.

RODRIGUES JÚNIOR, F.A.; VIEIRA, L.B.; QUEIROZ, D.M.; SANTOS, N.T. Geração de zonas de manejo para cafeicultura empregando-se sensor SPAD e análise foliar. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.15, n.8, p.778-787, 2011.

SANTOS, F.A.S.; PIERANGELI, M.A.P.; SILVA, F.L.; SERAFIM, M.E.; SOUZA, C.A. Atributos químicos, físicos e estoque de cálcio, magnésio, potássio e fósforo em solos de campos de murundus no Brasil. **Acta Amazonica**, Manaus, v.45, n.2, p.101-110, 2015.

SCHLINDWEIN, J.A.; ANGHINONI, I. Variabilidade horizontal de atributos de fertilidade e amostragem do solo no sistema plantio direto. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.24, n.1, p.85-91, 2000.

SEIDEL, E.J.; OLIVEIRA, M.S.D. A classification for a geostatistical index of spatial dependence. **Revista Brasileira de Ciência do Solo**, Viçosa, MG, v.40, p.1-10, 2016.

SILVA, J.P.S.; NASCIMENTO, C.W.A.; SILVA, D.J.; CUNHA, K.P.V.; BIONDI, C.M. Changes in soil fertility and mineral nutrition of mango orchards in São Francisco Valley, Brazil. **Revista Brasileira de Ciências Agrárias**, Recife, v.9, n.1, p.42-48, 2014.

SILVA, S.H.G.; TEIXEIRA, A.F.S.; MENEZES, M.D.; GUILHERME, L.R.G.; MOREIRA, F.M.S.; CURI, N. Multiple linear regression and random forest to predict and map soil properties using data from portable X-ray fluorescence spectrometer (pXRF). **Ciência e Agrotecnologia**, Lavras, v.41, n.6, p.648-664, 2017.

SÓRIA, J.E.; TAVANTI, R.F.R.; ALVES, M.; ANDREOTTI, M.; MONTANARI, R. Scaled semivariogram in the sample planning of soils cultivated with sugarcane. **Journal of Agricultural Science**,Cambridge, v.10, n.9, p.315-325, 2018.

SOUZA, Z.M.D.; MARQUES JÚNIOR, J.; PEREIRA, G.T.; MONTANARI, R. Otimização amostral de atributos de latossolos considerando aspectos solo-relevo. **Ciência Rural,** Santa Maria, v.36, p.829-836, 2006

STADLER, A.; RUDOLPH, S.; KUPISCH, M.; LANGENSIEPEN, M.; KRUK, J-V-D.; EWERT, F. Quantifying the effects of soil variability on crop growth using apparent soil electrical conductivity measurements. **European Journal of Agronomy**, Amsterdam, v.64, p.8-20, 2015.

STOLF, R.; THURLER, Á.D.M.; BACCHI, O.O.S.; REICHARDT, K. Method to estimate soil macroporosity and microporosity based on Areia content and bulk density. **Revista Brasileira de Ciência do Solo**, Viçosa, v.35, p.447-459, 2011.

TAGARAKIS, A.; LIAKOS, V.; FOUNTAS, S.; KOUNDOURAS, S.; GEMTOS, T.A. Management zones delineation using fuzzy clustering techniques in grapevines. **Precision Agriculture**, Dordrecht, v.14, p.18-39, 2013.

TEIXEIRA, D.D.B.; MARQUES JÚNIOR, J.; SIQUEIRA, D.S.; VASCONCELOS, V.; CARVALHO JÚNIOR, O.A.; MARTINS, E.S.; PEREIRA, G.T. Sample planning for quantifying and mapping magnetic susceptibility, Argila content, and base saturation using auxiliary information. **Geoderma**, Amsterdam, v.305, p.208-219, 2017a.

TEIXEIRA, P.C.; DONAGEMMA, G.G.; FONTANA, A.; TEIXEIRA, W.G. **Manual de métodos de análise de solos**. 3.ed. Brasília: Embrapa, 2017b. 573 p.

TESFAHUNEGN, G.B.; TAMENE, L.T.; VLEK, P.L.G. Catchment-scale spatial variability of soil properties and implications onsite-specific soil management in northern. **Soil & Tillage Research**, Amsterdam, v.117, p.124-139, 2011.

TREVISAN, R.G.; FREDDI, O.S.; WRUCK, F.J.; TAVANTI, R.F.R.; PERES, F.S.C. Variability of physical properties of soil and rice grown under cover crops in crop-livestock integrated system. **Bragantia**, Campinas, v.76, p.145-154, 2017.

UDEIGWE, T.K.; YOUNG, J.; KANDAKJI, T.; WEINDORF, D.C.; MAHMOUD, M.A.; STIETIYA, M.H.Elemental quantification, chemistry, and source apportionment in golf course facilities in a semi-arid urban landscape using a portable X-ray fluorescence spectrometer. **Solid Earth**, Munique, v.6, p.415–424, 2015.

VALENTE, D.S.M.; QUEIROZ, D.M.; PINTO, F.A.C.; SANTOS, N.T.SANTOS, F.L. Definition of management zones in coffee production fields based on apparent soil electrical conductivity. **Scientia Agricola**, Piracicaba, v.69, n.3, p.173-179, 2012.

VIEIRA, S.R.; CARVALHO, J.R.P.D; CEDDIA, M.B.; GONZÁLEZ, A.P. Detrending non stationary data for geostatistical applications. **Bragantia**, Campinas, v.69, p.1-8, 2010.

WANG, C.; ZUO, Q.; ZHANG, R. Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. **Journal of Hydrology**, Amsterdam, v.352, n.3-4, p.309-321, 2008.

YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística**: conceitos e aplicações. São Paulo: Oficina de Textos, 2013. 216 p.

YU, D-S.; ZHANG, Z-Q.; YANG, H.; SHI, X-Z.; TAN, M-Z.; SUN, W-Z; WANG, H-J. Effect of soil sampling density on detected spatial variability of soil organic carbon in a red soil region of china. **Pedosphere**, Beijing, v.21, n.2, p.207-213, 2011.

ZŮVALA, R.; FIŠEROVÁ, E.; MAREK, L. Mathematical aspects of the kriging applied on landslide in Halenkovice (Czech Republic). **Open Geosciences**, Varsóvia, v.8, n.1, p.275-288, 2016.