Smart Agricultural Technology 7 (2024) 100418

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Summarizing soil chemical variables into homogeneous management zones – case study in a specialty coffee crop

César de Oliveira Ferreira Silva [a], Célia Regina Grego [b], Rodrigo Lilla Manzione [c], Stanley Robson De Medeiros Oliveira [a,b], Gustavo Costa Rodrigues [b], Cristina Aparecida Gonçalves Rodrigues [d], Eduardo Antonio Speranza [b], Ariovaldo Luchiari Junior [b], Luciano Vieira Koenigkan [b]

[a] Universidade de Campinas, School of Agricultural Engineering, Av. Candido Rondon, 501, Campinas, 13083875, São Paulo, Brazil
[b] Embrapa Digital Agriculture, Av. Andre Tosello 209, Campinas, 13083886, São Paulo, Brazil
[c] Universidade Estadual Paulista "Júlio de Mesquita Filho", School of Science, Technology and Education, Av. Renato da Costa Lima, 451, Ourinhos, 19903-302, São Paulo, Brazil
[d] Embrapa Territorial, Av. Soldado Passarinho, 303, Campinas, 13070-115, São Paulo, Brazil

A R T I C L E I N F O

Editor: Spyros Fountas

Keywords:
Coffea arabica L.
Precision agriculture
Geostatistics
Indicator kriging

A B S T R A C T

Homogeneous management zones (HMZs) delineation is important for the application of precision agriculture because farm management decisions are based on it. Diverse soil chemical characteristics are important for the HMZs delineation. However, summarizing several variables into homogeneous zoning while taking into account the spatial distribution pattern of soil chemical characteristics is a challenge. Addressing this challenge is important to produce HMZs oriented for practical use for the farmers. In this work, 17 soil chemical variables were jointly analyzed for HMZ delineation by using indicator kriging (IK) to interpolate a soil fertility index (SFI). Soil samples were taken from a 4.5 ha area in a quasi-regular grid at 0 - 0.20 m depths in November 2019 (66 samples) and May 2021 (40 samples). Soil P, K, Ca, Mg, S, Na, H, Mn, Fe, Zn, B, cation exchange capacity, aluminum saturation, total organic carbon, base saturation, and organic matter were analyzed. In May 2021 the coffee yield was sampled together with the soil. Applying the SFI and then interpolating it using IK were effective for summarizing soil chemical variables into binary HMZs, showing a zone with higher priority for fertilization (therefore, lower general soil fertility) and another zone with low priority for fertilization. The summarizing process of several variables into binary HMZ was validated by evaluating the boxplots of each variable in each HMZ. Also, higher soil fertility areas presented a higher average coffee yield. Results indicated that joint use of SFI and IK was adequate to delineate HMZs in terms of summarizing soil fertility and separating coffee yield average variability. Delineating management zones by using the SFI approach is flexible for relatively limited sampled studies (less than one hundred samples) where machine learning and geostatistical methods may fail for lack of data.

1. Introduction

According to Barros et al. [12], coffee is one of the most important crops for the Brazilian economy. According to the Specialty Coffee Association [63] definitions, specialty coffee "refers to the highest quality green coffee beans roasted to their greatest flavor potential by true craftspeople and then properly brewed to well-established SCAA developed standards." These standards include scoring higher than 80 points on the quality scale and excellent or outstanding quality in fragrance, aroma, flavor, aftertaste, acidity, body, uniformity, balance, clean cup, sweetness, and overall better taste [63]. Specialty coffee refers to a modern demand for exceptional quality coffee, both farmed and brewed to a significantly higher than average standard and is related to the farmers and the brewer in what is known as the third wave of coffee [3].

Spatial and temporal variation in soil properties and meteorological conditions may affect coffee growth, grain development, quality, and yield [5,68,26,28,70,27,11,8,7,44,6]. To increase farmers' profitability and environmental protection, management practices need to adapt to

E-mail address: cesar.silva@feagri.unicamp.br (C.O.F. Silva).

https://doi.org/10.1016/j.atech.2024.100418
Received 18 December 2023; Received in revised form 16 February 2024; Accepted 17 February 2024
Available online 21 February 2024
2772-3755/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

variable site conditions [29,21], following the principles of Precision agriculture (PA) of considering its spatial variability [39].

A relevant topic on PA is the use of homogeneous management zones (HMZs), which are defined as sub-regions of a field and within which the effects on the crop of seasonal differences in weather, soil, management, etc. are expected to be uniform [56,18]. HMZs delineation is important for the application of Precision Agriculture because farm management decisions are based on it to make decisions on where and how much to apply when scheduling fertilization strategies [67]. However, high heterogeneity between different soil chemical variables makes summarizing it into unique zones a challenge that involves agronomical, mathematical, and computational aspects [22]. Ignoring the heterogeneity may result in the under-application or over-application of fertilizers at specific sites [30,16].

Most of the researches about HMZs focus on using ordinary kriging to interpolate variables to use them as input for clustering such as fuzzy k-means cluster analysis [33,48], fuzzy c-means clustering algorithm [10,46] or apply a dimensionality reduction tool over these interpolated maps as the principal components analysis (PCA) [53,78,41,23,48], spatially-weighting PCA [9,32], or factorial kriging [18,20,19]. According to the geostatistical paradigm, any soil or crop attribute is considered a random regionalized variable that varies continuously and its variation can be described by a spatial covariance function [47]. The kriging-based technique can provide the best linear unbiased estimation for the soil properties at unsampled locations [38,25]. However, the use of geostatistical methods as the ordinary kriging for interpolating maps is not enough to ensure the delineation of HMZs, since it requires summarizing several maps into a single one in a rational way in terms of agronomical and practical reasons, what is not always achievable by using pure computation methods like clustering [67,66,2]

Synthesizing different chemical variables is a challenge because when adding high variability to a model, the interpretation of the model becomes more complex [36]. Addressing this challenge leads to a gain of knowledge and applicability of the HMZs maps because usually, HMZs have an arbitrary number of zones, decided by the agronomical expert together with the farmer [40]. Nevertheless, in PA it may be sensible to divide the field into a few practical management zones. To ensure spatial contiguity because of spatially continuous variation [52,64,22], a probability-based approach in a smoothed manner may be used to make the HMZs more useful and applicable in the field.

In this context, the objective of this study was to propose a novel approach named soil fertility index (SFI) for summarizing several spatial variables in order to delineate binary homogeneous zones (HMZs). In the case study, we summarized 17 chemical soil variables in a coffee crop from Southern Minas Gerais over two years. Finally, we analyze the relation between HMZs delineation and observed coffee yield samples over the last year.

2. Material and methods

2.1. Description of the study site and agronomic practices

The field experiment was carried out at the municipality of Paraguaçu, southern Minas Gerais, Brazil, in a 4.5 ha area of coffee cultivation (*Coffea arabica L.*), cultivar Catucai Amarelo SL 134, transplanted in 2012, with spacing of 3.8 m between rows and 0.75 m between plants, totaling 3500 plants ha^{-1}. Uniform fertilization over the entire coffee plot was done directly on the soil in November 2019 and May 2021, by applying around 42, 10, and 42 kg.ha^{-1} of N, P, K.

The maximum altitude of this area is 894.3 m. The field boundaries and grid sampling point were delimited in the studied area using a GPS (Garmin GPSmap 62s). The soil of the area was classified as Oxisol according to the Brazilian soil classification system [62] and the local climate is characterized as mild, tropical high-altitude, with moderate temperatures, hot and rainy summer, classified as Cwa according to Köppen's classification. Fig. 1 shows the study area and soil sampling distribution in November 2019 and May 2021 overlapping an image taken on April 11 2019 by an original non-interchangeable camera onboard an unmanned aerial vehicle (UAV) model DJI Mavic 2 Pro, using an interface software "DJI Ground Station Pro". The mosaic building software was the OpenDroneMap [51] with the flight at 40 m height. The image has a ground sample distance of 0.94 cm/$pixel^{-1}$.

2.2. Soil sampling and laboratory analyses

A quasi-regular sampling grid was delimited in the studied area with points spaced by grid variable by 20 m x 10 m, totaling 66 georeferenced sampling points in November 2019, and with points spaced by grid variable by 30 m × 10 m, totaling 40 georeferenced sampling points in May 2021. In May 2021 coffee yield was sampled together with soil in groups of 2 trees around the spatial position of soil samples; for this reason, the sampling grid changed. Silva et al. [67] explored how soil samplings for geostatistical studies in Brazil are low sampled, in most of the cases less than one hundred sample points are collected, even for areas larger than 100 meters.

Soil samplings were performed by collecting subsamples under the crown projection in the layer of 0-20 cm, using a Dutch auger, in each plant composing the sampling point. These subsamples were homogenized to form a composite sample representative of the point in question and sent to the Laboratory of Soil Analysis. The following soil chemical attributes were evaluated: pH ($CaCl_2$ extractor), availability of phosphorus (P) (Mehlich), availability of potassium (K) (Mehlich 1 extractor), availability of sodium (Na) (Mehlich 1 extractor), availability of iron (Fe) (Mehlich 1 extractor), availability of manganese (Mn) (Mehlich 1 extractor), availability of zinc (Zn) (Mehlich 1 extractor), availability of boron (B) (wet extractor), exchangeable calcium (Ca^{2+}) (1 mol L^{-1} KCL extractor), exchangeable magnesium (Mg^{2+}) (1 mol L^{-1} KCL extractor), sulfur (S) (phosphate extractor), aluminum saturation (m), hydrogen (H), cation exchange capacity (CEC), base saturation (V), and organic matter (OM), following the methodology described by Alvarez et al. [4].

2.3. Indicator kriging (IK)

Indicator kriging (IK) is based on assigning a binary indicator of 1 or 0 depending on whether the observation is greater or less than a threshold value. Also, IK is a non-parametric approach. Non-parametric geostatistics is based on no assumed parametric model of error distribution and makes the modeling of uncertainty a priority [35]. This uncertainty is modeled as a probability distribution of the variable of interest rather than an estimation error, as in kriging [65]. For the IK approach, the experimental indicator semivariogram was estimated by Eq. (1):

$$\gamma I(h, z_k) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(u_a, z_k) - I(u_a + h, z_k)]^2 \tag{1}$$

where $\gamma I(h, z_k)$ is the estimated semivariogram; $I(u_a, z_k)$ and $I(u_a + h, z_k)$ are the observed values of the indicator coding at location u_a and $u_a + h$; $N(h)$ is the number of observation pairs separated by distance h.

The indicator semivariogram measures how often two values of Z separated by a vector h are on opposite sides of the threshold value z_k.

IK requires that the attribute values be modified according to a non-linear transformation, which is called indicator coding. Lloyd and Atkinson [45] cite that indicator kriging can also be calculated from cutoffs of a continuous variable. Statistical parameters and geostatistical analyses were performed for all variables, focusing on the spatial continuity and dependence of chemical soil properties.

2.4. The soil fertility index (SFI)

For this work, using a new binary variable F for each soil variable observed value is proposed according to Eq. (2):

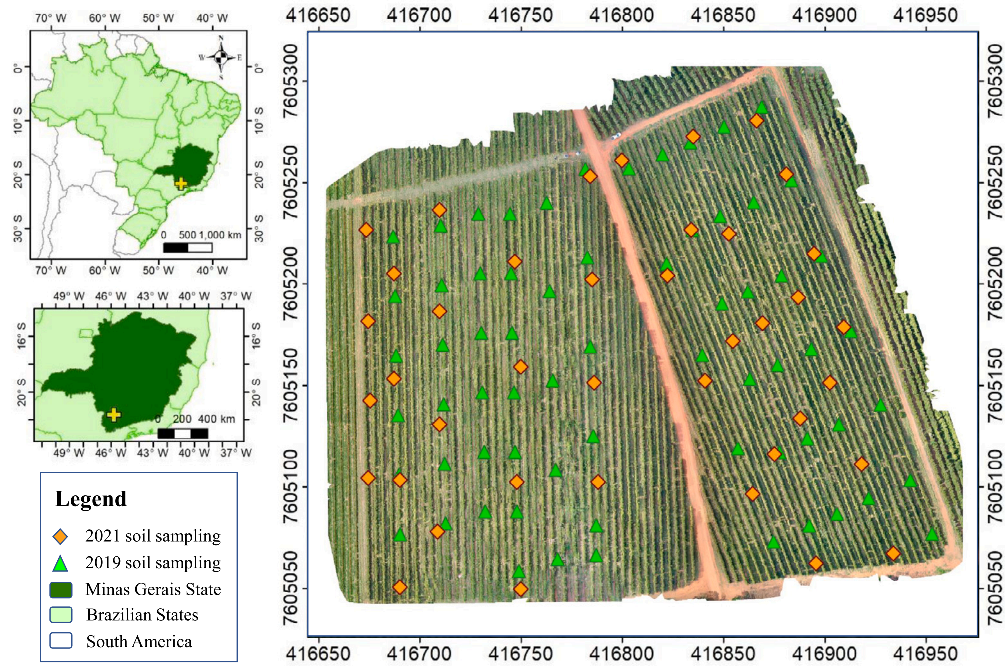

Fig. 1. UAV image of the study area with soil sampling points in November 2019 and May 2021.

$$\left\{ \begin{array}{ll} F = 0 & \text{if Z is classified as "bad"} \\ F = 1 & \text{if Z is classified as "good"} \end{array} \right\} \tag{2}$$

where F is the binary indicator of fertility status and Z is the observed values of each variable.

Using a simple percentage index (Eq. (3)) to convert single indicator rules (from Eq. (2)) into the "fertilization needs" index (SFI), the higher this percentage is, the higher the number of variables are classified as "higher fertilization need":

$$SFI = 100 \sum_{i=1}^{N} \frac{F_i}{N} \tag{3}$$

where F_i is the binary indicator of fertilization needs for each soil variable observed value, N is the number of variables.

The IK semivariogram will be set by applying an arbitrary threshold value (z_k) for the SFI.

2.5. Performance evaluation

Prior to the application of IK algorithms, the SFI observed values were checked by the Kolmogorov–Smirnov (K–S) test to verify whether they were normally distributed [42].

The evaluation of the performance of IK application was performed based on the mean error (ME), mean squared error (MSE), kriged reduced mean error (KRME), and kriged reduced mean squared error (KRMSE) generated by the cross-validation procedure. The ME (Eq. (4)), MSE (Eq. (5)), KRME (Eq. (6)), and KRMSE (Eq. (7)) of the IK estimates were calculated accordingly:

$$ME = \frac{1}{N} \sum_{i=1}^{N} (z_{0,1} - I_{p,i}) \tag{4}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (z_{0,1} - I_{p,i})^2 \tag{5}$$

$$KRME = \frac{1}{N} \sum_{i=1}^{N} \frac{(z_{0,1} - I_{p,i})}{s} \tag{6}$$

$$KRMSE = \frac{1}{N} \sum_{i=1}^{N} \left[\frac{(z_{0,1} - I_{p,i})}{s} \right]^2 \tag{7}$$

where $z_{0,1}$ is the observed value at location i, $I_{p,i}$ is the predicted value at location i, N is the number of pairs of observed and predicted values, and s is the standard deviation of the observed values.

The mean error and kriged reduced mean error values near zero indicates a better model prediction performance [66]. As a practical rule, the MSE should be less than the variance of the sample values and KRMSE should be in the range $1 \pm (2\sqrt{2})/\mathrm{N}$ [1].

The spatial dependence ratio was calculated as the proportion (percentage) of the nugget effect in relation to the sill (Eq. (8)), which, according to [17], can be classified as: (a) strong spatial dependence, < 25%; (b) moderate spatial dependence, 25 to 75%; and (c) weak spatial dependence, > 75%.

$$DD = \frac{C_0}{C_0 + C} \times 100 \tag{8}$$

where C_0 is the nugget effect and $C_0 + C$ is the sill.

The nugget effect (C_0) is a parameter of the semivariogram that shows the non-explained variability, which may occur due to measurement errors or non-detected variation by the sampling scale. The structural component (C) corresponds to the difference between the sill and the nugget effect and represents the spatially structured semivariance [17]. The nugget effect was expressed in relation to the sill to simplify the comparison of the spatial dependence degree (DD) of the studied variables.

2.6. Methodology overview and validation of the HMZ

Fig. 2 presents an overview of the methodology presented in previous sections. This is a general methodology, applicable for different datasets and summarizing needs. A validation strategy for applications in coffee crops is presented here.

Boxplots of the soil chemical variable samples in each HMZ obtained by IK were generated to evaluate each variable individually. It could indicate the ability of the methodology to summarize and generalize several soil chemical variables without losing local information, crucial for the actual application of fertilizers.

A boxplot of coffee yield in each HMZ in May 2021 was generated to assess the ability of the methodology to significantly separate higher from lower yields following the soil fertility. Also, analysis of variance

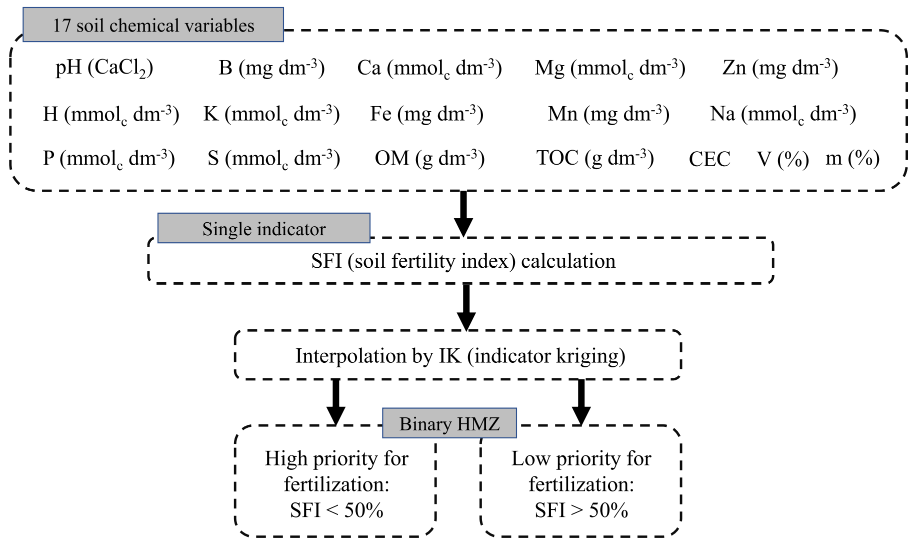

Fig. 2. Overview of the methodology for summarizing several soil chemical variables and producing binary HMZs.

(ANOVA) was performed to assess the significance of coffee yield differences in each HMZ.

2.7. Open-source software for research

All statistical and modeling procedures were performed with the R software [58] using the following packages: data.table [24], devtools [77], ggplot2 [76], gstat [55], and tidyverse [75].

3. Results

3.1. Plot characterization

The basic statistical parameters of all the analyzed variables are presented in Table 1. The higher the CV, the more heterogeneous the data set. Soil pH (extracted using $CaCl_2$), OM, TOC, CEC, Fe, and H had medium variability, with a coefficient of variation (CV) between 10% and 30%. All other soil chemical variables present high variability, with a coefficient of variation (CV) higher than 30%. pH values showed a trend in soil acidity, presenting an average pH of 4.55. The results referring to the variables pH, Ca, and Mg are similar to those reported by Santos et al. [61]. The result of the variables P, K, CEC, *m* and OM is similar to that found in [69].

A Pearson correlation matrix summarized the relationship between the soil chemical variables and yield data of the coffee datasets (Table 2). The correlation matrix showed both positive and negative correlations between some soil nutrients. However, the correlation matrix did not provide a clear interpretation regarding soil nutrients concerning yield suggesting the relationships may be non-linear and a more complex approach is needed to find HMZs.

Considering the high variability of almost all soil chemical variables, the study area can be considered suitable for receiving different and site-specific management practices as long the soil samples show different chemical behaviors in different spatial positions of the coffee crop. As there are heterogeneous variables that generate variability in crop yields, the most used techniques in data analysis for developing HMZs are the geomathematical and geostatistical models [13].

The 2021 frost caused damage to crops, and one of the factors is the low level of K, which is one of the main drivers of frost resistance [71,34].

3.2. Conversion of soil chemical variables into SFI

The soil chemical variables were transformed into probability indicators following Eq. (2) by using the cutoff thresholds presented in Table 3.

After performing the coding step, the SFI (Eq. (3)) was calculated. Mean, variance, coefficient of variation, minimum value, maximum value, K-S value, skewness, and kurtosis of SFI were obtained (Table 4) to verify the existence of a central trend and to determine the dispersion of the index values. According to [14], these standard statistical parameters are useful in evaluating the magnitude of the data dispersion around a central tendency value. Based on the values presented in Table 3, the SFI is distributed normally, as indicated by the proximity to zero for the skewness and kurtosis coefficients [66]. Also, K-S values are inside the range of its critical-values ($\alpha = 0.34$). Data normality does not constitute a requisite in geostatistics [54] but it has been noticed empirically as an important property for improving semivariogram structuring [14]. According to the preliminary exploratory analysis, it was assumed that distributions were symmetric enough and long-tailed. It was also assumed a non-occurrence of proportional effect, enabling the development of well-defined semivariograms.

3.3. Definition of the HMZs by using indicator kriging (IK)

The geostatistical analysis of SFI for interpolation was performed by fitting the theoretical semivariogram model to the experimental semivariogram. The IK interpolation is based on this model. The main variables of this analysis are the nugget effect (C_0) and the structural component (C). The range (A) of the experimental semivariogram can determine the spatial dependence limit; thus, samples that show distances between them higher than the value of the range have random distributions and are independent of each other, with no restrictions for the use of classic statistics [54].

Anisotropic semivariograms of SFI were simultaneously calculated for four directions with 45° (not shown) angular increments and ± 22.5° angular tolerance. No sign of relevant anisotropy was found in the observations. A relevant anisotropy sign could be found when the sill, range, and nugget are different in different directions. The spherical model was best adjusted to the experimental semivariogram. Fig. 3 presents the experimental omnidirectional semivariogram (points) and the spherical model fitted (curve). Table 5 shows the parameters for the theoretical semivariogram model. Here, a DD of almost 48% and 51% were achieved in November 2019 and May 2021 respectively, showing that the SFI presented moderate spatial dependence.

Table 1
Descriptive statistics of soil chemical variables in the coffee crop.

	November 2019				May 2021			
Attributes	Min	Mean ± SD	Max	CV	Min	Mean ± SD	Max	CV
pH ($CaCl_2$)	3.80	4.55 ± 0.52	5.60	11.34	3.8	4.34 ± 0.41	5.3	9.45
CEC	68.00	101.76 ± 17.39	144.30	17.09	72.1	100.79 ± 12.81	130.8	12.71
P (mmolc dm^{-3})	2.80	73.83 ± 70.63	388.6	95.66	28	105.78 ± 58.31	332	55.12
K (mmolc dm^{-3})	1.70	4.72 ± 1.90	10.30	40.26	0.52	1.34 ± 0.84	4.2	62.69
Ca (mmolc dm^{-3})	5.00	22.35 ± 12.37	56.00	55.35	3	17.23 ± 13.14	50	76.26
Mg (mmolc dm^{-3})	1.00	6.56 ± 4.17	17.00	63.56	1	5.95 ± 4.3	21	72.27
S (mmolc dm^{-3})	8.00	22.47 ± 11.57	62.00	51.48	5	31.2 ± 32.51	201	99.23
Na (mmolc dm^{-3})	0.00	0.16 ± 0.13	0.80	78.67	0.07	0.13 ± 0.08	0.4	61.54
H (mmolc dm^{-3})	20.00	63.67 ± 24.96	117.00	39.20	30	72.15 ± 21.87	105	30.31
V (%)	6.00	35.36 ± 20.32	76.00	57.46	5	25.6 ± 19.09	70	74.57
m (%)	0.00	15.34 ± 14.51	63.03	94.60	0	21.03 ± 20.91	64.52	99.43
OM (g dm^{-3})	17.00	23.33 ± 3.57	38.00	15.32	15	26.38 ± 3.22	33	12.21
TOC (g dm^{-3})	10.00	13.56 ± 2.12	22.00	15.64	9	15.38 ± 1.81	19	11.77
Zn (mg dm^{-3})	1.90	7.25 ± 4.9	22.50	67.56	2.4	7.06 ± 4.35	21.5	61.61
Mn (mg dm^{-3})	1.00	9.70 ± 5.91	24.50	60.87	3.5	12.16 ± 7.62	28.5	62.66
Fe (mg dm^{-3})	11.00	21.52 ± 4.76	34.00	22.11	27	35.25 ± 5.92	52	16.79
B (mg dm^{-3})	0.67	1.34 ± 0.46	2.12	34.07	0.17	0.54 ± 0.26	1.63	48.15
Yield (kg.2 $trees^{-1}$)	-	-	-	-	0.27	3.05 ± 2.60	10.18	85.26

Min – Minimum value; Max – Maximum value; CEC – cation exchange capacity; SD – Standard deviation; CV – Coefficient of variation; OM – organic matter; m – Aluminum saturation; TOC – Total Organic Carbon; V – Base saturation.

Table 2
Pearson correlation matrix for coffee soil chemical attributes and yield (Y) in May 2021 at significance of $p < 0.05$. Bold values are higher or equal to the absolute value of 0.7.

	P	OM	TOC	pH	K	Ca	Mg	Na	H	CEC	V	m	S	B	Fe	Mn	Zn	Y
P	1.0																	
OM	0.4	1.0																
TOC	0.4	**1.0**	1.0															
pH	-0.3	-0.2	-0.2	1.0														
K	-0.3	-0.3	-0.2	0.5	1.0													
Ca	-0.2	-0.1	-0.1	**1.0**	0.5	1.0												
Mg	-0.1	0.0	0.0	**0.9**	0.3	**0.9**	1.0											
Na	-0.2	0.0	0.0	0.5	0.1	0.5	0.6	1.0										
H	0.4	0.4	0.4	**-0.9**	-0.6	**-0.9**	**-0.7**	-0.4	1.0									
CEC	0.5	0.6	0.6	-0.5	-0.5	-0.3	-0.2	0.0	**0.7**	1.0								
V	-0.3	-0.2	-0.2	**1.0**	0.6	**1.0**	**0.9**	0.5	**-0.9**	-0.5	1.0							
m	0.3	0.2	0.1	**-0.8**	-0.5	**-0.8**	**-0.7**	-0.3	**0.8**	0.6	-0.8	1.0						
S	0.2	0.3	0.4	-0.1	-0.4	0.0	0.0	0.1	0.1	0.2	-0.1	0.0	1.0					
B	-0.3	-0.2	-0.2	0.5	0.5	0.4	0.3	0.0	-0.3	-0.2	0.4	-0.2	-0.1	1.0				
Fe	0.1	-0.1	-0.1	-0.5	-0.3	-0.5	-0.4	-0.1	0.5	0.3	-0.5	0.6	0.1	-0.2	1.0			
Mn	-0.4	-0.1	-0.1	**0.7**	0.5	**0.7**	0.5	0.4	**-0.7**	-0.4	**0.7**	-0.6	-0.2	0.3	-0.6	1.0		
Zn	-0.3	-0.1	-0.1	**0.7**	0.4	**0.7**	0.6	0.5	-0.6	-0.2	**0.7**	-0.4	-0.2	0.2	-0.5	**0.7**	1.0	
Y	0.1	0.2	0.2	-0.3	-0.3	-0.3	-0.3	0.1	0.4	0.3	-0.4	0.4	0.3	-0.2	0.3	-0.4	-0.2	1.0

CEC – cation exchange capacity; OM – organic matter; m – Aluminum saturation; TOC – Total Organic Carbon; V – Base saturation; Y - Coffee yield.

Table 3
Conditional cutoff thresholds for soil chemical attributes under coffee cultivation. If conditional is satisfied the $F = 1$, elsewise, 0.

Attributes	Condition	Attributes	Condition
V (%)	> 50.00[a]	m (%)	< 50.00[a]
pH ($CaCl_2$)	> 5.40[a]	OM (g dm^{-3})	> 4.00[a]
CEC	> 80.00[b]	TOC (g dm^{-3})	> 15.00[c]
H (mmolc dm^{-3})	> 36.00[b]	Zn (mg dm^{-3})	> 3.00[c]
P (mmolc dm^{-3})	> 50.00[b]	Mn (mg dm^{-3})	> 20.00[c]
K (mmolc dm^{-3})	> 4.00[b]	Fe (mg dm^{-3})	> 1.50[c]
Ca (mmolc dm^{-3})	> 30.00[b]	B (mg dm^{-3})	> 1.00[c]
Mg (mmolc dm^{-3})	> 10.00[b]	S (mmolc dm^{-3})	> 10.00[c]
Na (mmolc dm^{-3})	> 0.50[c]		

[a] [4]
[b] [74]
[c] [37]

Table 4
Descriptive statistics for SFI under a coffee crop.

Statistical parameters	SFI (2019)	SFI (2021)
N	66	40
Minimum	37.50	31.25
Maximum	81.25	75.00
μ	56.15	51.09
Median	56.25	50.00
σ	12.23	8.94
Variance	149.63	79.90
CV (%)	21.78	17.50
Kurtosis	-0.48	-0.40
Skewness	0.34	0.16
K-S	0.25	0.29

N – number of samples; CV - Coefficient of variation equals standard deviation (σ) divided by sample mean (μ); $K - S$ – Kolmogorov–Smirnov test.

The range of the semivariogram is the maximum distance of the correlated measurements. It can be a sufficient criterion for the selection of sampling design in mapping soil properties [73,15]. In both years, the

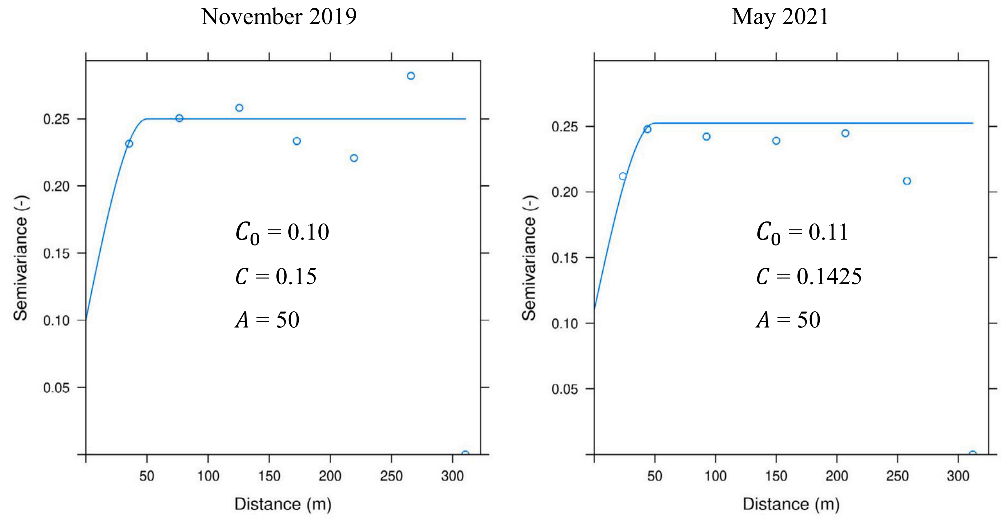

Fig. 3. Experimental semivariogram and theoretical semivariogram model for the SFI under coffee cultivation in November 2019 and May 2021.

Table 5
Parameters for spherical semivariograms models for SFI in November 2019 and May 2021.

Statistical parameters	C_0	C	A	DD (%)
2019	0.10	0.11	50.00	47.62
2021	0.15	0.1425	50.00	50.93

C_0 – nugget effect; C – sill; A – range; DD - spatial dependency index from [17].

Table 6
Cross-validation results for IK.

Evaluation metric	ME	MSE	KRME	KRMSE
2019	0.010	0.212	0.023	0.965
2021	0.017	0.235	0.031	0.984

ME – mean error; MSE – mean square error; KRME – kriged reduced mean error; KRMSE – kriged reduced mean square error; IK – indicator kriging.

best semivariogram model was found with the same range, confirming the spatial dependency with values under a sampling interval lower than 50 m (Table 5).

Fig. 4 presents the kriged map created based on a semivariogram analysis of SFI. Fig. 4a;c shows the estimation of SFI spatial distribution across the coffee crop. In a binary approach, two zones can be produced to separate the field into a higher priority for agronomical interventions and a lower priority. The HMZs map (Fig. 4b;d) was developed by considering pixels with SFI < 0.5 as an area with higher priority for fertilization, otherwise, the pixels were considered as an area with lower priority for fertilization.

The error statistics such as ME, MSE, KRME, and KRMSE were estimated and presented in Table 6. The error terms ME, and KRME were close to zero and KRMSE were close to one. It indicates the IK modeling presents good results.

The remarkable continuity and smoothness with increasing distance between samples is of importance. Together with the persistence of the range (A) of the semivariogram in both years, this shows that HMZs delineation was effective and the main future improvement will be sample densification with distances between samples smaller than A.

Summing up, the maps of IK graphically realize the fusion of the 17 soil chemical variables into a single indicator, providing a partition of the field into binary zones with a higher fertilization need (SFI < 50%) and a lower need (SFI > 50%). However, in order to interpret the variations shown in the previous maps (Fig. 4) in terms of actual soil characteristics, which could have an impact on agronomic management, a direct comparison between HMZs based on the soil sampling of the shallow 0–0.20-m depth was necessary.

3.4. Validation of the HMZs

Figs. 5 and 6 show the boxplot of each soil chemical variable in each HMZ in November 2019 and May 2021 respectively. The soils from both HMZs had coherent soil chemical behaviors. Soils on low priority for fertilization zones presented higher values of macronutrients and micronutrients, and lower values of H and m. Soils on higher priority for fertilization zones effectively presented lower nutrient concentrations and higher values of H and m. These coherent results demonstrate that summarizing information by using the SFI preserved the spatial information from the soil samples and avoided arbitrary choices about the number of HMZs. Incorporating information improves automatization and promotes a more focused participation of the farmers in the final decisions instead of modeling decisions, making the process less subjective [49].

Fig. 7 shows the spatial distribution of coffee yield samples across the HMZs and the boxplot in each HMZ. The outlier P20 (20th soil sample in the plot b) is closer to the boundary between the binary HMZs, indicating a lack of information in this area.

15 coffee yield samples are included in the low priority for fertilization area, presenting an average yield of 3.2 ± 2.87 kg.2 trees^{-1}, while 23 coffee yield samples are included in the high priority for fertilization area, presenting an average yield of 2.21 ± 1.93 kg.2 trees^{-1}. According to the ANOVA analysis (Table 7), no significant difference can be inferred from the coffee yield when comparing the different HMZs. The reason is soil fertility cannot explain all the coffee yield variation since weather extreme events in Southeast Brazil negatively impacted and devastated crops in mid-2021.

4. Discussion

4.1. Incorporating soil chemical summarized information into precision agriculture

In the present article, the methodology is focused on converting the relevant variables on soil fertility based on agronomical knowledge. The results of the present methodology (Fig. 2) showed a good summarizing ability, which produced HMZs that separated several soil chemical variables into two zones and the information was well preserved as shown by the boxplots (Fig. 5 and 6). This separability between HMZs can be seen in the boxplots, as the boxplot in the high-priority fertilization

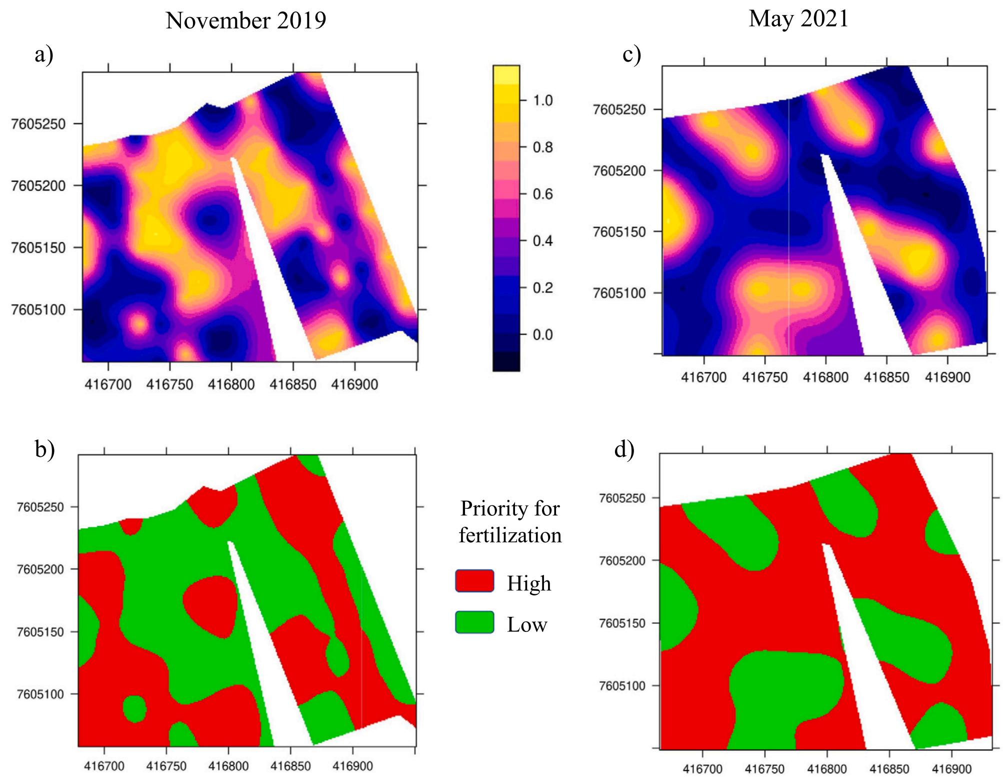

Fig. 4. Maps of IK (a;c) of the SFI index and binary HMZs (b, d) for a coffee crop from Southern Minas Gerais in November 2019 and May 2021. Legend is in the middle.

Table 7
ANOVA (single factor) analysis grouping the coffee yield in the two HMZs (high and low priority for fertilization) in May 2021. Significant different is identified for F > Critical value of F. SS – sum of squares; df – degrees of freedom; MS – mean squares.

Source of Variation	SS	df	MS	F	P-value	Critical F-value
Between HMZs	4.42	1	4.42	0.76	0.40	4.11
Within HMZs	210.38	36	5.84			
Total	214.80	37				

zone shows lower values for the soil chemical variables, i.e. these are the zones that concentrate the lowest fertility.

Different soil variables can indicate different aspects of soil fertility. According to [60], pH is an indicator of the biological-physical-chemical condition of the soil. An excessively acidic soil (pH very low) or excessively alkaline soil (pH very high) is less favorable for agriculture because there will be less oxygen, less organic matter, less water retention and infiltration, and more toxic ions. High soil acidity can generate high levels of Al^{3+} and Ca while Mg deficiency in the plant, affects the full development of plants and the achievement of high yields since the low pH reduces the availability of some nutrients and increases the toxic effect of aluminum on plants [74]. The macronutrients Ca and Al are constituents of the minerals and organic matter of the substrate where the plant develops and are also found dissolved in the soil solution [60]. On the other hand, the values of m are the percentage of soluble aluminum in relation to the exchangeable base (Ca^{2+}, Mg^{2+}, K^{+}) and aluminum content in the soil CEC. Therefore, it determines the alitic or aluminaic characteristics of the soil.

V represents the percentage of CEC occupied by bases (Ca^{2+}, Mg^{2+}, K^{+}, and Na^{+}) about the exchange capacity determined at pH 7. At pH 7 soils were considered 100% base-saturated and had zero base saturation at pH 4. Soils with V equal to or greater than 50% are denominated eutrophic soils (tending to present higher fertility). Soil with base saturation values less than 50% are denominated dystrophic soils (tending to present lower fertility). Base saturation (V) can indicate the amount of cations, such as Ca, Mg, and K, and identify if the soil is acidic at a level that is harmful to the crop. The calculation of Al saturation (m) is considered the most correct form to evaluate Al toxicity in the soil. The soil OM contributes to an increase in soil CEC which can serve to retain and increase the reserve of soil cations and improve soil structure physics and soil water relations. Soils with higher OM content are associated with increased population and diversity of microorganisms [50].

It is important to underline that the coffee yield showed a numerical difference (although not statistically significant) on the different HMZs. The coffee yield is an independent variable in the present methodology, different from the methods based on PCA and clustering. Here, the indicators are based on individual cutoffs and agronomic knowledge, and the variables used in the modeling are not necessarily chosen based on their correlation with coffee yield. This methodology can be particularly useful in situations where resources are scarce for large samples, and agronomic knowledge can be used to compensate for the smaller number of samples.

4.2. Applicability of HMZs from SFI and IK

By identifying fields where the relationship between soil and yields is fundamentally different from others, unique management areas could

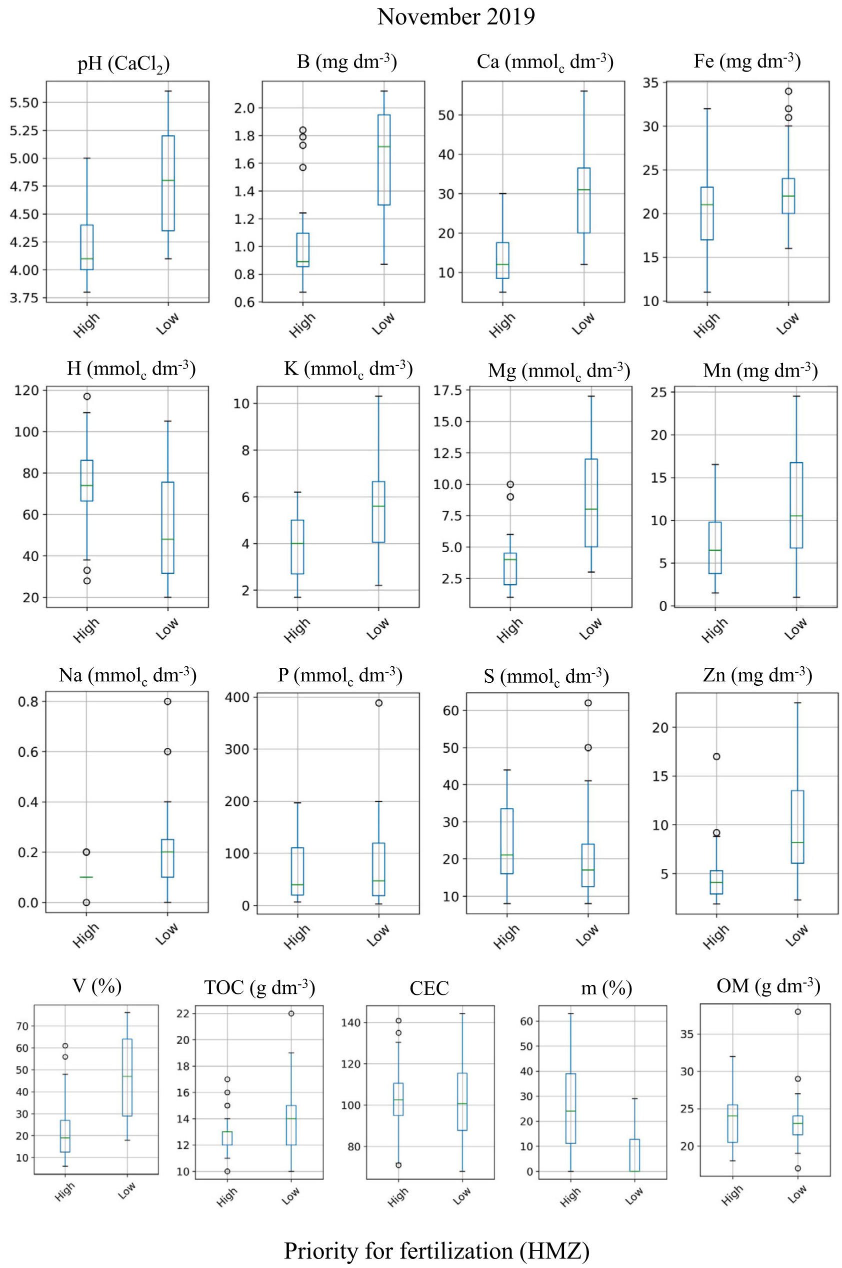

Fig. 5. Boxplots of 17 soil chemical variables sampled in November 2019 grouped by the binary HMZs for a coffee crop from Southern Minas Gerais.

be delineated. Usually, HMZs are delineated from a single variable, such as altitude [40] or soil apparent electrical conductivity [56,64]. However, multivariate approaches for delineating HMZs are needed for a more general and comprehensive application of PA [52,49,22]. Also, making the HMZ a brief spatial summary of several characteristics from the field is a challenge because both spatial and intrinsic heterogeneity make the HMZs too general for real applications [22]. Reduce the scope of the HMZs into a group of variables like the soil fertility, or weather drivers, yield factor, etc, could make multivariate approaches easier to understand in the field and more appropriate for the practical use by the farmers in the field.

In the multivariate approach, the most commonly used methodologies are the joint use of dimensionality reduction techniques such as PCA and clustering techniques such as c-means or k-means. When applying a PCA, the research objective usually is to find the key loading factors to explain a main independent factor. For example, in [56] soil properties and nutrient concentrations were compared with apparent electrical conductivity (ECa) using principal components (PC)-stepwise regression and ANOVA. The dimensionality reduction in this case means using fewer variables to explain another one with regression modeling. In a second moment, clustering is applied to group the principal components into groups (the HMZs).

Our results suggest that the binary HMZs are well suited for use in precision agriculture where the assessment of within-field variation is assumed as a basis for variable-rate application (VRT) of agronomic inputs. For farming by HMZs to be effective, there should be a strong and persistent relationship between the HMZs, delineated with soil and crop properties, and the spatial patterns in yield over time.

The area of most intensive effort within PA has been the site-specific management of nutrients (SSNM) guided for HMZs, generally via variable rate technology (VRT). This concept of site-specific management is intuitively appealing [57], however, its adoption has been slower than expected because it involves social, agronomic, economic, and technological changes [59].

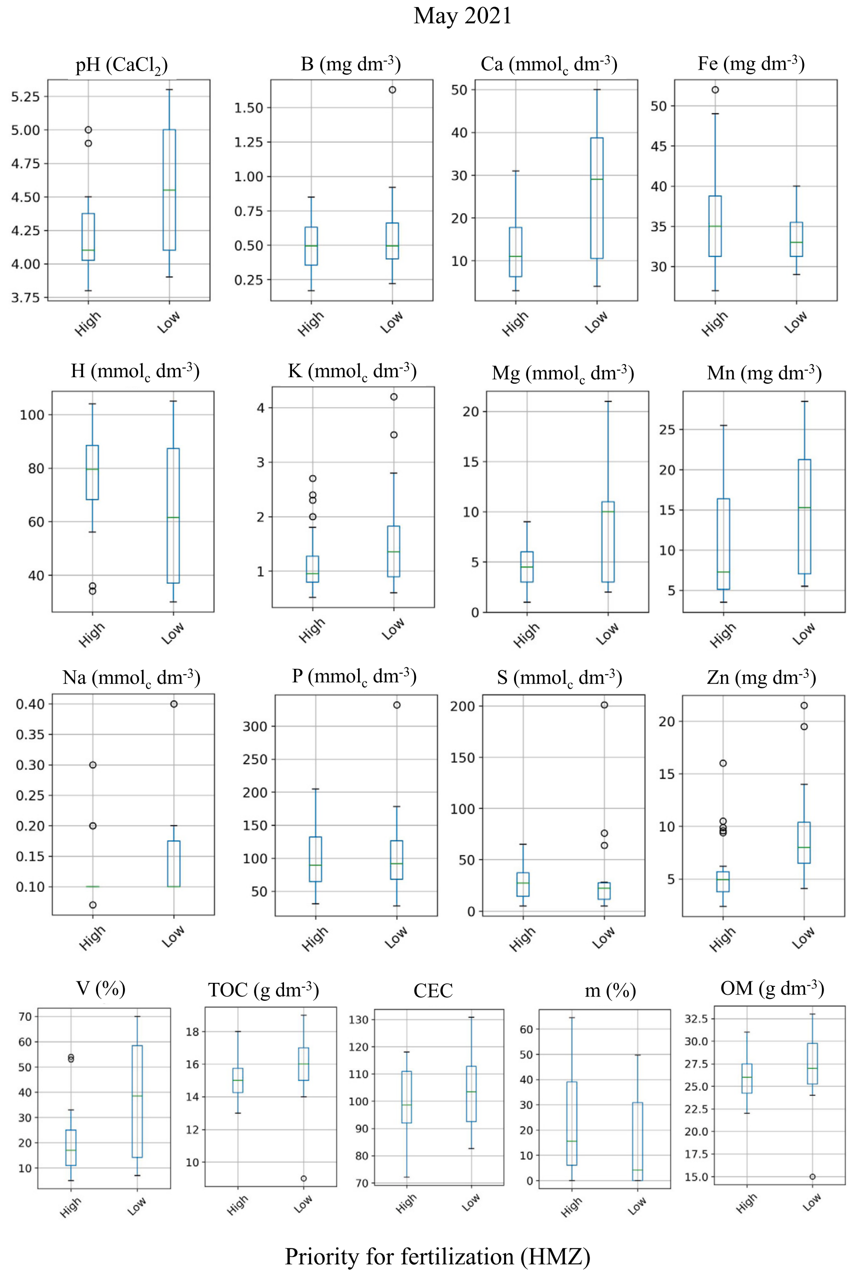

Fig. 6. Boxplots of 17 soil chemical variables sampled in May 2021 grouped by the binary HMZs for a coffee crop from Southern Minas Gerais.

A smoothed zoning as presented here, without fragmented spatial distribution, is highly desirable, because usually high accuracy classifications and clustering methods produce unrealistic maps for site-specific management, with isolated pixels and corners which are ignored in real applications. Also, a smoothed zoning doesn't need to be redesigned by hand at the end of modeling, just needs a threshold to separate different zones. A binary approach was adopted, but to produce more zones, the maps of IK (Fig. 4a;c) of the SFI index can be split into more than one cutoff value, producing more than two zones.

In this work, two years of data were used separately from coffee plots under uniform management. For this reason, in practice, HMZs obtained from May 2021 would be actually used in a PA strategy.

4.3. Temporal instability of HMZs and nutrient mobility on soil

The temporal stability of the HMZs is another important issue in site-specific agriculture, which might require the extension of the current methodology for new experiments with the same sampling design over different periods. The instability of spatial variability is due to the mobility of soil nutrients, which are transferred from year to year, and nutrient transport rates depend on local conditions [43,72]. The most mobile nutrients are usually phosphorus (P), potassium (K), and nitrogen (N) [72]. For example, [43] found that P transport rates were heterogeneous due to local topographic and chemical variations in the soil.

Checking the plot characterization shown in Table 1, it is clear the relevant difference of the chemical soil variables along the time. The mobility of P can be detected by the decreasing of CV from 95.66% to 55.12%, while for S the CV increased from 51.48% to 99.23%, and Ca, Na, K, and V had a CV difference between 2019 and 2021 around 15-20%. On the other hand, OM, *m*, and TOC are highly sensible to tillage and fertilization (using single rate application for this case). In Table 1 it is shown that the CV for *m* increased from 63.03% to 99.43%, for OM the CV decreased from 38% to 12.21%, and for TOC

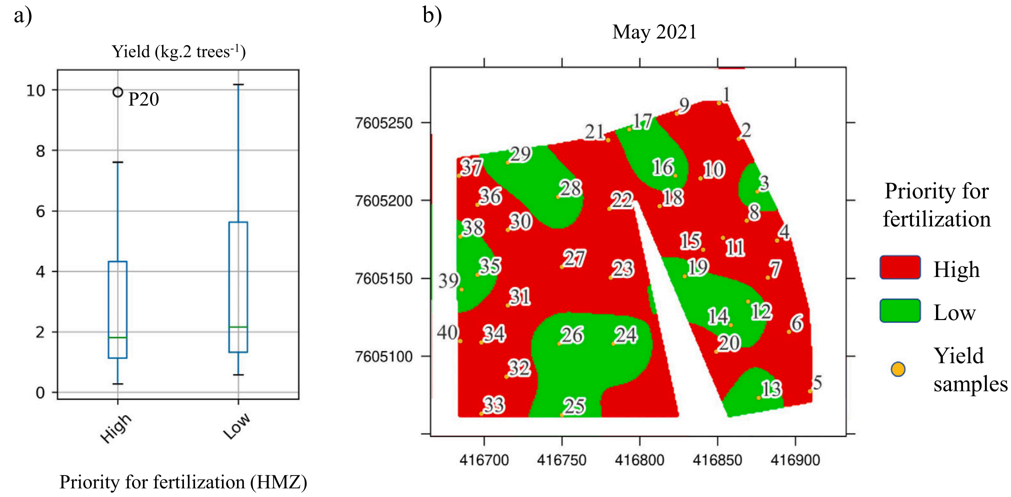

Fig. 7. Validation of HMZs in May 2021. This variable is independent of the HMZ delineation procedure. a) is the boxplot of yield in each HMZ in May 2021 and b) is the HMZ together with the spatial position of yield samples following its original enumeration. The outlier P20 in the plot a) is the 20th soil sample in the plot b).

from 22% to 11.77%. Consequently, the HMZs can change from year to year.

Knowing the mobility of nutrients in the soil is intrinsic to the dynamics of the soil-water-plant interface [72] makes the PA paradigm even more important, as it demonstrates the need for periodic mapping, and consequently the possibility of rethinking the recommendations made and changing strategies, as well as detecting more complex problems that will require further studies.

4.4. Current limitations and future research

Several questions remain to be answered before farmers adopt a site-specific approach for orchard management, such as: Are patterns of yield variation stable enough from year to year, to be predictable? Do yield patterns of variation match homogeneous zones? Are there economic and environmental benefits for targeted management? Several authors have shown the inconsistency over time of spatial variability patterns of important crop and soil properties such as yield, protein content, and plant available N [56,49,22,2]. A common finding is that temporal variability is generally much higher than spatial variability and the definition of stable low- and high-yielding potential zones is very uncertain [14,54,67,59,72]. In this context, a stable configuration of HMZs is highly desirable but not always achievable.

The combined analysis of MN, SSC, and VR results confirms the recommendation of the field partition into two classes. However, other scenarios and combinations can be tested. Gavioli et al. [31] assessed the field partition into 2, 3, and 4 HMZs using a PCA-based method and found the split into 2 zones obtained smoother boundaries with the best multidimensional variance reduction. Córdoba et al. [22] and Peralta and Costa [56] also considered two classes HMZs as the best choice because is a easier field operations choice.

Future research might focus on new experiments with the same sampling design along different periods, looking for more evidence about the robustness of the methodology to ensure the assurance of the farmers in its application. Farmers can focus on managing the variation within coffee complete cycle (around 1 year), and a better strategy might be to combine the use of HMZs with crop-based in-season remote sensing. The latter information could be incorporated efficiently into a decision support system software aimed at supporting farmers in their agricultural management. Also, in future research this methodology will be compared with others, for example with those based on PCA and clustering.

5. Conclusions

In this work, a simple percentage index was proposed for summarizing several soil chemical variables into a single index and then interpolating it by using IK to produce binary HMZ for PA purposes. On more practical grounds, we found that the joint use of SFI and IK was especially useful when the available data were limited and smoothed HMZs were easier to communicate to the farmers.

In light of the experimental results, it is important to underline that the main objective of the paper was to propose a methodology to summarize different soil chemical variables in order to produce a partition of the field into HMZs characterized by similarities in soil chemical properties. The challenge that was faced was to prove that using a simple indicator approach can produce HMZs preserving the main soil chemical characteristics. Also, indirectly it was possible to separate the coffee yield in the different HMZs. From this perspective, the results were quite positive and promising.

The concluding remarks from this study are:

1. SFI was able to summarize several soil chemical variables into a single index;
2. The IK technique was efficient in delineating binary management zones in Oxisols under coffee cultivation by interpolating the soil fertility index;
3. Boxplots of each soil chemical variable showed coherent behaviors into each HMZ, demonstrating that no agronomical information were lost after summarizing data using SFI;
4. Boxplot of coffee yield in each HMZ showed the soil fertility deficit drove a decrease on coffee yield.
5. This methodology can be particularly useful in situations where resources are scarce for large samples, and agronomic knowledge can be used to compensate for the smaller number of samples.

CRediT authorship contribution statement

César de Oliveira Ferreira Silva: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Célia Regina Grego:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation. **Rodrigo Lilla Manzione:** Writing – review & editing, Validation, Supervision. **Stanley Robson De Medeiros Oliveira:** Writing – review & editing, Resources, Project administration. **Gustavo Costa Rodrigues:** Validation, Investigation. **Cristina Aparecida Gonçalves Rodrigues:** Writing – review & editing, Investigation, Funding acquisition. **Eduardo Antonio Speranza:** Writing – review &

editing, Validation. **Ariovaldo Luchiari:** Writing – review & editing. **Luciano Vieira Koenigkan:** Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Celia Regina Grego reports financial support was provided by Brazilian Coffee Research and Development Consortium. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Data availability

Data will be made available on request.

Acknowledgements

A special thanks to Josiane Moraes, the farm manager, and her team, who provided the experimental area and all the necessary support and infrastructure for the execution of the field activities. I would also like to thank the ConCafé Consortium for funding the research project, coordinated by ConCafé Consortium Digital Agriculture, under the code SEG 10 18 20 01200000. This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel – Brazil (in Portuguese: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—CAPES) – Funding Code 001.

References

[1] P.P. Adhikary, C. Dash, R. Bej, H. Chandrasekharan, Indicator and probability kriging methods for delineating Cu, Fe, and Mn contamination in groundwater of Najafgarh Block, Delhi, India, Environ. Monit. Assess. 176 (2011) 663–676, https://doi.org/10.1007/s10661-010-1611-4.

[2] K. Aggelopooulou, A. Castrignanò, T. Gemtos, D. Benedetto, Delineation of management zones in an apple orchard in Greece using a multivariate approach, Comput. Electron. Agric. 90 (2013) 119–130, https://doi.org/10.1016/j.compag.2012.09.009.

[3] L.F. de Almeida, E.E. Spers, Coffee Consumption and Industry Strategies in Brazil: a Volume in the Consumer Science and Strategic Marketing Series, Woodhead Publishing, 2019.

[4] V.V.H. Alvarez, R.F. Novais, N.F. Barros, R.B. Cantarutti, A.S. Lopes, Interpretação dos resultados das análises de solo, in: A.C. Ribeiro, P.T.G. Guimarães, V.V.H. Alvarez (Eds.), Recomendações Para o Uso de Corretivos e Fertilizantes em Minas Gerais, Comissão de Fertilidade do Solo do Estado de Minas Gerais, Viçosa, 1999, pp. 25–32.

[5] M.C. Alves, F.M. Silva, E.A. Pozza, M.S. Oliveira, Modeling spatial variability and pattern of rust and brown eye spot in coffee agroecosystem, J. Pestic. Sci. 82 (2009) 137–148, https://doi.org/10.1007/s10340-008-0232-y.

[6] A.D. Andrade, R. de Oliveira Faria, D.J.C. Alonso, G. Araújo, S. Ferraz, M.A.D. Herrera, F.M. da Silva, Spatial variability of soil penetration resistance in coffee growing, Coffee Sci. 13 (2018) 341–348, https://doi.org/10.25186/cs.v13i3.1456.

[7] G. Araújo, S. Ferraz, M.S. de Oliveira, F.M. da Silva, R.S. Sales, L.C.C. Carvalho, Plant sampling grid determination in precision agriculture in coffee field, Coffee Sci. 13 (2018) 112–121, https://doi.org/10.25186/cs.v13i1.1391.

[8] G. Araújo, S. Ferraz, F.M. da Silva, M.S. de Oliveira, F.C. da Silva, L.C.C. Carvalho, Comparativo entre os atributos químicos do solo amostrados de forma convencional e em malha, Coffee Sci. 12 (2017) 17–29, https://doi.org/10.25186/cs.v12i1.1188.

[9] D. Arrouays, N.P. Saby, J. Thioulouse, C. Jolivet, L. Boulonne, C. Ratié, Large trends in French topsoil characteristics are revealed by spatially constrained multivariate analysis, Geoderma 161 (2011) 107–114, https://doi.org/10.1016/j.geoderma.2010.12.002.

[10] B.S. Bansod, O. Pandey, An application of PCA and fuzzy C-means to delineate management zones and variability analysis of soil, Eurasian Soil Sci. 46 (2013) 556–564, https://doi.org/10.1134/S1064229313050165.

[11] L.S. Barros, E.R.R. Silva, M.N.M. Maciel, V.S.d. Melo, R.E. Cicerelli, T.D. Almeida, Dispersão espacial de atributos químicos do solo de um açaizeiro na região amazônica, Anu. Inst. Geociênc. 45 (2022) 5–25.

[12] P.H.B. Barros, R.A. Oliveira, I.S. Baggio, Coffee productivity and regional development in Brazil, Rev. Polít. Agríc. 28 (2019) 76.

[13] C.L. Bazzi, E.G. Souza, M.A. Uribe-Opazo, L.H. Nóbrega, D.M. Rocha, Definição de unidades de manejo usando atributos químicos e físicos do solo em uma área de soja, Eng. Agric. 33 (2013) 952–964, https://doi.org/10.1590/S0100-69162013000500007.

[14] A.C.C. Bernardi, C.R. Grego, R.G. Andrade, L.M. Rabello, R.Y. Inamasu, Variabilidade espacial de índices de vegetação e propriedades do solo em sistema de integração lavoura-pecuária, Rev. Bras. Eng. Agric. Ambient. 21 (2017) 513–518, https://doi.org/10.1590/1807-1929/agriambi.v21n8p513-518.

[15] H. Burdett, C. Wellen, Statistical and machine learning methods for crop yield prediction in the context of precision agriculture, Precis. Agric. (2022) 1–22, https://doi.org/10.1007/s11119-022-09897-0.

[16] G. Buttafuoco, A. Castrignanò, A.S. Colecchia, N. Ricca, Delineation of management zones using soil properties and a multivariate geostatistical approach, Ital. J. Agron. 5 (2010) 323–332, https://doi.org/10.4081/ija.2010.323.

[17] C. Cambardella, E. Elliott, Carbon and nitrogen dynamics of soil organic matter fractions from cultivated grassland soils, Soil Sci. Soc. Am. J. 58 (1994) 123–130, https://doi.org/10.2136/sssaj1994.03615995005800050033x.

[18] A. Castrignanò, G. Buttafuoco, R. Quarto, D. Parisi, R.V. Rossel, F. Terribile, G. Langella, A. Venezia, A geostatistical sensor data fusion approach for delineating homogeneous management zones in precision agriculture, Catena 167 (2018) 293–304, https://doi.org/10.1016/j.catena.2018.05.011.

[19] A. Castrignanò, E.A. Costantini, R. Barbetti, D. Sollitto, Accounting for extensive topographic and pedologic secondary information to improve soil mapping, Catena 77 (2009) 28–38, https://doi.org/10.1016/j.catena.2008.12.004.

[20] A. Castrignanò, R. Quarto, A. Venezia, G. Buttafuoco, A comparison between mixed support kriging and block cokriging for modelling and combining spatial data with different support, Precis. Agric. 20 (2019) 193–213, https://doi.org/10.1007/s11119-018-09630.

[21] A.F. Colaço, R.G. Bramley, Do crop sensors promote improved nitrogen management in grain crops?, Field Crops Res. 218 (2018) 126–140, https://doi.org/10.1016/j.fcr.2018.01.007.

[22] M.A. Córdoba, C.I. Bruno, J.L. Costa, N.R. Peralta, M.G. Balzarini, Protocol for multivariate homogeneous zone delineation in precision agriculture, Biosyst. Eng. 143 (2016) 95–107, https://doi.org/10.1016/j.biosystemseng.2015.12.008.

[23] N. Davatgar, M. Neishabouri, A. Sepaskhah, Delineation of site specific nutrient management zones for a paddy cultivated area based on soil fertility using fuzzy clustering, Geoderma 173 (2012) 111–118, https://doi.org/10.1016/j.geoderma.2011.12.005.

[24] M. Dowle, A. Srinivasan, data.table: extension of 'data.frame', R package version 1.14.2. URL https://CRAN.R-project.org/package=data.table, 2022.

[25] M. Emadi, A.R. Shahriari, F. Sadegh-Zadeh, B. Jalili Seh-Bardan, A. Dindarlou, Geostatistics-based spatial distribution of soil moisture and temperature regime classes in Mazandaran province, northern Iran, Arch. Agron. Soil Sci. 62 (2016) 502–522, https://doi.org/10.1080/03650340.2015.1065607.

[26] G.A. Ferraz, F. Da Silva, M. De Oliveira, A.A.P. Custódio, P.F.P. Ferraz, Spatial variability of plant attributes in a coffee plantation, Rev. Cienc. Agron. 48 (2017) 81–91, https://doi.org/10.5935/1806-6690.20170009.

[27] G.A. Ferraz, F.M.d. Silva, M.S.d. Oliveira, F.C.d. Silva, R.d.L. Bueno, et al., Variabilidade espacial da força de desprendimento de frutos do cafeeiro, Eng. Agric. 34 (2014) 1210–1223, https://doi.org/10.1590/S0100-69162014000600016.

[28] G.A.S. Ferraz, F.M. Silva, L.C. Carvalho, M.C. Alves, B.C. Franco, Variabilidade espacial e temporal do fósforo, potássio e da produtividade de uma lavoura cafeeira, Eng. Agric. 32 (2012) 140–150, https://doi.org/10.1590/S0100-69162012000100015.

[29] K. Fleming, D. Westfall, D. Wiens, M. Brodahl, Evaluating farmer defined management zone maps for variable rate fertilizer application, Precis. Agric. 2 (2000) 201–215, https://doi.org/10.1023/A:1011481832064.

[30] W. Fu, H. Tunney, C. Zhang, Spatial variation of soil nutrients in a dairy farm and its implications for site-specific fertilizer application, Soil Tillage Res. 106 (2010) 185–193, https://doi.org/10.1016/j.still.2009.12.001.

[31] A. Gavioli, E.G. de Souza, C.L. Bazzi, L.P.C. Guedes, K. Schenatto, Optimization of management zone delineation by using spatial principal components, Comput. Electron. Agric. 127 (2016) 302–310, https://doi.org/10.1016/j.compag.2016.06.029.

[32] A. Gavioli, E.G. de Souza, C.L. Bazzi, K. Schenatto, N.M. Betzek, Identification of management zones in precision agriculture: an evaluation of alternative cluster analysis methods, Biosyst. Eng. 181 (2019) 86–102, https://doi.org/10.1016/j.biosystemseng.2019.02.019.

[33] A. Gili, C. Álvarez, R. Bagnato, E. Noellemeyer, Comparison of three methods for delineating management zones for site-specific crop management, Comput. Electron. Agric. 139 (2017) 213–223, https://doi.org/10.1016/j.compag.2017.05.022.

[34] P.A. Gómez-Ruiz, R. Lindig-Cisneros, E. de la Barrera, C. Martorell, Potassium enhances frost tolerance in young individuals of three tropical dry forest species from Mexico, Funct. Plant Biol. 43 (2016) 461–467, https://doi.org/10.1071/FP15329.

[35] P. Goovaerts, Geostatistics for Natural Resources Evaluation, Oxford University Press, New York, 1997.

[36] F. Guastaferro, A. Castrignanò, D. De Benedetto, D. Sollitto, A. Troccoli, B. Cafarelli, A comparison of different algorithms for the delineation of management zones, Precis. Agric. 11 (2010) 600–620, https://doi.org/10.1007/s11119-010-9183-4.

[37] P.T.G. Guimarães, A.W.R. Garcia, V.H. Alvarez, L.C. Prezotti, A.S. Viana, A.E. Miguel, E. Malavolta, J.B. Corrêa, A.S. Lopes, F.D. Nogueira, A.V.C. Monteiro, J.A. Oliveira, Cafeeiro, in: A.C. Ribeiro, P.T.G. Guimarães, V.V.H. Alvarez (Eds.), Recomendações Para o Uso de Corretivos e Fertilizantes em Minas Gerais, Comissão de Fertilidade do Solo do Estado de Minas Gerais, Viçosa, 1999, pp. 289–302.

[38] E.H. Isaaks, M.R. Srivastava, Applied Geostatistics, 551.72 ISA, Oxford University Press, New York, 1989.

[39] ISPAG, Precision Ag definition, https://www.ispag.org/about/definition/, 2019. (Accessed 21 April 2022).

[40] J.L. Jacintho, G.A. Ferraz, F.M. Silva, S.A. Santos, Definição de zonas de manejo para cafeicultura, Rev. Bras. Eng. Agric. Ambient. 21 (2017) 94–99, https://doi.org/10.1590/1807-1929/agriambi.v21n2p94-99.

[41] H.L. Jiang, G.S. Liu, S.D. Liu, E.H. Li, R. Wang, Y.F. Yang, H.C. Hu, Delineation of site-specific management zones based on soil properties for a hillside field in central China, Arch. Agron. Soil Sci. 58 (2012) 1075–1090, https://doi.org/10.1080/03650340.2011.570337.

[42] A. Kolmogorov, Sulla determinazione empirica di una lgge di distribuzione, Inst. Ital. Attuari, Giorn. 4 (1933) 83–91.

[43] D. Lambert, J. Lowenberg-DeBoer, G. Malzer, Understanding phosphorous in Minnesota soils, Agric. Econ. 37 (2007) 43–53, https://doi.org/10.1111/j.1574-0862.2007.00221.x.

[44] J.S.d.S. Lima, S.d.A. Silva, R.B.d. Oliveira, A.S.d. Fonseca, Estimativa da produtividade de café conilon utilizando técnicas de cokrigagem, Ceres 63 (2016) 54–61, https://doi.org/10.1590/0034-737X201663010008.

[45] C. Lloyd, P.M. Atkinson, Assessing uncertainty in estimates with ordinary and indicator kriging, Comput. Geosci. 27 (2001) 929–937, https://doi.org/10.1016/S0098-3004(00)00132-1.

[46] J.A. Martínez-Casasnovas, A. Escolà, J. Arnó, Use of farmer knowledge in the delineation of potential management zones in precision agriculture: a case study in maize (Zea mays L.), Agriculture 8 (2018) 84, https://doi.org/10.3390/agriculture8060084.

[47] G. Matheron, La théorie des variables régionalisées et ses applications, Cahors Cent. Morphol. Math. 5 (1970) 1–212.

[48] M.S. Metwally, S.M. Shaddad, M. Liu, R.J. Yao, A.I. Abdo, P. Li, J. Jiao, X. Chen, Soil properties spatial variability and delineation of site-specific management zones based on soil fertility using fuzzy clustering in a hilly field in Jianyang, Sichuan, China, Sustainability 11 (2019) 7084, https://doi.org/10.3390/su11247084.

[49] F. Morari, A. Castrignanò, C. Pagliarin, Application of multivariate geostatistics in delineating management zones within a gravelly vineyard using geo-electrical sensors, Comput. Electron. Agric. 68 (2009) 97–107, https://doi.org/10.1016/j.compag.2009.05.003.

[50] P. Núñez, A. Pimentel, I. Almonte, D. Sotomayor-Ramírez, N. Martínez, A. Pérez, C. Céspedes, Soil fertility evaluation of coffee (coffea spp.) production systems and management recommendations for the Barahona Province, Dominican Republic, J. Soil Sci. Plant Nutr. 11 (2011) 127–140.

[51] OpenDroneMap/ODM, Opendronemap/odm github repository, https://github.com/OpenDroneMap/ODM/, 2020. (Accessed 21 April 2022).

[52] R. Ortega, O. Santibáñez, Agronomic evaluation of three zoning methods based on soil fertility in corn crops (Zea maize L.), Comput. Electron. Agric. 58 (2007) 49–59, https://doi.org/10.1016/j.compag.2006.12.011.

[53] R.A. Ortega, O.A. Santibanez, Determination of management zones in corn (Zea mays L.) based on soil fertility, Comput. Electron. Agric. 58 (2007) 49–59, https://doi.org/10.1016/j.compag.2006.12.011.

[54] J.O. Paris, I. Gontijo, F.L. Partelli, A.G. Facco, Variability and spatial correlation of soil micronutrients and organic matter with macadamia nut production, Rev. Bras. Eng. Agric. Ambient. 24 (2019) 31–36, https://doi.org/10.1590/1807-1929/agriambi.v24n1p31-36.

[55] E. Pebesma, B. Graeler, gstat: spatial and spatio-temporal geostatistical modelling, prediction and simulation, R package version 2.0-9 https://CRAN.R-project.org/package=gstat, 2022.

[56] N.R. Peralta, J.L. Costa, Delineation of management zones with soil apparent electrical conductivity to improve nutrient management, Comput. Electron. Agric. 99 (2013) 218–226, https://doi.org/10.1016/j.compag.2013.09.014.

[57] F.J. Pierce, P. Nowak, Aspects of precision agriculture, Adv. Agron. 67 (1999) 1–85, https://doi.org/10.1016/S0065-2113(08)60513-1.

[58] R Development Core Team, R: a Language and Environment for Statistical Computing, 2022.

[59] P.C. Robert, Precision agriculture: a challenge for crop nutrition management, in: Progress in Plant Nutrition: Plenary Lectures of the XIV International Plant Nutrition Colloquium, 2002, pp. 143–149.

[60] C.C. Ronquim, Conceitos de fertilidade do solo e manejo adequado para as regiões tropicais, in: Embrapa Territorial-Boletim de Pesquisa e Desenvolvimento (INFOTECA-E), 2010.

[61] E.O.J. Santos, I. Gontijo, M.B. Silva, Spatial variability of soil acidity attributes and liming requirement for conilon coffee, Coffee Sci. 9 (2014) 275–283.

[62] H.G. Santos, P.K.T. Jacomine, L.H.C. Anjos, V.A. Oliveira, M.R. Coelho, J.R. Lumbrelas, Sistema Brasileiro de Classificação de Solos, Centro Nacional de Pesquisa de Solos, Rio de Janeiro, 2006.

[63] SCA, What is specialty coffee?, https://www.ispag.org/about/definition/. (Accessed 6 November 2022), 2021.

[64] E. Scudiero, P. Teatini, D.L. Corwin, R. Deiana, A. Berti, F. Morari, Delineation of site-specific management units in a saline region at the Venice Lagoon margin, Italy, using soil reflectance and apparent electrical conductivity, Comput. Electron. Agric. 99 (2013) 54–64, https://doi.org/10.1016/j.compag.2013.08.023.

[65] S.M. Shaddad, G. Buttafuoco, A. Castrignanò, Assessment and mapping of soil salinization risk in an Egyptian field using a probabilistic approach, Agronomy 10 (2020) 85, https://doi.org/10.3390/agronomy10010085.

[66] C.d.O.F. Silva, C.R. Grego, R.L. Manzione, S.R.d.M. Oliveira, Improving coffee yield interpolation in the presence of outliers using multivariate geostatistics and satellite data, AgriEngineering 6 (2024) 81–94, https://doi.org/10.3390/agriengineering6010006.

[67] C.d.O.F. Silva, R.L. Manzione, S.R.d.M. Oliveira, Exploring 20-year applications of geostatistics in precision agriculture in Brazil: what's next?, Precis. Agric. (2023) 1–34, https://doi.org/10.1007/s11119-023-10041-9.

[68] F.M. Silva, M.C. Alves, J.C.S. Souza, M.S.d. Oliveira, Effects of manual harvesting on coffee (Coffea arabica L.) crop biannuality in Ijaci, Minas Gerais, Cienc. Agrotecnol. 34 (2010) 625–632, https://doi.org/10.1590/S1413-70542010000300014.

[69] S.A. Silva, J.S.S. Lima, Multivariate analysis and geostatistics of the fertility of a Humic Rhodic Hapludox under coffee cultivation, Rev. Bras. Ciênc. Solo 36 (2012) 467–474.

[70] S.A. Silva, J.S.S. Lima, Relação espacial entre o estoque de nutrientes e a densidade de solo cultivado com cafeeiro, Pesqui. Agropecu. Trop. 43 (2013) 377–384, https://doi.org/10.1590/S1983-40632013000400002.

[71] R.L. Snyder, J.d. Melo-Abreu, Frost Protection: Fundamentals, Practice and Economics, volume 1, 2005.

[72] G.W. Thomas, Soil and climatic factors which affect nutrient mobility, Nutr. Mobil. Soils Accumul. Losses 4 (1970) 1–20.

[73] A. Utset, M.E. Ruiz, J. Herrera, D.P. de Leon, A geostatistical method for soil salinity sample site spacing, Geoderma 86 (1998) 143–151, https://doi.org/10.1016/S0016-7061(98)00037-8.

[74] H.D. Vieira, Café Rural, Interciência, Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, Rio de Janeiro, 2017.

[75] H. Wickham, tidyverse: easily install and load the 'tidyverse', R package version 1.3.1. URL https://CRAN.R-project.org/package=tidyverse, 2021.

[76] H. Wickham, ggplot2: create elegant data visualisations using the grammar of graphics, R package version 3.3.6. URL https://CRAN.R-project.org/package=ggplot2, 2022.

[77] H. Wickham, J. Hester, W. Chang, J. Bryan, devtools: tools to make developing R packages easier, R package version 2.4.3. URL https://CRAN.R-project.org/package=devtools, 2022.

[78] W. Xin-Zhong, L. Guo-Shun, H. Hong-Chao, W. Zhen-Hai, L. Qing-Hua, L. Xu-Feng, H. Wei-Hong, L. Yan-Tao, Determination of management zones for a tobacco field based on soil fertility, Comput. Electron. Agric. 65 (2009) 168–175, https://doi.org/10.1016/j.compag.2008.08.008.