Revista Brasileira de Ciência do Solo

Division - Soil In Space and Time | Commission - Pedometrics

Proposal and validation of geostatistical-based metrics to quantify within-field variability

Enio Júnior Seidel[(1)*] and **Marcelo Silva de Oliveira**[(2)]

[(1)] Universidade Federal do Amapá, Departamento de Ciências Exatas e Tecnológicas, Macapá, Amapá, Brasil.
[(2)] Universidade Federal de Lavras, Departamento de Estatística, Lavras, Minas Gerais, Brasil.

ABSTRACT: Metrics are fundamental to quantify and classify the spatial dependence of soil and agricultural attributes. This study aimed to propose and validate metrics based on two distinct approaches, one additive, which considers the arithmetic mean of the vertical and horizontal components, and the other multiplicative, which considers the geometric mean of the vertical and horizontal components of the semivariogram. Furthermore, we intend to propose the classification of spatial dependence based on the categorization of these metrics. Finally, a function in R language is presented to calculate the metrics and classify spatial dependence. The spatial dependence arithmetic index 1 (SDAI1) and spatial dependence arithmetic index 2 (SDAI2) are constructed in a dimensionless way, in the range between 0 and 100 %, considering the sum (arithmetic mean) between the vertical and horizontal components of the semivariogram. The spatial dependence geometric index 1 (SDGI1) and the spatial dependence geometric index 2 (SDGI2) are constructed in a dimensionless way, in the range between 0 and 100 %, considering the multiplication (geometric mean) between the vertical and horizontal components of the semivariogram. The SDAI1, SDAI2, SDGI1, and SDGI2 metrics are compared with other metrics existing in the literature, such as the spatial dependence degree (SPD), the integral scales J1 and J2, the mean correlation distance (MCD), the spatial dependence index (SDI), and the spatial dependence measure (SDM). For different spatial dependence scenarios, correlations are calculated between the geostatistical-based metrics and the performance measures Moran's I, mean squared error (MSE), and kriging variance (KV). The metrics perform well in describing spatial dependence, with the exception of J1 (or MCD) and J2. However, the SDAI1, SDAI2, and SDGI1 metrics have slightly better correlations with the Moran's I, MSE, and KV measures, when compared to the SDI, SDM, SDGI2, and SPD metrics. Furthermore, the SDAI1 and SDAI2 metrics show superior performance in capturing the vertical and horizontal effects of the semivariogram. Finally, a function in R language was developed to calculate the metrics and classify spatial dependence.

Keywords: precision agriculture, spatial variability, autocorrelation, semivariograms.

*** Corresponding author:**
E-mail: enioseidel@gmail.com

Received: February 19, 2025
Approved: June 05, 2025

How to cite: Seidel EJ, Oliveira MS. Proposal and validation of geostatistical-based metrics to quantify within-field variability. Rev Bras Cienc Solo. 2025;49:e0250042.
https://doi.org/10.36783/18069657rbcs20250042

Editors: José Miguel Reichert and Quirijn de Jong van Lier.

INTRODUCTION

Geostatistics is widely applied in agricultural sciences to predict and map soil and agricultural attributes, and is essential in the context of Agriculture 4.0 (Rodrigues et al., 2020). At the heart of the use of Geostatistics is spatial dependence, which conceptually is characterized by the way in which geospatial locations correlate.

Quantification of spatial dependence is generally performed using metrics derived from semivariogram parameters (Padilha et al., 2024) or by measures of spatial autocorrelation, such as Moran's I (Moran, 1950) or Geary's C (Geary, 1954). Among the metrics derived from the semivariogram, the following stand out: the relative nugget effect (RNE), presented in Cambardella et al. (1994), the integral scales J1 and J2, presented in Russo and Jury (1987), the mean correlation distance (MCD), given in Han et al. (1994), the spatial dependence degree (SPD), presented in Biondi et al. (1994), the spatial dependence index (SDI), proposed by Seidel and Oliveira (2014), and the spatial dependence measure (SDM), proposed by Appel Neto et al. (2020). Spatial dependence metrics play a prominent role for two main purposes: (I) they allow the early assessment of the quality of kriging predictions, since the stronger the spatial dependence, the more accurate the kriging prediction; (II) they allow the comparison of the spatial variability observed in different attributes in different fields.

Among the semivariogram parameters, the range parameter (*a*) describes the spatial variability in the horizontal direction of the graph, while the nugget effect (C_0), partial sill (C_1), and sill (*C*) describe the spatial variability in the vertical direction of the semivariogram (Santos et al., 2018). Seidel and Oliveira (2012) presented a measure called the combined spatial dependence index (CSDI), composed of a component that captures the variability in the vertical direction of the semivariogram (VSDI) and a component that captures the variability in the horizontal direction of the semivariogram (HSDI). The VSDI considers the relationship between the C_1 and the *C* in the same way as the index of Biondi et al. (1994). The HSDI considers the relationship between the *a* and half of the greatest distance between sampled points (0.5 *MD*). The combination of vertical and horizontal parameters of the semivariogram is also present in a multiplicative form in the J1 and J2 (Russo and Jury, 1987), MCD (Han et al., 1994), SDI (Seidel and Oliveira, 2014; Appel Neto et al., 2018; Uribe-Opazo et al., 2023), SDM (Appel Neto et al., 2020), and adapted RNE (Zhao et al., 2023) measures.

Assessing spatial variability within the field is a concern for many researchers (Cambardella et al., 1994; Taylor et al., 2007; Herrero-Langreo et al., 2019; Leroux and Tisseyre, 2019; Sandonís-Pozo et al., 2022; Jang et al., 2023; Zhao et al., 2023), as there may be heterogeneities in field factors that directly influence agricultural productivity (Leroux and Tisseyre, 2019) and production profitability (Zhao et al., 2023). Thus, the construction and comparison of different metrics for quantifying spatial dependence, derived from the semivariogram, are important.

This study aimed to propose and validate metrics based on two distinct approaches: an additive one, which considers the arithmetic mean of the vertical and horizontal components $\left(metric = \frac{vertical\ component\ + horizontal\ component}{2}\right)$, and a multiplicative one, which considers the geometric mean of the vertical and horizontal components of the semivariogram $\left(metric = \sqrt[2]{vertical\ component\ \times horizontal\ component}\right)$. Furthermore, we intend to propose the classification of spatial dependence based on the categorization of such metrics. Finally, a function in R language is presented to calculate the metrics and classify spatial dependence.

MATERIALS AND METHODS

Vertical components describe the spatial variability in the vertical direction of the semivariogram, being given in the interval from zero are given by equations 1 and 2, respectively.

$$VC_1 = \left(\frac{C_1}{C_0 + C_1}\right) \quad \text{Eq. 1}$$

and

$$VC_2 = \sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)} \quad \text{Eq. 2}$$

in which: C_0 is the nugget effect and C_0 is the partial sill.

The horizontal component *HC* is given by equation 3.

$$HC = \min\left\{1; \frac{a}{0.5\ MD}\right\} \quad \text{Eq. 3}$$

in which: *a* is the range; *MD* is the maximum distance between points on the sampling grid; and *min*{*} is the minimum function.

The HC describes the spatial variability in the horizontal direction of the semivariogram, being given in the interval from zero to one, that is, applying the *min*{*} function ensures that this component assumes values only in the interval [0, 1]. The 0.5 *MD* factor is based on practical recommendations for using pairs of locations up to half of the largest sampling distance to estimate semivariances (Journel and Huijbregts, 2003; Olea, 2006). The construction of the metrics is based on the simple arithmetic mean (equation 4) and the simple geometric mean (equation 5).

$$\overline{X}_a = \frac{1\ VC_i\ +\ 1\ HC}{2} \quad \text{Eq. 4}$$

$$\overline{X}_g = \sqrt[2]{VC_i^{\ 1}\ HC^1} \quad \text{Eq. 5}$$

in which: VC_i is the i-th vertical component, i=1,2; and *HC* is the horizontal component.

Spatial dependence arithmetic index 1 (SDAI1) and the spatial dependence arithmetic index 2 (SDAI2) are constructed in a dimensionless way, in the interval between 0 and 100 %, considering the arithmetic mean between the vertical and horizontal components of the semivariogram.

The SDAI1 is given by equation 6.

$$SDAI1\ (\%) = \frac{\left(\frac{C_1}{C_0+C_1}\right) + \left(\min\left\{1; \frac{a}{0.5\ MD}\right\}\right)}{2} \times 100 \quad \text{Eq. 6}$$

in which: C_0 is the nugget effect; C_1 is the partial sill; a is the range; MD is the maximum distance between points on the sampling grid; and $min\{*\}$ is the minimum function. $0 \leq \left(\frac{C_1}{C_0 + C_1}\right) \leq 1$. $0 \leq \min\left\{1; \frac{a}{0.5\ MD}\right\} \leq 1$. The SDAI1 was originally presented as CSDI by Seidel and Oliveira (2012).

The SDAI2 is given by equation 7.

$$SDAI2\ (\%) = \frac{\sqrt[2]{\left(\frac{C_1}{C_0+C_1}\right)} + \left(\min\left\{1; \frac{a}{0.5\ MD}\right\}\right)}{2} \times 100 \qquad \text{Eq. 7}$$

in which: C_0 is the nugget effect; C_1 is the partial sill; a is the range; and MD is the maximum distance between points on the sampling grid; and $min\{*\}$ is the minimum function: $0 \leq \sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)} \leq 1$; $0 \leq \min\left\{1; \frac{a}{0.5 \times MD}\right\} \leq 1$.

Spatial dependence geometric index 1 (SDGI1) and the spatial dependence geometric index 2 (SDGI2) are constructed in a dimensionless way, in the interval between 0 and 100 %, considering the geometric mean between the vertical and horizontal components of the semivariogram.

The SDGI1 is given by equation 8.

$$SDGI1\ (\%) = \sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)\ \left(\min\left\{1; \frac{a}{0.5\ MD}\right\}\right)} \times 100 \qquad \text{Eq. 8}$$

in which: C_0 is the nugget effect; C_1 is the partial sill; a is the range; MD is the maximum distance between points on the sampling grid; and $min\{*\}$ is the minimum function: $0 \leq \left(\frac{C_1}{C_0 + C_1}\right) \leq 1$; $0 \leq \min\left\{1; \frac{a}{0.5\ MD}\right\} \leq 1$.

The SDGI2 is given by equation 9.

$$SDGI2\ (\%) = \sqrt[2]{\sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)}\ \left(\min\left\{1; \frac{a}{0.5\ MD}\right\}\right)} \times 100 \qquad \text{Eq. 9}$$

in which: C_0 is the nugget effect; C_1 is the partial sill; a is the range; MD is the maximum distance between points on the sampling grid; and min{*} is the minimum function. $0 \leq \sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)} \leq 1$. $0 \leq \min\left\{1; \frac{a}{0.5\ MD}\right\} \leq 1$.

To carry out the classification of the SDAI1, SDAI2, SDGI1, and SDGI2, a methodology adapted from Seidel and Oliveira (2016), Appel Neto et al. (2018), and Appel Neto et al. (2020) was used. Values were generated for the components $\left(\frac{C_1}{C_0 + C_1}\right)$ or $\sqrt[2]{\left(\frac{C_1}{C_0 + C_1}\right)}$ and $\min\left\{1; \frac{a}{0.5\ MD}\right\}$, from 0.05 to 1.00, varying by 0.05. Subsequently, these values were combined by arithmetic means for SDAI1 and SDAI2, and by geometric means for SDGI1 and SDGI2, which generated a vector of 400 values that was increased by a zero value. The first and the third quartiles of the vector were considered as cut-offs to categorize the metrics, and classify the spatial dependence as weak, moderate, and strong.

The SDAI2 and SDGI2 metrics are constructed to give greater value to the vertical component of the semivariogram because it is given by the square root of values in the interval [0, 1], in the same way as performed in J2 (Russo and Jury, 1987) and in SDM (Appel Neto et al., 2020). Equation 10 shows the result for $0 \leq \left(\frac{C_1}{C_0 + C_1}\right) \leq 1$.

$$\begin{cases} \sqrt[2]{\left(\frac{C_1}{C_0+C_1}\right)} = \left(\frac{C_1}{C_0+C_1}\right), & if \left(\frac{C_1}{C_0+C_1}\right) = 0 \ or \ \left(\frac{C_1}{C_0+C_1}\right) = 1, \\ \sqrt[2]{\left(\frac{C_1}{C_0+C_1}\right)} > \left(\frac{C_1}{C_0+C_1}\right), & if \ 0 < \left(\frac{C_1}{C_0+C_1}\right) < 1. \end{cases} \quad \text{Eq. 10}$$

in which: C_0 is the nugget effect; and C_1 is the partial sill.

In addition to proposing the measures SDAI1, SDAI2, SDGI1, and SDGI2 and their respective classifications of spatial dependence, the SDI, SDM, SPD, and MCD, and the integral scales J1 and J2 are also evaluated. It follows that MCD and J1 are identical, that is, MCD is the closed form of the integral scale J1. Furthermore, it is decided to use the SPD and not the RNE, because *SPD*(%) = 100 - *RNE*(%).

To evaluate the performance of the metrics, the associations with their components, and compare them with some of the indices used in the literature, scenarios of spatial variability (considering weak to strong spatial dependence) were simulated in the geoR package (Ribeiro Junior et al., 2020), from R software (R Development Core Team, 2021) in the same way as in Appel Neto et al. (2020). The scenarios were composed with the following parameters: C_1 values of 10, 25, 50, 75, and 90 % of C = 50; a values of 10, 25, 50, 75, and 90 % of 0.5 MD = 70.71 m, in a random field with mean equal to zero, with n = 169, in a 100 × 100 m regular sampling grid, totaling 25 scenarios for each of the exponential, Gaussian and spherical models. Each scenario was replicated 100 times to mitigate possible variations in the simulation algorithm. With the 100 values generated from each scenario, the means of the metrics and, as performance measures, the means of the Moran's I, mean squared error (MSE), and kriging variance (KV) generated by cross-validation were calculated. Finally, Pearson correlation between the metrics and performance measures was calculated.

In addition, a function in R is presented to calculate the metrics and classify spatial dependence. Finally, to exemplify the application of the metrics in the proposal for data classification, the *soilmoisture* and the *NVDI* data sets from the geotoolsR package (Rossoni and Felix, 2020) were used.

RESULTS AND DISCUSSION

The simulations of the values of SDAI1, SDAI2, SDGI1, and SDGI2 show statistically symmetric behaviors (Figure 1), with skewness coefficients of -0.02 (p = 0.8510), -0.17 (p = 0.1688), 0.23 (p = 0.0628), and -0.08 (p = 0.4990), respectively. The fact that there is symmetry in the metric distributions allows the classification of spatial dependence based on the first and third quartiles, since symmetry gives both sides of the distribution the same frequency (probabilistic) weight. The proposed classification of spatial dependence for the metrics SDAI1, SDAI2, SDGI1, and SDGI2 is presented in table 1. For categorization, the quartile values were rounded to integer values. Cambardella et al. (1994) also used quartiles (1st quartile = 25 %; 3rd quartile = 75 %) in the RNE categorization cut-offs to generate three classes (weak; moderate; strong) of spatial dependence. Similarly, Barbosa et al. (2017) used quartiles (1st quartile = 25 %; 3rd quartile = 75 %) in the categorization cut-offs in the proposal of a measure equivalent to the SPD, in the power semivariogram model, to generate three classes (weak; moderate; strong) of spatial dependence.

Table 2 presents the correlations between the spatial dependence measures (SDI, SDM, SDAI1, SDAI2, SDGI1, and SDGI2) and the performance measures (Moran's I, MSE, and KV). We see that the SDI and SDM measures present strong negative correlations with MSE and KV in the Gaussian and spherical models. In the exponential model, SDI and SDM are strongly negatively correlated with KV and moderately negatively correlated with MSE. Furthermore, SDI and SDM are strongly positively correlated with Moran's I in the Gaussian and spherical models, and moderately positively correlated in the exponential model.

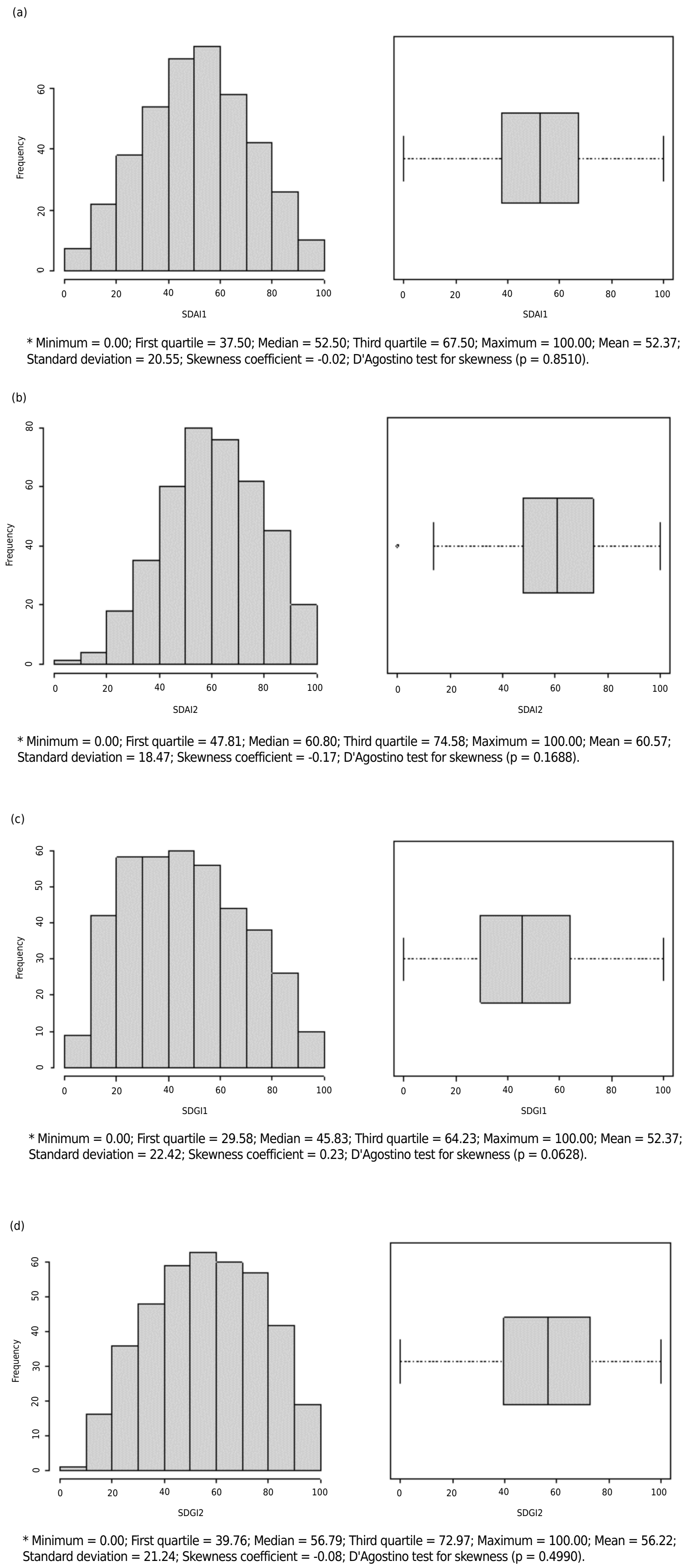

Figure 1. Simulated behavior (Histogram and Boxplot) of Spatial dependence arithmetic index 1 (SDAI1) (a), Spatial dependence arithmetic index 2 (SDAI2) (b), Spatial dependence geometric index 1 (SDGI1) (c), and Spatial dependence geometric index 2 (SDGI2) (d).

Table 1. Proposed classification of spatial dependence for the metrics Spatial dependence arithmetic index 1 (SDAI1), Spatial dependence arithmetic index 2 (SDAI2), Spatial dependence geometric index 1 (SDGI1), and Spatial dependence geometric index 2 (SDGI2)

Metric	Classification		
	Weak	Moderate	Strong
SDAI1	0 to 38	38 to 68	68 to 100
SDAI2	0 to 48	48 to 75	75 to 100
SDGI1	0 to 30	30 to 64	64 to 100
SDGI2	0 to 40	40 to 73	73 to 100

Seidel and Oliveira (2014) observed an excellent relationship between the SDI and MSE in the spherical and Gaussian models, and a good relationship between the SPD and the MSE for the exponential model. Appel Neto et al. (2020) observed strong negative correlations between the SDI and SDM indices and the MSE and KV performance measures for the spherical, exponential, Gaussian, and cubic models, and strong negative correlations between the SPD and the MSE and KV performance measures in the exponential, Gaussian, and wave models.

The SDAI1 and SDAI2 measures present strong negative correlations with MSE and KV and strong positive correlations with Moran's I in all three models. The SDGI1 measure has strong negative correlations with MSE and KV, and a strong positive correlation with Moran's I in all three models. The SDGI2 measure shows a strong negative correlation with the MSE in the Gaussian and spherical models, and a moderate negative correlation in the exponential model. In addition, the SDGI2 measure has a strong negative correlation with the KV and a strong positive correlation with Moran's I in all three models.

The J1 (or MCD) and J2 measures show weak positive correlations with MSE and KV and weak negative correlations with Moran's I in all three models. The SPD shows strong negative correlations with MSE and KV for the Gaussian model, and moderate negative correlations in the exponential and spherical models. Furthermore, the SPD presents a moderate positive correlation with Moran's I in all three models. Fu et al. (2011) observed a strong correlation (R^2 = 0.8578) between the RNE and Moran's I (or strong correlation between SPD and Moran's I).

In theory, it is expected that higher values of spatial dependence should be related to lower values of MSE and KV and higher values of Moran's I. Thus, the measures SDAI1, SDAI2, and SDGI1 tend to present slightly better results when compared to SDI and SDM, SDGI2, SPD, J1 (or MCD), and J2, in this aspect. These relevant metrics correlations for models corroborate the perception that metrics that summarize the entire spatial dependency structure in a single number are possible.

Spatial dependence metrics are associated with cross-validation performance measures, and according to Appel Neto et al. (2018), spatial dependence metrics can indicate the quality of kriging prediction maps. However, Amaral and Della Justina (2019), in a study with soil samples from sugarcane fields, observed that the SPD, SDI, and Moran's I metrics presented limitations when trying to assess the accuracy of predictions by interpolation. Even considering these limitations, the search for an metric that allows the classification of the spatial variability structure present in a phenomenon is of notorious importance, since only with the use of a single number (the calculated value of the metric) is it possible to compare and hierarchize spatial phenomena, and, therefore, to recommend the appropriate management of different phenomena, especially when a trade-off decision regarding the application of resources needs to be made.

The component $\min\left\{1; \frac{a}{0.5\,MD}\right\}$ presents strong positive correlations with SDI, SDM, SDAI1, SDAI2, SDGI1, and SDGI2, and a weak correlation with SPD in the three models.

Table 2. Pearson correlation between spatial dependence index (SDI), spatial dependence measure (SDM), spatial dependence arithmetic index 1 (SDAI1), spatial dependence arithmetic index 2 (SDAI2), spatial dependence geometric index 1 (SDGI1), spatial dependence geometric index 2 (SDGI2), integral scales J1 and J2, mean correlation distance (MCD), spatial dependence degree (SPD), component $\left(\frac{C_1}{C_0+C_1}\right)$ ($), component $\sqrt[2]{\left(\frac{C_1}{C_0+C_1}\right)}$ (£), component $min\left\{1; \frac{a}{0.5 \times MD}\right\}$ (#), and Moran's I, Mean square error (MSE), and Kriging variance (KV). Simulated data

X	SDI	SDM	SDAI1	SDAI2	SDGI1	SDGI2	J1 (or MCD)	J2	SPD
Exponential semivariogram model									
SDM	0.985								
SDAI1	0.906	0.923							
SDAI2	0.909	0.942	0.992						
SDGI1	0.966	0.987	0.963	0.976					
SDGI2	0.945	0.984	0.943	0.969	0.991				
J1 (or MCD)	0.257	0.217	0.004	0.015	0.095	0.126			
J2	0.257	0.217	0.003	0.015	0.095	0.126	0.999		
SPD	0.332	0.288	0.603	0.525	0.407	0.318	-0.444	-0.444	
$	0.332	0.288	0.603	0.525	0.407	0.318	-0.444	-0.444	1.000
£	0.261	0.239	0.564	0.503	0.371	0.291	-0.541	-0.541	0.980
#	0.921	0.969	0.870	0.908	0.945	0.975	0.279	0.279	0.131
I Moran	0.681	0.698	0.793	0.796	0.781	0.732	-0.379	-0.379	0.654
MSE	-0.672	-0.673	-0.770	-0.764	-0.754	-0.694	0.377	0.378	-0.683
KV	-0.709	-0.707	-0.789	-0.784	-0.780	-0.720	0.336	0.336	-0.667
Gaussian semivariogram model									
SDM	0.982								
SDAI1	0.942	0.938							
SDAI2	0.946	0.961	0.992						
SDGI1	0.985	0.986	0.968	0.979					
SDGI2	0.958	0.992	0.940	0.970	0.981				
J1 (or MCD)	0.016	0.089	0.014	0.049	0.013	0.112			
J2	0.016	0.091	0.015	0.050	0.015	0.114	0.999		
SPD	0.474	0.358	0.619	0.529	0.495	0.341	-0.359	-0.361	
$	0.474	0.358	0.619	0.529	0.495	0.341	-0.359	-0.361	1.000
£	0.437	0.329	0.588	0.508	0.472	0.320	-0.397	-0.399	0.989
#	0.864	0.939	0.837	0.889	0.883	0.954	0.268	0.270	0.087
I Moran	0.908	0.849	0.873	0.857	0.911	0.825	-0.193	-0.193	0.684
MSE	-0.838	-0.776	-0.822	-0.802	-0.860	-0.765	0.245	0.246	-0.713
KV	-0.883	-0.822	-0.856	-0.838	-0.894	-0.804	0.221	0.221	-0.702
Spherical semivariogram model									
SDM	0.992								
SDAI1	0.940	0.929							
SDAI2	0.955	0.954	0.994						
SDGI1	0.989	0.983	0.970	0.983					
SDGI2	0.983	0.996	0.941	0.967	0.986				
J1 (or MCD)	0.436	0.510	0.362	0.411	0.388	0.386			
J2	0.435	0.508	0.361	0.410	0.510	0.509	0.999		
SPD	0.416	0.341	0.642	0.572	0.496	0.364	-0.327	-0.328	
$	0.416	0.341	0.642	0.572	0.496	0.364	-0.327	-0.328	1.000
£	0.388	0.318	0.615	0.553	0.478	0.344	-0.369	-0.371	0.989
#	0.925	0.963	0.845	0.887	0.908	0.963	0.696	0.695	0.133
I Moran	0.880	0.838	0.840	0.838	0.893	0.827	-0.024	-0.026	0.598
MSE	-0.825	-0.774	-0.796	-0.787	-0.845	-0.767	0.129	0.131	-0.634
KV	-0.864	-0.816	-0.817	-0.813	-0.989	-0.803	0.058	0.060	-0.595

In addition, this component presents moderate positive correlations with J1 (or MCD) and J2 in the spherical model and weak positive correlations in the exponential and Gaussian models. Santos et al. (2018) also observed strong and positive correlations of the component $\min\left\{1; \dfrac{a}{0.5\ MD}\right\}$ with the SDI, and negative correlations (weak or moderate) with the SPD.

Component $\left(\dfrac{C_1}{C_0 + C_1}\right)$ has a correlation equal to 1 with the SPD. In addition, it has a moderate positive correlation with the SDI in the Gaussian and spherical models and a weak positive correlation in the exponential model. In Santos et al. (2018), weak to moderate correlations were also observed between the SDI and the component $\left(\dfrac{C_1}{C_0 + C_1}\right)$. In Padilha et al. (2024), there was a strong positive correlation in the spherical model and a moderate positive correlation in the exponential and Gaussian models between the SDI and the SPD (or $\left(\dfrac{C_1}{C_0 + C_1}\right)$).

Component $\left(\dfrac{C_1}{C_0 + C_1}\right)$ presents weak positive correlations with the SDM and SDGI2 for the three models and moderate positive correlations with the SDAI1, SDAI2, and SDGI1. Padilha et al. (2024) observed a moderate positive correlation between the SDM and the SPD (or $\left(\dfrac{C_1}{C_0 + C_1}\right)$) in the spherical and exponential models, and a weak correlation in the Gaussian model.

Furthermore, the component $\left(\dfrac{C_1}{C_0 + C_1}\right)$ presents moderate negative correlations with J1 (or MCD) and J2 in the exponential model, and weak negative correlations in the Gaussian and spherical models.

Component $\sqrt[2]{\left(\dfrac{C_1}{C_0 + C_1}\right)}$ has a moderate positive correlation with the SDI in the Gaussian model and weak positive correlations in the exponential and spherical models. The component has a moderate positive correlation with the SDGI1 in the Gaussian and spherical models, and a weak positive correlation in the exponential model. Furthermore, this component presents weak positive correlations with SDM and SDGI2, a strong positive correlation with SPD, and moderate positive correlations with SDAI1 and SDAI2 in all three models.

Finally, it is verified that the component $\sqrt[2]{\left(\dfrac{C_1}{C_0 + C_1}\right)}$ presents moderate negative correlations with J1 (or MCD) and J2 in the exponential model and weak negative correlations in the Gaussian and spherical models.

In theory, spatial dependence metrics are expected to have good correlations with all components to capture the variability in the vertical and horizontal directions of the semivariogram. Thus, SDAI1 and SDAI2 measures tend to have better results when compared to SDI, SDM, SDGI1, SDGI2, SPD, J1 (or MCD), and J2, in this regard.

It should be noted that SDAI1, SDAI2, SDGI1, and SDGI2 are direct metrics (by arithmetic mean and geometric mean) and dimensionless, dependent only on elements inherent to the semivariogram and the sampling grid. The SDI and SDM measures also depend on model factors and require classification of the spatial dependence exclusive to each model (Seidel and Oliveira, 2016; Barbosa et al., 2017; Appel Neto et al., 2018, 2020; Uribe-Opazo et al., 2023). The J1 (or MCD) and J2 measures also depend on the model factors and are given in the range measurement unit (Russo and Jury, 1987; Han et al., 1994). In addition, the SPD measure considers only a vertical component of the semivariogram (Biondi et al., 1994), not considering any horizontal component.

A ranking of the metrics (based on our results and literature findings), in order of performance, is: 1st - SDAI1 and/or SDAI2; 2nd - SDGI1; 3rd - SDGI2 and/or SDI and/or SDM; 4th - SPD; 5th - J1 (MCD) and/or J2.

However, it would not be problematic to use all metrics together, as they all aim to quantify spatial variability. Furthermore, the SDAI1, SDAI2, SDGI1, SDGI2, SDI and SDM proposals are attempts to improve the quantification of spatial variability originally done by SPD (or RNE).

Because the semivariogram is very informative, metrics are needed to quickly describe, through a number, the magnitude of spatial dependence (Seidel and Oliveria, 2014). These measures allow the comparison of different situations or scenarios of spatial dependence (Biondi et al., 1994; Appel Neto et al., 2020). Thus, the quantification of variability within the field can be understood as essential in the current context of precision agriculture.

Below is the *metrics(data, obj)* function for calculating the SDAI1, SDAI2, SDGI1, and SDGI2 measurements and obtaining the respective spatial dependence classifications. The *metrics()* function has the commands *data* and *obj*. The command *data* receives the result of applying the function *read.geodata()*, and the command *obj* receives the result of applying the *variofit()* function from the geoR package (Ribeiro Junior et al., 2020).

```
##########################################
########################
# Function:
metrics <- function(data, obj) {
# VC (Vertical component) and HC (Horizontal component)
VC <- obj$cov.pars[1]/(obj$nugget+obj$cov.pars[1])
HC <- min(1,(obj$practicalRange/(0.5*summary(data)[[3]][[2]])))
# SDAI1 and SDAI2
SDAI1 <- ((VC+HC)/2)*100
SDAI2 <- ((sqrt(VC)+HC)/2)*100
# SDGI1 and SDGI2
SDGI1 <- sqrt(VC*HC)*100
SDGI2 <- sqrt(sqrt(VC)*HC)*100
# classifications
cat("\nMetrics:\n")
cat("\nSDAI1(%) =\n", SDAI1, "\n")
if (SDAI1<38) {
cat("classification: Weak spatial dependence\n")
} else if (SDAI1>=38 && SDAI1<68) {
cat("classification: Moderate spatial dependence\n")
} else {
cat("classification: Strong spatial dependence\n")}
cat("\nSDAI2(%) =\n", SDAI2, "\n")
if (SDAI2<48) {
cat("classification: Weak spatial dependence\n")
} else if (SDAI2>=48 && SDAI2<75) {
cat("classification: Moderate spatial dependence\n")
} else {
cat("classification: Strong spatial dependence\n")}
cat("\nSDGI1(%) =\n", SDGI1, "\n")
if (SDGI1<30) {
cat("classification: Weak spatial dependence\n")
} else if (SDGI1>=30 && SDGI1<64) {
cat("classification: Moderate spatial dependence\n")
```

```
} else {
cat("classification: Strong spatial dependence\n")}
cat("\nSDGI2(%) =\n", SDGI2, "\n")
if (SDGI2<40) {
cat("classification: Weak spatial dependence\n")
} else if (SDGI2>=40 && SDGI2<73) {
cat("classification: Moderate spatial dependence\n")
} else {
cat("classification: Strong spatial dependence\n")}
}
#
metrics(data, obj)
###########################################
```

Finally, we exemplify the application of the *metrics* function to the *soilmoisture* and *NVDI* data from the geotoolsR package (Rossoni and Felix, 2020). For both *soilmoisture* and *NVDI* there is moderate spatial dependence based on the classification and calculation of the metrics SDAI1, SDAI2, SDGI1, and SDGI2 (Table 3).

CONCLUSIONS

This study proposed and validated new metrics to measure spatial dependence, based on two distinct approaches: an additive formulation using the arithmetic mean of the vertical and horizontal components of the semivariogram; and a multiplicative formulation using the geometric mean. In addition, a classification for spatial dependence was proposed. Finally, a function in R was developed to calculate the metrics and classify spatial dependence.

The proposed metrics effectively describe spatial dependence. However, the metrics based on the additive formulation, called spatial dependence arithmetic index 1 and spatial dependence arithmetic index 2, stood out for their strong correlations with performance measures and for capturing more effectively the vertical and horizontal behaviors of the semivariogram.

Thus, faced with the challenge of summarizing all the graphical information of the semivariogram in a single number, thus allowing the classification and comparison of spatial dependencies, we believe that this study represents a significant advance in the approach to this topic.

Table 3. Applying the metrics function to *soilmoisture* and *NVDI* data sets

Data	Partial sill	Sill	Range	MD
soilmoisture	1.62	4.12	78.83	401.62
NVDI	0.0036	0.0065	7.45	24.19
Data	SDAI1	SDAI2	SDGI1	SDGI2
soilmoisture	39.30 (Moderate)	50.99 (Moderate)	39.30 (Moderate)	49.62 (Moderate)
NVDI	58.59 (Moderate)	68.09 (Moderate)	58.51 (Moderate)	67.78 (Moderate)

MD: maximum distance; SDAI1: spatial dependence arithmetic index 1; SDAI2: spatial dependence arithmetic index 2; SDGI1: spatial dependence geometric index 1; SDGI2: spatial dependence geometric index 2.

DATA AVAILABILITY

The data will be provided upon request.

AUTHOR CONTRIBUTIONS

Conceptualization: Enio Júnior Seidel (equal) and Marcelo Silva de Oliveira (equal).

Formal analysis: Enio Júnior Seidel (lead).

Methodology: Enio Júnior Seidel (lead).

Software: Enio Júnior Seidel (lead).

Writing – original draft: Enio Júnior Seidel (equal) and Marcelo Silva de Oliveira (equal).

Writing – review & editing: Enio Júnior Seidel (equal) and Marcelo Silva de Oliveira (equal).

REFERENCES

Amaral LR, Della Justina DD. Spatial dependence degree and sampling neighborhood influence on interpolation process for fertilizer prescription maps. Eng Agric. 2019;39:85-95. https://doi.org/10.1590/1809-4430-Eng.Agric.v39nep85-95/2019

Appel Neto E, Barbosa IC, Seidel EJ, Oliveira MS. Spatial dependence index for cubic, pentaspherical and wave semivariogram models. Bol Cienc Geod. 2018;24:142-51. https://doi.org/10.1590/S1982-21702018000100010

Appel Neto E, Seidel EJ, Oliveira MS. Geostatistical-based index for spatial variability in soil properties. Rev Bras Cienc Solo. 2020;44:e0200086. https://doi.org/10.36783/18069657rbcs20200086

Barbosa IC, Appel E, Seidel EJ, Oliveira MS. Proposal of the spatial dependence evaluation from the power semivariogram model. Bol Cienc Geod. 2017;23:461-75. https://doi.org/10.1590/S1982-21702017000200031

Biondi F, Myers DE, Avery CC. Geostatistically modeling stem size and incremente in an old-growth forest. Can J Forest Res. 1994;24:1354-68. https://doi.org/10.1139/x94-176

Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, Konopka AE. Field-scale variability of soil properties in central Iowa soils. Soil Sci Soc Am J. 1994;58:1501-11. https://doi.org/10.2136/sssaj1994.03615995005800050033x

Fu W, Zhao K, Zhang C, Tunney H. Using Moran's I and geostatistics to identify spatial patterns of soil nutrients in two different long-term phosphorus-application plots. J Plant Nutr Soil Sc. 2011;174:785-98. https://doi.org/10.1002/jpln.201000422

Geary RC. The contiguity ratio and statistical mapping. J Roy Stat Soc D-Sta. 1954;5:115-41. https://doi.org/10.2307/2986645

Han S, Hummel JW, Goering CE, Cahn MD. Cell size selection for site-specific crop management. Trans Amer Soc Agr Eng. 1994;37:19-26. https://doi.org/10.13031/2013.28048

Herrero-Langreo A, Gorretta N, Tisseyre B, Gowen A, Xu JL, Chaix G, Roger JM. Using spatial information for evaluating the quality of prediction maps from hyperspectral images: A geostatistical approach. Anal Chim Acta. 2019;1077:116-28. https://doi.org/10.1016/j.aca.2019.05.067

Jang G, Kim D-W, Kim H-J, Chung YS. Short communication: Spatial dependence analysis as a tool to detect the hidden heterogeneity in a kenaf field. Agronomy. 2023;13:428. https://doi.org/10.3390/agronomy13020428

Journel AG, Huijbregts CJ. Mining geostatistics. Caldwell: Blackburn Press; 2003.

Leroux C, Tisseyre B. How to measure and report within-field variability: a review of common indicators and their sensitivity. Precis Agric. 2019;20:562-90. https://doi.org/10.1007/s11119-018-9598-x

Moran PAP. Notes on continuous stochastic phenomena. Biometrika. 1950;37:17-23. https://doi.org/10.1093/biomet/37.1-2.17

Olea RA. A six-step practical approach to semivariogram modeling. Stoch Env Res Risk A. 2006;20:307-18. https://doi.org/10.1007/s00477-005-0026-1

Padilha LC, Pazini JB, Seidel EJ. Relationship between some geostatistical-based measures for agricultural attributes. Rev Cienc Agrovet. 2024;23:782-7. https://doi.org/10.5965/223811712342024782

R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021. Available from: http://www.R-project.org/.

Ribeiro Junior PJ, Diggle PJ, Schlather M, Bivand R, Ripley B. geoR: Analysis of Geostatistical Data. R package version 1.8-1. CRAN: Package geoR; 2020. Available from: https://CRAN.R-project.org/package=geoR.

Rodrigues MS, Castrignanò A, Belmonte A, Silva KA, Lessa BFT. Geostatistics and its potential in Agriculture 4.0. Rev Cienc Agron. 2020;51:e20207691. https://doi.org/10.5935/1806-6690.20200095

Rossoni DF, Felix VB. geotoolsR: Tools to Improve the use of Geostatistic. R package version 1.1. CRAN: Package geotoolsR; 2020. Available from: https://CRAN.R-project.org/package=geotoolsR

Russo D, Jury W. A theoretical study of the estimation of the correlation scale in spatially variable fields: 1. Stationary fields. Water Resour Res. 1987;23:1257-68. https://doi.org/10.1029/WR023i007p01257

Sandonís-Pozo L, Llorens J, Escolà A, Arnó J, Pascual M, Martínez-Casasnovas JA. Satellite multispectral indices to estimate canopy parameters and within-field management zones in super-intensive almond orchards. Precis Agric. 2022;23:2040-62. https://doi.org/10.1007/s11119-022-09956-6

Santos ELE, Seidel EJ, Pazini JB, Oliveira MS, Appel Neto E, Barbosa IC. Some aspects about the spatial dependence index for variability of soil attributes. Cienc Rural. 2018;48:e20170710. https://doi.org/10.1590/0103-8478cr20170710

Seidel EJ, Oliveira MS. Descrição da dependência espacial em Geoestatística através da construção de dois índices. In: Anais da 57ª Reunião Anual da Região Brasileira da Sociedade Internacional de Biometria; 2012; Piracicaba. RBRAS; 2012.

Seidel EJ, Oliveira MS. Novo índice geoestatístico para a mensuração da dependência espacial. Rev Bras Cienc Solo. 2014;38:699-705. https://doi.org/10.1590/S0100-06832014000300002

Seidel EJ, Oliveira MS. A classification for a geostatistical index of spatial dependence. Rev Bras Cienc Solo. 2016;40:e0160007. https://doi.org/10.1590/18069657rbcs20160007

Taylor JA, Praat JP, Bollen AF. Spatial variability of kiwifruit quality in orchards and its implications for sampling and mapping. HortScience. 2007;42:246-50. https://doi.org/10.21273/HORTSCI.42.2.246

Uribe-Opazo MA, Dalposso GH, Galea M, Johann GA, Bastiani F, Moyano ENC, Grzegozewski DM. Spatial variability of wheat yield using the gaussian spatial linear model. Aust J Crop Sci. 2023;17:179-89. https://doi.org/10.21475/ajcs.23.17.02.p3742

Zhao M, Guerrero A, Munnaf MA, Lauwers L, Mouazen AM. Within-field spatial variability and potential for profitability of variable rate applications. Precis Agric. 2023;24:2248-63. https://doi.org/10.1007/s11119-023-10039-3