




# Identifying Covariates to Assess the Spatial Variability of Saturated Soil Hydraulic Conductivity Using Robust Cokriging at the Watershed Scale

Mauricio Fornalski Soares<sup>1</sup> · Luana Nunes Centeno<sup>2</sup> · Luís Carlos Timm<sup>3</sup>  · Carlos Rogério Mello<sup>4</sup> · Douglas Rodrigo Kaiser<sup>5</sup> · Samuel Beskow<sup>1</sup>

Received: 30 September 2019 / Accepted: 11 March 2020 / Published online: 27 April 2020  
© Sociedad Chilena de la Ciencia del Suelo 2020

## Abstract

The mapping of saturated soil hydraulic conductivity ( $K_{\text{Sat}}$ ) is essential to understanding soil water dynamics and is a sensitive input in hydrological modeling. The objectives of this study were to provide a reference for the selection of soil hydrology and other environmental attributes that can be used as covariates for estimating  $K_{\text{Sat}}$  and to compare the efficiency of univariate ordinary kriging versus ordinary robust cokriging, using selected soil hydrology and environmental attributes. Data sets were obtained from a sample grid of 179 points established in the Ellert creek watershed (ECW), located in Rio Grande do Sul state, Southern Brazil.  $K_{\text{Sat}}$ , macroporosity, microporosity, total porosity, and bulk density were determined from soil sampled at each point. Data of land use and elevation were also applied. All data sets were firstly submitted to classical statistics. Boxplot graphics were constructed to evaluate the relationship between  $K_{\text{Sat}}$  and land uses. Spearman coefficient of correlation between  $K_{\text{Sat}}$  and the other attributes was also assessed. For the assortment of covariates, cluster analysis was applied. Classical and robust estimators were applied to calculate the auto and cross-semivariograms and hereafter the ordinary kriging and cokriging. The Spearman coefficient showed some inconsistencies among the applied variables, suggesting that the multivariate method was more appropriate. All cross-semivariograms, except for land use, showed results with better accuracy than the auto-semivariograms. From the methods applied, the best estimates of  $K_{\text{Sat}}$  were obtained using the robust cokriging method, using macroporosity and soil bulk density as covariates.

**Keywords** Cross-semivariogram · Robust estimation · Covariate

---

✉ Luís Carlos Timm  
lctimm@ufpel.edu.br

Mauricio Fornalski Soares  
mauriciofornalski@gmail.com

Luana Nunes Centeno  
luanununescenteno@gmail.com

Carlos Rogério Mello  
crmello@deg.ufla.br

Douglas Rodrigo Kaiser  
douglasrodrigokaiser@gmail.com

Samuel Beskow  
samuelbeskow@gmail.com

<sup>1</sup> Center of Technological Development, Federal University of Pelotas, Campus Porto, Rua Gomes Carneiro, n. 01, Pelotas, Rio Grande do Sul 96010-610, Brazil

<sup>2</sup> Center of Technological Development, Federal University of Pelotas, Pelotas, Rio Grande do Sul, Brazil

<sup>3</sup> Department of Rural Engineering, Faculty of Agronomy, Federal University of Pelotas, Campus Universitário s/n, Capao do Leão, Rio Grande do Sul 96010-900, Brazil

<sup>4</sup> Water Resources Department, Federal University of Lavras, Lavras, Minas Gerais, Brazil

<sup>5</sup> Federal University of Fronteira Sul, Cerro Largo, RS, Brazil

## 1 Introduction

The variability of soils, both spatially and temporally, has long been recognized in soil science (Jenny 1941; Lin 2006; Simonson 1959; Wang and Shi 2018). Techniques that capture soil spatial and temporal variability have recently evolved to focus the attention of soil scientists on soil attributes related to hydrological models (Qiao et al. 2018; She et al. 2017). According to Libohova et al. (2018), the success of predictions when modeling hydrological processes depends on the accurate representation of the spatial and temporal variability of soil hydrological attributes and the main external factors, such as climate, land use and management, and soil water dynamics.

Saturated soil hydraulic conductivity ( $K_{\text{Sat}}$ ) is a key soil attribute for understanding the soil water movement phenomenon and is therefore one of the main inputs used in hydrological models. Several studies have reported that  $K_{\text{Sat}}$  values are affected by the variation of pore space geometry (Baiamonte et al. 2017), topography (Wang et al. 2013), land use (Price et al. 2010; Salemi et al. 2013; Pinto et al. 2019), and scale (Picciafuoco et al. 2019). It has also been well documented in the literature that  $K_{\text{Sat}}$  is highly variable in space (Wang et al. 2013; Papanicolaou et al. 2015; She et al. 2017). Therefore, due to its high variability as well as to the effort required for  $K_{\text{Sat}}$  data sampling, representativeness of the spatial variability of  $K_{\text{Sat}}$  at the watershed scale is often difficult to obtain (Reichardt and Timm 2020).

Geostatistical tools have been used to quantify and map the spatial variability of soil hydrology attributes aimed at supporting hydrological modeling at the watershed scale. The classical estimator of Matheron is the most used semivariogram estimator to characterize and quantify the spatial variability structure of soil attributes. However, due to its high sensitivity to the existence of outliers, and its requirement of a normal distribution in order to estimate the underlying process (Lark 2000), its use is difficult in field conditions, especially for attributes related to soil water dynamics (Lebrez and Bárdossy 2017). In order to overcome these drawbacks, data transformation (logarithm, exponential, root fourth, etc.) from the original scale of measurement has been adopted (Song et al. 2019; Wallin and Bolin 2015). However, Goovaerts (1997) indicated that data transformation is not ideal if the aim is prediction, since in general those who ultimately use the predictions, such as land managers, hydrological modelers, and environmental scientists and so on, require values on the original scale of measurement which involves a back-transformation. Therefore, several robust estimators, such as that of Cressie and Hawkins (1980), have been proposed to minimize the effects of the presence of outliers or for use on datasets with non-normal distributions for calculating experimental semivariance values.

The determination of  $K_{\text{Sat}}$  can be costly, and much effort is required to obtain representative spatial variability at the watershed scale. Therefore, multiple regression models (namely pedotransfer functions (PTFs)) have been developed to estimate  $K_{\text{Sat}}$  from easily available soil attributes from databases. For instance, Boadu (2000) developed regression-based models for estimating the  $K_{\text{Sat}}$  of compacted soils using fractal dimension, entropy, porosity, percent of fine soil particles, and the bulk density as predictors. However, some values (e.g., fractal dimension, entropy, and soil porosity) are not readily available from soil databases.

Another alternative is to use geostatistical tools such as the cross-semivariogram to estimate  $K_{\text{Sat}}$  from an auxiliary variable of simple determination and/or low cost, since both variables are well correlated. In this sense, robust estimators can be applied for the calculation of the cross-semivariogram between  $K_{\text{Sat}}$  and other soil (such as porosity, textural fractions, bulk density, etc.), topographical (elevation, slope, aspect, etc.), and environmental (climate, land use, etc.) attributes when such data are available.

Adhikary et al. (2017) mentioned that the ordinary cokriging method in general reduces the variance of the prediction error and improves estimations compared with univariate ordinary kriging method, when the auxiliary variable is well correlated with the primary one. Wang and Shi (2018) used robust cross-semivariograms in the cokriging process to estimate soil particles (clay, silt, and sand) in a large watershed, applying some of the estimators proposed by Lark (2003).

The spatial variability of  $K_{\text{Sat}}$  in watersheds is recognized as a complex phenomenon, due mainly to the natural and anthropogenic processes that influence soil water dynamics, predominantly at the surface layer (Becker et al. 2018). In this sense, an auxiliary variable (or covariate) should consider some characteristics of the soil or the environment that are quantifiable and related to the occurrence or the impediment of the target phenomenon. Acquiring these auxiliary data should be operationally more viable compared with the acquisition of the target variable. Nowadays, topographical attributes are widely available and can be readily obtained through the geographic information system (GIS) using digital elevation models (DEMs) and digital terrain analysis techniques (She et al. 2017). Additionally, if soil hydrology attributes such as  $K_{\text{Sat}}$  can have their estimates optimized using land features as covariates obtained by GIS, costs and time can be saved (She et al. 2017).

The development of the Ellert Creek Watershed (ECW) is regionally strategic as it is the main watercourse that flows directly to the Pelotas River, supplying drinking water for the population of Pelotas, a city with some 329,000 inhabitants (Beskow et al. 2016). The ECW is a headwater watershed of the Pelotas River Watershed (PRW), where local events are possibly reflected in the hydrological behavior of

the downstream PRW. Furthermore, the ECW has suffered from the impacts of human activities, being vulnerable to water erosion and floods, causing soil and nutrient losses resulting in economic and social damage.

Soil hydrology attributes (e.g.,  $K_{\text{Sat}}$ ) are key to understanding the processes of the hydrological cycle and are critical information for the application of hydrological models aimed at supporting decisions on water resource management (Beskow et al. 2016; She et al. 2017). Among them,  $K_{\text{Sat}}$  is a sensitive model input for the application of distributed hydrological models (Wang et al. 2013; Hu et al. 2015). Therefore, the objectives of this study were to (i) provide a reference for the selection of soil hydrology and environmental attributes that can be used as covariates for  $K_{\text{Sat}}$  estimations at the watershed scale using the non-linear Spearman correlation coefficient and multivariate analysis, (ii) assess the performance of ordinary kriging for mapping the spatial distribution of  $K_{\text{Sat}}$  using Matheron and Cressie and Hawkins estimators, and (iii) assess the performance of the ordinary cokriging method to map  $K_{\text{Sat}}$  using the Cressie and Hawkins cross-semivariogram estimator and each selected attribute as an auxiliary variable.

## 2 Materials and Methods

### 2.1 Description of the Study Site

The study was carried out in a headwater sub-basin of the Pelotas River Watershed, known as the ECW, located in the municipality of Canguçu, Southern Rio Grande do Sul (RS) state, approximately 50 km northeast of municipality of Pelotas (Fig. 1). The ECW has an area of approximately 0.7 km<sup>2</sup>, and the altitude varies from 310 to 419 m. According to the Köppen climate classification, ECW is Cfa type, a mesothermal climate indicating wet subtropical conditions, and characterized by an annual average temperature of 18 °C, with hot summers and cold winters (Kuinchtner and Buriol 2001). The precipitation is well distributed throughout the year, and its mean annual value is around 1350 mm (IBGE 1986). The regional relief varies from undulating to strong undulating, with a predominance of native forest or sparse shrub, and shallow soils.

The watershed comprises only one soil class, the *Neossolos* (Brazilian Soil Classification System) or Entisols in soil taxonomy, identified using the soil map developed by Embrapa (1999), and updated by the latest version of the Brazilian Soil Classification System (Embrapa 2006). The ECW was selected due to its economic and social importance to the municipality of Canguçu, an area greatly dominated by family farming systems. The ECW main watercourse flows directly to the Pelotas River, located in the southern part of the Rio Grande do Sul state, with a total area of approximately 940 km<sup>2</sup>,

including the municipalities of Pelotas, Morro Redondo, Arroio do Padre, and Canguçu.

### 2.2 Sample Grid

A sample grid of 106 sampling points spaced 50 m in the west by 75 m in the south direction was established at the beginning of field work. Subsequently, in order to better model the spatial variability structure of soil hydrology attributes, 78 soil samples were collected in a specific area of the ECW watershed spaced 25 m at both directions, totaling 184 sampling points (Fig. 1). This specific landscape was selected due to its diverse features, including different land uses, soil textural classes, and features of the topographical terrain. The geographical location of each point in the ECW area was recorded using GPS navigation equipment. ArcGIS software was used to establish the sample grid and to obtain the UTM coordinates of each point (Environmental Systems Research Institute, Redlands, CA).

### 2.3 Soil and Topographic Attributes

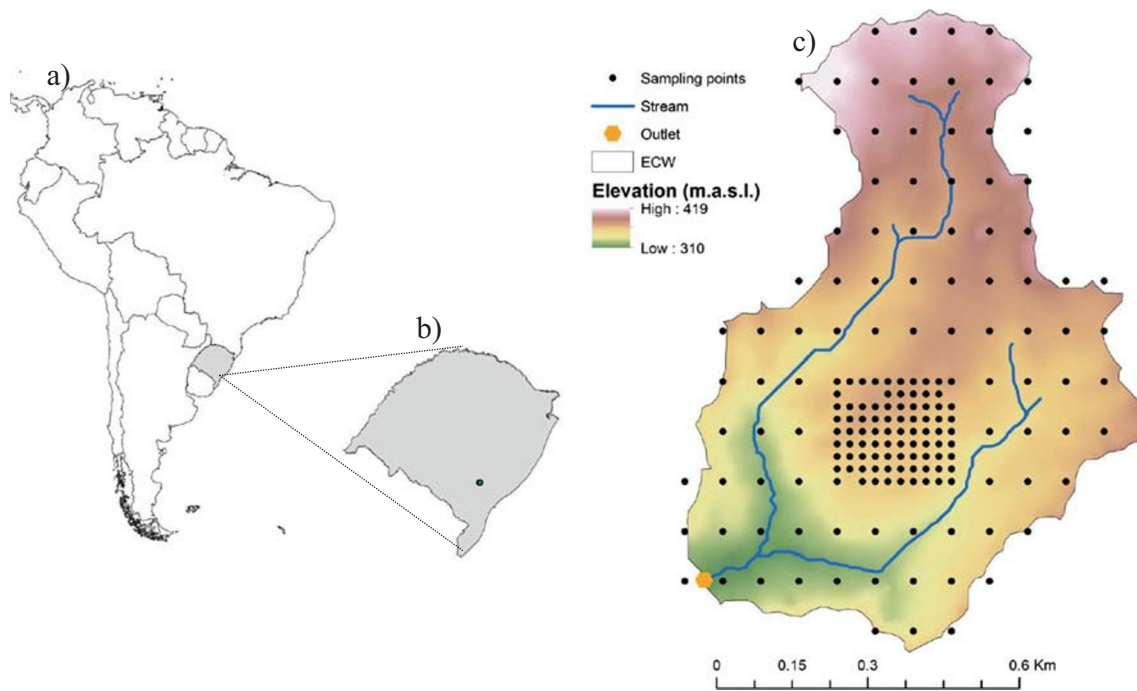
Undisturbed soil samples were collected at all 184 sampling points from the 0–0.20-m layer using metallic cylinders with a 5.0-cm height and a 4.8-cm internal diameter. It is important to highlight that the samples were collected from the topsoil layer (0–20 cm), the area of the watershed most affected by different land management systems (Alvarez and Steinbach 2009).

The following soil hydrology attributes were determined: soil bulk density (BD) (Blake and Hartge 1986), soil total porosity (TP), macroporosity (Mac), and microporosity (Mic) (Klute 1986). Saturated soil hydraulic conductivity ( $K_{\text{Sat}}$ ) was measured at each point using the constant head method (Klute and Dirksen 1986).

The main land uses identified in the ECW were forest, silviculture, annual cropping, and pasture. Forest here included native forest sites which accounted for around 10% of the ECW's area; areas of silviculture were characterized by the occurrence of *Pinus* sp., *Eucalyptus* sp., and *Acacia* sp., equivalent to 12% of the ECW. Annual cropping areas included the cultivation of *Glycine max*, *Zea mays*, and *Nicotiana tabacum* which occupied the largest area of the ECW (71%). Pastures were mainly characterized by grazing areas, accounting for 7% of the total area of the ECW.

### 2.4 Exploratory Analysis

All data sets were submitted to basic and descriptive statistics, calculating position (arithmetic average and median) and dispersion (standard deviation, variance, and coefficient of variation) measures as well as skewness and kurtosis coefficients, which are related to the characteristics of the data distribution.



**Fig. 1** Location, topography, hydrography, and sampling points of the Ellert Creek Watershed. **a** South America, **b** Rio Grande do Sul State, and **c** Ellert Creek Watershed

The Kolmogorov-Smirnov (K-S) test at 5% significance was applied to verify normality of all data sets. Additionally, boxplots were constructed to evaluate the relationship between  $K_{\text{Sat}}$  and the identified land uses. The Spearman coefficient of correlation between  $K_{\text{Sat}}$  and the other soil hydrology attributes and environmental features was also calculated.

## 2.5 Cluster Analysis

Before the multivariate statistical analyses, each data set was standardized (mean = 0 and standard deviation = 1), with the aim of transforming the original data to the same order of magnitude. Because the variables present different orders of magnitude, the use of the standardized variables requires that the variance and covariance matrixes are the same as the correlation matrix. Cluster analysis was applied to construct dendrograms with the aim of grouping the standardized variables into homogeneous groups (Everitt and Hothorn 2011), using the Mojena's (1977) criterion for cutoff. The dissimilarity between the standardized variables is a distance measure. The hierarchical cluster technique uses a Euclidean distance (metric) to separate a set of objects (variables) into groups according to chosen criteria. The group-average method was used as an aggregation criterion. It is a weighted method which calculates the distance between two clusters as the arithmetic average distance from observations in one cluster to observations in another cluster. This method tends to combine clusters that have small variances and may produce new clusters that have the same variance (Ramos et al. 2007). The

number of groups was defined based on the criterion of the square root of the number of variables (Bitencourt et al. 2016).

## 2.6 Geostatistical Analysis

The semivariograms used for geostatistical analysis may explain the spatial variability of  $K_{\text{Sat}}$ , and are calculated from the set  $(Z_u(x_i), i = 1, 2, \dots, N)$ , usually by the Matheron classical estimator (Webster and Oliver 2007), defined by Eq. (1):

$$\gamma_{Z_u}^M(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z_u(x_i) - z_u(x_i + h)]^2 \quad (1)$$

where  $\gamma_{Z_u}^M(h)$  is the semivariance value using the Matheron estimator,  $z_u(x_i)$  and  $z_u(x_i + h)$  are values of  $Z_u$  at locations  $x_i$  and  $x_i + h$ , respectively, and  $N(h)$  is the number of pairs  $[z_u(x_i), z_u(x_i + h)]$  separated by the lag distance  $h$ .

Lark (2000) studied several robust estimators for experimental semivariograms, which deal better with the presence of outliers in the datasets. One of these estimators was developed by Cressie and Hawkins (1980) and is presented in Eq. (2):

$$\gamma_{Z_u}^{CH}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \frac{\left\{ \left[ |z_u(x_i) - z_u(x_i + h)|^{\frac{1}{2}} \right]^4 \right\}}{0.457 + \frac{0.494}{N(h)}} \quad (2)$$

where  $\gamma_{Z_u}^{CH}(h)$  is the semivariance value using the Cressie-Hawkins estimator. This estimator was designed to calculate the semivariogram of a primary process with differences

normally distributed even in the presence of contaminants of a secondary process.

Cressie and Hawkins (1980) showed that the square root of the absolute difference had a normal distribution with no significant bias. Therefore, the effect of outliers is damped by the square root of the absolute differences and then increasing that values to the fourth power to resize the semivariogram scale. The denominator of Eq. (2) is a correction factor based on the assumption that the underlying process has the differences normally distributed over all lags (Lark 2000). The breakdown point of this estimator is zero, and for this reason, it may present some difficulty by not establishing any limits on the effect of outliers. In this case, a very large outlier could make the estimate non-representative (Genton 1998). However, in practice, it is an estimator that smooths the discrepant values and does not set absolute limits on its effect.

The classical cross-semivariogram of Matheron is defined by Eq. (3), being an extension of Eq. (1), where  $Z_u$  and  $Z_v$  are the target and the auxiliary variables, respectively.

$$\gamma_{Z_u, Z_v}^M(h) = \frac{1}{2N(h)} \times \sum_{i=1}^{N(h)} [z_u(x_i) - z_u(x_i + h)]^2 \cdot [z_v(x_i) - z_v(x_i + h)]^2 \tag{3}$$

where  $\gamma_{Z_u, Z_v}^M(h)$  is the cross-semivariance value between the variables  $Z_u$  and  $Z_v$ , the latter obtained by means of the Matheron estimator.

For the Cressie and Hawkins estimator, the cross-semivariogram follows the same rule described above, where the covariate assists in robust semivariance estimates, according to Eq. (4):

$$\gamma_{Z_u, Z_v}^{CH}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \frac{\left\{ [z_u(x_i) - z_u(x_i + h)]^{\frac{1}{2}} \right\}^4 \left\{ [z_v(x_i) - z_v(x_i + h)]^{\frac{1}{2}} \right\}^4}{0.457 + \frac{0.494}{N(h)}} \tag{4}$$

where  $\gamma_{Z_u, Z_v}^{CH}(h)$  is the cross-semivariance value between  $Z_u$  and  $Z_v$  variables, the latter obtained by means of the Cressie-Hawkins estimator.

### 2.7 Cross-Validation

The cross-validation technique was used to verify the quality of the geostatistical results, comparing estimated data by kriging and cokriging with  $K_{Sat}$  measurements. The statistical parameters used to evaluate the performance of the cross-validation technique were the mean error (ME, Eq. (5)), the square root of the mean squared error (RMSE, Eq. (6)), and the coefficient of determination ( $r^2$ ).

$$ME = \frac{1}{n} \sum_{i=1}^n (e_i - m_i) \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (e_i - m_i)^2} \tag{6}$$

where  $n$  = number of sampling points,  $e_i$  = estimated value of  $K_{Sat}$ , and  $m_i$  = observed  $K_{Sat}$  value. The ME is an indicator of the accuracy of the estimation, showing the tendency of the interpolator to overestimate the values if positive or to underestimate if negative. RMSE quantifies the dispersion of the measured and estimated values around the 1:1 line.

## 3 Results

### 3.1 Exploratory Analyses

The summary statistics of soil saturated hydraulic conductivity ( $K_{Sat}$ ) and the other six associated attributes (BD, TP, Mac, Mic, Elev, and Land Use) for the 179 points (five soil samples were discarded during laboratory procedures) are presented in Table 1. The frequency distribution of  $K_{Sat}$  and  $TK_{Sat}$  (normalized  $K_{Sat}$  data set using the fourth root transformation) is presented in Fig. 2. According to the classification of Wilding and Drees (1983), the coefficients of variation (CV) for BD, TP, Elev, and Land Use can be considered low ( $CV \leq 15\%$ ), whereas the Mac and Mic CV values were considered moderate ( $15\% < CV \leq 35\%$ ).

The crossbar of the boxplot graphics (Fig. 3) shows the median for each dataset for different land uses. The length of the box reflects the interquartile range, and the vertical tails (or bars) of the boxplot are marked by the extremes (either the highest or lowest observed data); however, such do not qualify as outliers. Outliers are defined as data  $\geq 1.5$  times greater (upper fence) or  $\geq 1.5$  times lower (lower fence) than the interquartile range.

### 3.2 Cluster Analysis

Table 2 presents the correlation matrix between attributes, showing that between  $K_{Sat}$  and the associated attributes (BD, TP, Macro, Micro, Elev, and Land Use), all linear correlations were significant at the 1% probability level, with the exception of elevation.

The dendrogram plotted from topographic and soil hydrology attributes and land use in the ECW are shown in Fig. 4. The cutoff was defined at 70% taking into account the statistical stopping rule suggested by Mojena (1977).

### 3.3 Geostatistical Analysis

In this study, isotropic experimental semivariograms are estimated for  $K_{Sat}$ , assuming an identical spatial correlation in all

**Table 1** Exploratory analysis of the data sampled in the study watershed

Attrib	Number	$\bar{X} \pm SE$	Med	$s$	CV (%)	Skew	Kurt	K-S	Norm
$K_{Sat}$	179	$0.81 \pm 0.06$	0.52	0.80	99.4	1.38	1.46	0.175	No
Mac	179	$20.6 \pm 0.38$	20.4	5.05	24.3	1.17	4.10	0.102	No
Mic	179	$26.1 \pm 0.59$	26.4	7.89	30.3	0.11	0.79	0.075	Yes
TP	179	$46.7 \pm 0.49$	46.4	6.62	14.1	1.01	1.61	0.137	No
BD	179	$1.41 \pm 0.01$	1.42	0.17	12.4	-1.01	1.61	0.151	No
Elev	179	$362.5 \pm 1.32$	362.2	17.6	4.8	0.01	0.79	0.096	No

*Attrib* attributes,  $K_{Sat}$  saturated soil hydraulic conductivity ( $m\ h^{-1}$ ), *Mac* macroporosity (%), *Mic* microporosity (%), *TP* total porosity (%), *BD* soil bulk density ( $g\ cm^{-3}$ ), *Elev* elevation (m), *Number* number of data points,  $\bar{X}$  mean value, *SE* standard error, *Med* median value,  $s$  standard deviation, *CV* (%) coefficient of variation, *Skew* coefficient of skewness, *Kurt* coefficient of kurtosis, *K-S* Kolmogorov-Smirnov test value (K-S critical value 0.086), *Norm* normality

directions and neglecting the influence of anisotropy on the semivariogram parameters. Isotropy is assumed for methodological simplicity. Isotropy is a feature of a natural process or data in which directional influence is considered insignificant and spatial dependence (autocorrelation) changes only with the distance between two lags (Johnston et al. 2001).

In Table 3, the fitted parameters of the semivariogram and cross-variogram models and their respective degree of spatial dependence (DSD) values are presented.

The DSD calculated from the univariate semivariograms of  $K_{Sat}$  was 43.2% for the transformed data, applying the classical Matheron estimator  $[\gamma_{Z_u}^M(h)]$ . On the other hand, a DSD of 22% was obtained for the untransformed  $K_{Sat}$  data using the robust estimator of Cressie and Hawkins  $[\gamma_{Z_u, Z_v}^{CH}]$ , indicating that these data have a moderate and weak spatial dependence (Cambardella et al. 1994). Following the same classification, the cross-semivariograms presented moderate spatial dependence ( $K_{Sat} \times Macro = 33.2\%$ ;  $K_{Sat} \times TP = 34.8\%$ ;  $K_{Sat} \times BD = 36.2\%$ ), except for the  $K_{Sat} \times LU$  (land use) cross-semivariogram (2.1%), which presented a weak spatial dependence. The ranges for the univariate semivariograms of  $K_{Sat}$  were 168.5 m for  $\gamma_{Z_u}^M(h)$  and 178.8 m for  $\gamma_{Z_u}^{CH}(h)$ . For

$\gamma_{Z_u, Z_v}^{CH}(h)$ , the range values were 152.9, 213.1, 161.8, and 212.7 m.

The maps developed by kriging methods using the parameters of  $\gamma_{Z_u}^M(h)$  ( $TK_{Sat}$ ) and  $\gamma_{Z_u}^{CH}(h)$  ( $RK_{Sat}$ ) are presented in Fig. 5. Furthermore, the cross-validation statistical precision (ME, RMSE, and  $r^2$ ) between the observed and estimated  $K_{Sat}$  values using univariate kriging and cokriging are shown in Table 4.

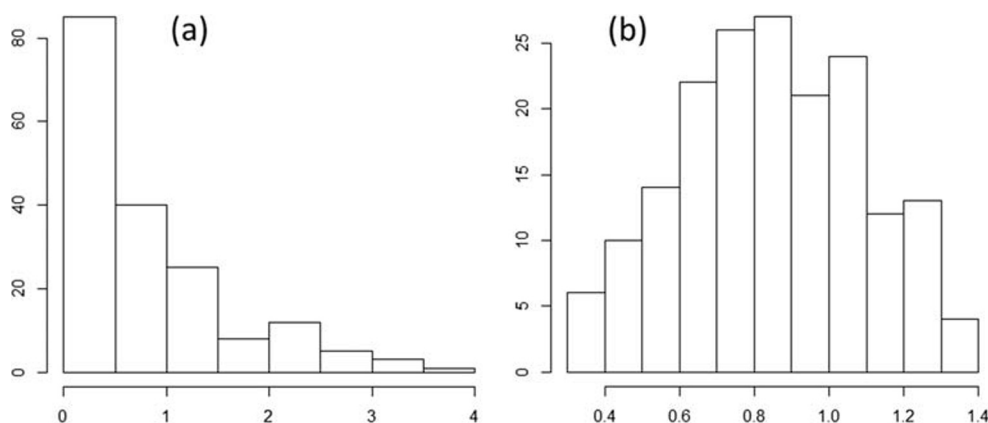
Figure 6 presents the scattered points around the 1:1 line from the cross-validation process, which was applied for each of the studied combinations, showing the accuracy of the geostatistical procedures applied in this study.

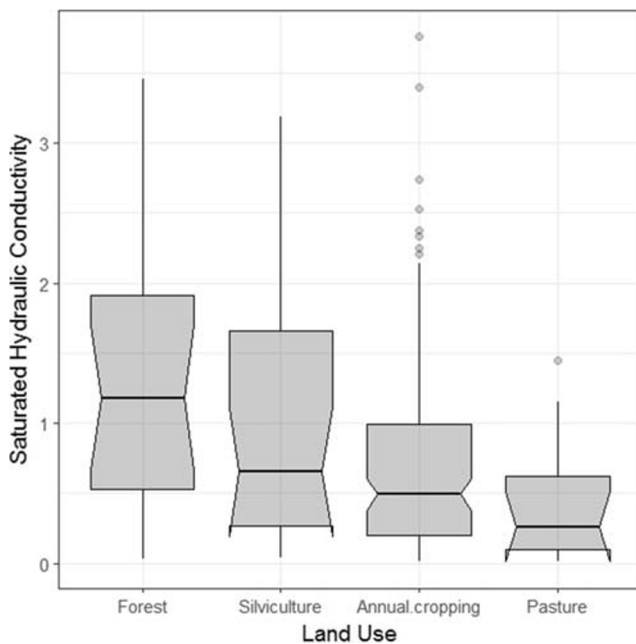
## 4 Discussion

### 4.1 Exploratory Analyses

The average  $K_{Sat}$  and the median corresponded to  $0.81\ m\ h^{-1}$  and  $0.52\ m\ h^{-1}$ , respectively. This contrast shows the great variability of this soil hydrology attribute, also demonstrated by the coefficient of variation (99.4%), being classified as having a high variability ( $CV > 35\%$ ) according to Wilding

**Fig. 2** **a** Histogram of the saturated soil hydraulic conductivity ( $K_{Sat}$ ) values and **b** histogram of transformed data by the fourth root indicating a Gaussian behavior.  $TK_{Sat}$ :  $K_{Sat}$  normalized by fourth root  $[K_{Sat}^{(1/4)}]$





**Fig. 3** Boxplot of  $K_{Sat}$  for the different land uses observed in SEW

and Drees (1983). Godoy et al. (2019) and She et al. (2014) found a CV for  $K_{Sat}$  of 122.0 and 50.8% in studies conducted in central-western Brazil and in the Loess Plateau region (China), respectively. Even when the data sets were transformed to logarithmic (ln), She et al. (2017) obtained a CV above 35%.

All datasets followed a normal distribution as determined by applying the K-S test.  $K_{Sat}$  data presented a non-normal distribution, agreeing with results widely reported in the literature (Ahuja et al. 2010; Godoy et al. 2018; She et al. 2017). Therefore, the original  $K_{Sat}$  data was fourth root transformed (Elsenbeer et al. 1992) in order to use Matheron’s classical semivariance estimator (Fig. 2).

The boxplot analysis (Fig. 3) of  $K_{Sat}$  in the different types of land uses shows a decrease in the saturated soil hydraulic

conductivity following the areas of forest, silviculture, annual cropping, and pasture. Forest soils are associated with a higher infiltration capacity (Archer et al. 2013; Kurnianto et al. 2018; Wood and King 1977) and a lower generation of surface runoff (Germer et al. 2010; Kurnianto et al. 2018; Liu et al. 2018) than soils under other types of vegetation. This occurs in nature due to preferential flows in the soil profile as a result of soil biota activity and higher organic matter content, as reported by Pinto et al. (2015), Ma et al. (2017), and Pinto et al. (2019).

The median  $K_{Sat}$  values for forest, silviculture, annual cropping, and pasture were 1.18, 0.65, 0.49, and 0.26  $m\ h^{-1}$ , respectively (Table 1). Annual cropping showed a greater number of outliers, suggesting a greater variability of  $K_{Sat}$  in this area of land use. Despite this variability being linked to the different farming activities practiced in the watershed, we did not investigate changes in  $K_{sat}$  patterns over the time within the same management system, which would lead to another line of investigation. In the agricultural areas of the watershed, different farming systems with differing levels of mechanization, and technology such as conventional tillage, no-tillage and minimum tillage were observed, increasing the variability of  $K_{sat}$ .

Several studies have shown that land use systems affect  $K_{Sat}$  values and therefore their spatial variation in a watershed (Price et al. 2010; Salemi et al. 2013; Pinto et al. 2019). Therefore, if properly quantified, it is possible to improve  $K_{Sat}$  estimates in non-sampled points when used as a covariate in a cokriging interpolation.

### 4.2 Cluster Analyses

The value of the Spearman correlation coefficient between  $K_{Sat}$  vs. Land Use was low (0.21), although it was significant at a probability level of 1%. A correlation of 0.5 was observed between  $K_{Sat}$  vs. Mac and  $K_{Sat}$  vs. BD; however, in the case of  $K_{Sat}$  vs. BD, the correlation was negative. The highest correlation value was found between  $K_{Sat}$  and TP (0.72). The values of Mac and TP were directly related with  $K_{Sat}$ , which is corroborated by several other studies, where values of hydraulic conductivity are reported to increase with the increase of total porosity. This is mainly due to macroporosity since it is a preferred path of water flow in the soil profile, especially in response to gravitational force (gravitational potential) (Kurnianto et al. 2018).

A negative correlation was observed between  $K_{Sat}$  and Micro (−0.51). According to Godoy et al. (2019), this would be an inconsistent pattern considering that the  $K_{Sat}$  of the soil has no physical relationship with soil micropores. This result indicates that a simple linear correlation matrix is not enough to explain the correlations between  $K_{Sat}$  and other attributes, and thus, multivariate statistical tools should be considered for the assortment of spatial covariates and then a more efficient cokriging process.

**Table 2** Spearman correlation matrix for the attributes studied

	$K_{Sat}$	BD	TP	Macro	Micro	Elev
BD	−0.50*					
TP	0.50*	−0.99*				
Mac	0.72*	−0.74*	0.74			
Mic	−0.51*	−0.02 <sup>NS</sup>	0.02 <sup>NS</sup>	−0.57*		
Elev	0.05 <sup>NS</sup>	−0.21*	0.21	0.07 <sup>NS</sup>	0.05 <sup>NS</sup>	
Land Use	0.28*	−0.25*	0.24	0.32*	−0.12 <sup>NS</sup>	−0.10 <sup>NS</sup>

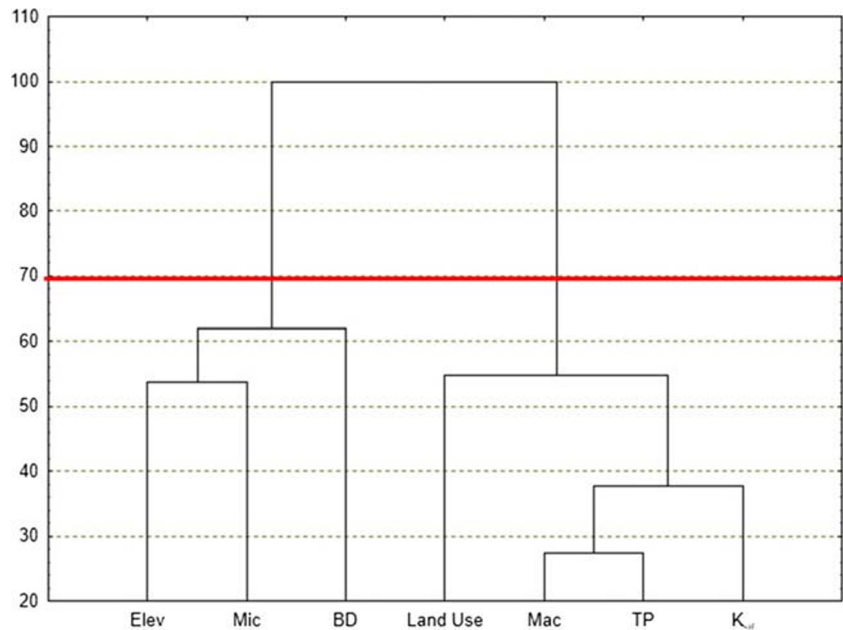
$K_{Sat}$  saturated soil hydraulic conductivity ( $m\ h^{-1}$ ), *Mac* macroporosity (%), *Mic* microporosity (%), *TP* total porosity (%), *BD* bulk density ( $g\ cm^{-3}$ ), *Elev* elevation (m)

\*Significant at 0.01 level significance

\*\*Significant at 0.05 level significance

<sup>NS</sup> No level significance

**Fig. 4** Dendrogram of all hydrology attributes, elevation, and land use, constructed from 179 data points sampled at SEW. The graph also shows the cutoff of 70% (red line)



Based on the dendrogram and the cluster analyses, two clusters were formed: the first was composed by  $K_{Sat}$ , TP, Mac, and Land Use, and the second formed by the BD, Micro, and Elev. In cluster 1, the most pronounced similarity occurred between Mac and TP and then between these two and  $K_{Sat}$ , followed by Land Use. The relationship between Mac and TP is justified by the fact that macroporosity is a straight component of total soil porosity. Soil porosity, mainly macroporosity, is the main water path in the soil profile, and in this way, the approximation of these attributes with  $K_{Sat}$  becomes clear (Godoy et al. 2019). Some studies (Kurnianto et al. 2018; Liu et al. 2018) suggest that  $K_{Sat}$  is affected by land use, mainly in the superficial soil layers. This is due to the physical modifications in the superficial layers by soil management, altering the dynamics of the water in the soil.

### 4.3 Geostatistical Analyses

The variographic analysis (Table 3) demonstrates the spatial dependence of univariate semivariograms fitted using the spherical theoretical model. The cross-semivariograms were also fitted by the spherical theoretical model, except for  $K_{Sat} \times LU$ , which had the experimental semivariogram fitted by the Gaussian theoretical model.

The values of the semivariogram parameter  $[\gamma_{Z_u}^{CH}(h)]$  contribution  $[C]$  and nugget effect  $[C_0 + C]$  are integers and positives, except when BD was used as a covariate ( $K_{Sat} \times BD$ ). The negative values for these parameters of  $[\gamma_{Z_u, Z_v}^{CH}(h)]$   $K_{Sat} \times BD$  occur due to the inversely proportional correlation between the attributes  $K_{Sat}$  and BD, demonstrated by the Spearman matrix. This is widely diffused in the literature,

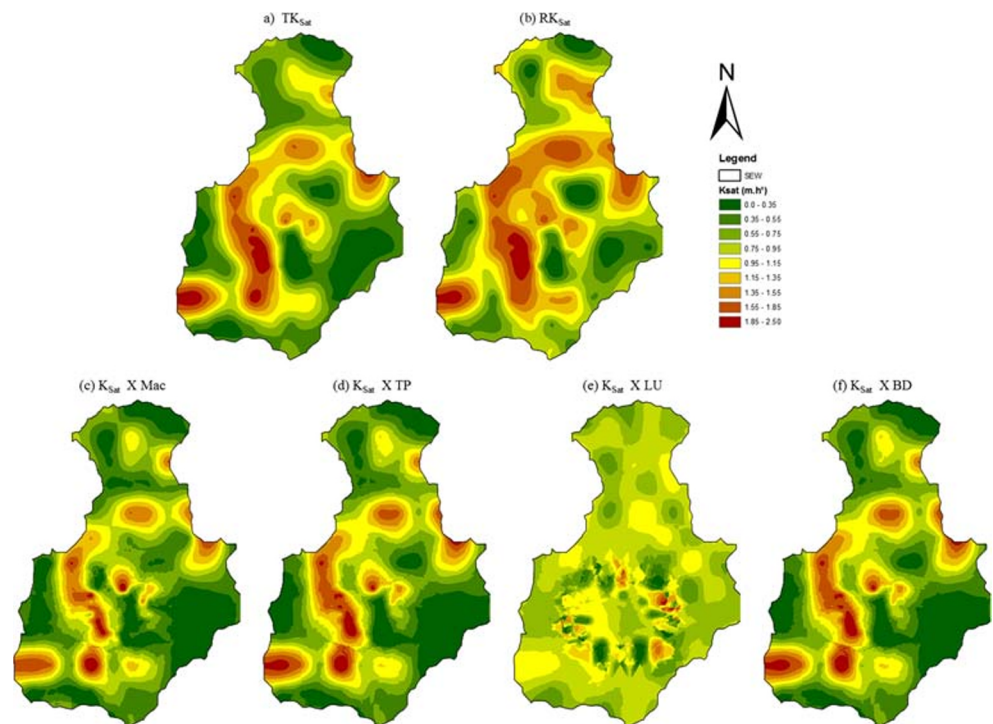
**Table 3** Parameters of univariate semivariograms and cross-semivariograms adjusted by theoretical models

Attributes	Model	$C_0$	$C_0 + C$	$s^2$	$a$ (m)	DSD (%)	Est.
Univariate variogram							
$TK_{Sat}$	Spherical	0.03	0.07	0.06	168.5	43.2	M
$RK_{Sat}$	Spherical	0.36	0.73	0.65	178.84	22.0	CH
Cross-variogram							
$K_{Sat} \times Mac$	Spherical	1.457	4.38	4.05	152.9	33.2	CH
$K_{Sat} \times TP$	Spherical	0.009	0.03	0.02	213.1	34.8	CH
$K_{Sat} \times Land\ use$	Gaussian	0.002	0.11	0.08	161.8	2.1	CH
$K_{Sat} \times BD$	Spherical	-0.024	-0.07	-0.06	212.7	36.2	CH

$K_{Sat}$  saturated soil hydraulic conductivity ( $m\ h^{-1}$ ),  $Mac$  macroporosity (%),  $TP$  total porosity (%),  $BD$  bulk density ( $g\ cm^{-3}$ ),  $C_0$  nugget effect,  $C_0 + C$  sill,  $a$  range,  $s^2$  variance,  $DSD$  degree of spatial dependence  $[C_0/(C_0 + C)] \times 100$ ,  $TK_{Sat}$   $K_{Sat}$  normalized by fourth root  $[K_{Sat}^{(1/4)}]$ ,  $RK_{Sat}$   $K_{Sat}$  by robust estimator,  $Est.$  semivariance estimator,  $M$  Matheron estimator,  $CH$  Cressie and Hawkins estimator



**Fig. 5** Spatial distribution of  $K_{Sat}$  in the Sanga Ellert Watershed using classic ordinary kriging (a), robust ordinary kriging (b), and robust ordinary cokriging (c–f)



where it is assumed that changes in BD mainly affect the volume and the relationship of the macropores of a soil, thus influencing the flow of soil water (Godoy et al. 2019).

Although the highest range of values were observed for  $\gamma_{Z_u, Z_v}^{CH}(h)$  for  $K_{Sat} \times TP$  and  $K_{Sat} \times BD$ , it is not possible to state categorically that cross-semivariograms significantly increase their spatial range when compared with univariate semivariograms, mainly in this case, where for this parameter,  $\gamma_{Z_u}^M(h)$  and  $\gamma_{Z_u}^{CH}(h)$  of  $K_{Sat}$  presented better performance than  $\gamma_{Z_u, Z_v}^{CH}(h)$ , when the covariates were Mac and LU. This may be an indication that for a more realistic assessment of the

efficiency of each method, the analysis of parameters of cross-validation could be essential.

The kriging maps ( $TK_{Sat}$  and  $RK_{Sat}$ ) exhibit distinct patterns (Fig. 5), where it is possible to observe that for  $TK_{Sat}$ , high values occupy a more extensive area of the ECW, whereas for  $RK_{Sat}$ , the highest values are restricted to areas near the drainage network and with a greater forest presence. These results are in accordance with (Kurnianto et al. 2018).

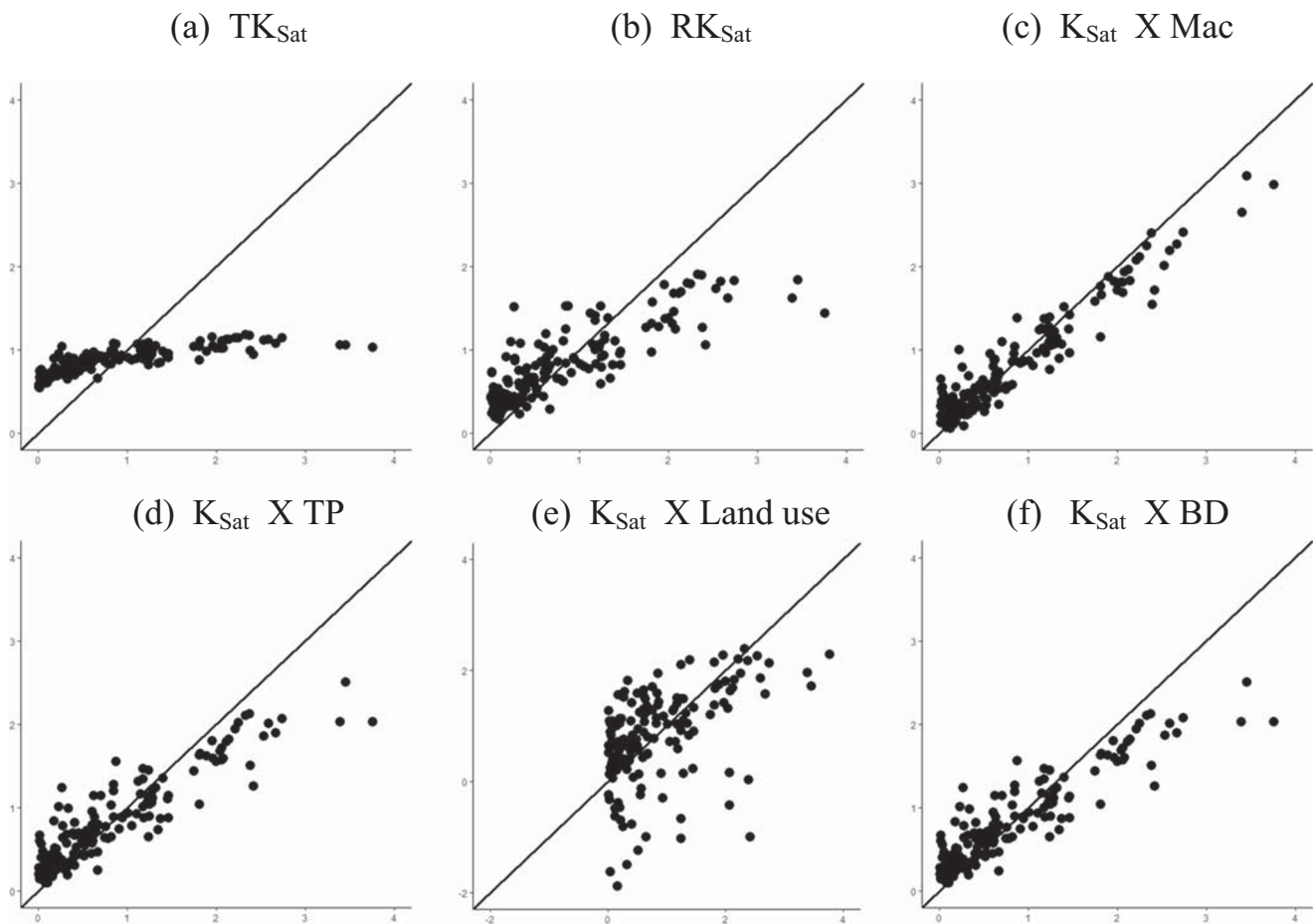
The map generated by the parameters of  $\gamma_{Z_u, Z_v}^{CH}(h)$  called  $K_{Sat} \times LU$  presents a random pattern very different from the others, indicating a more coherent saturated flow accumulation when compared with each other. This phenomenon can be explained by the low performance of the  $\gamma_{Z_u, Z_v}^{CH}(h)$  parameters when using land use as the covariate, indicating that this environmental attribute is not suitable for cokriging interpolation in the ECW.

Maps generated from  $K_{Sat} \times Macro$ ,  $K_{Sat} \times TP$ , and  $K_{Sat} \times BD$  are more detailed than those generated by univariate ordinary kriging. Maps generated from  $K_{Sat} \times Macro$  are close to the original data in relation to the maximum and minimum values and amplitude. Additionally, maps generated by univariate kriging had less spatial detail (more uniform) than those generated by cokriging ( $K_{Sat} \times Mac$ ;  $K_{Sat} \times TP$ ;  $K_{Sat} \times BD$ ). This spatial detail is clearer in some specific areas of the watershed, such as seen in the middle area. A greater amplitude of the high and low values distributed over the ECW area is also observed in the three  $K_{Sat}$  maps which presented better performance in the cokriging process. This suggests that the

**Table 4** Summarized statistics of the cross-validation for observed vs. interpolated data of  $K_{Sat}$

Attributes	RMSE	ME	$r^2$
<b>Kriging</b>			
$TK_{Sat}$	0.6875	0.0285	0.64
$RK_{Sat}$	0.4948	-0.0033	0.69
<b>Cokriging</b>			
$K_{Sat} \times Mac$	0.2618	-0.0044	0.91
$K_{Sat} \times TP$	0.3755	-0.0047	0.83
$K_{Sat} \times Land\ use$	0.8373	-0.0281	0.22
$K_{Sat} \times BD$	0.3744	-0.0054	0.83

$K_{Sat}$  saturated soil hydraulic conductivity ( $m\ h^{-1}$ ),  $Mac$  macroporosity (%),  $TP$  total porosity (%),  $BD$  bulk density ( $g\ cm^{-3}$ ),  $TK_{Sat}$   $K_{Sat}$  normalized by fourth root [ $K_{Sat}^{(1/4)}$ ],  $RK_{Sat}$   $K_{Sat}$  by robust estimator,  $ME$  mean error,  $RMSE$  root mean square error,  $r^2$  coefficient of determination



**Fig. 6** Cross-validation and linear adjustment of observed vs. interpolated data of  $K_{\text{Sat}}$  using classic ordinary kriging (a), robust ordinary kriging (b), and robust ordinary cokriging (c–f)

smoothing effect is more pronounced in kriging than in cokriging. Furthermore, this demonstrates that cokriging, when used the correct covariates, can better describe the variability of  $K_{\text{Sat}}$  than ordinary univariate kriging.

The RMSE values (Table 4), which indicate the data dispersion behavior, presented a better absolute performance of the  $K_{\text{Sat}} \times \text{Mac}$  interpolation, with a value closer to zero. In contrast, the  $K_{\text{Sat}} \times \text{LU}$  cokriging and the univariate interpolations ( $TK_{\text{Sat}}$  and  $RK_{\text{Sat}}$ ) presented the worst results. On the other hand, the values of ME, which identify if the model underestimated (negative values) or overestimated (positive values) the target variable, were quite satisfactory, presenting low and negative values with a slight underestimating behavior, except for the maps generated by univariate kriging  $TK_{\text{Sat}}$  and cokriging “ $K_{\text{Sat}} \times \text{LU}$ .” Only the  $TK_{\text{Sat}}$  interpolation overestimated the observed data, showing the highest ME value.

The coefficient of determination ( $r^2$ ) had a very high performance demonstrating a good fit for almost all of the estimated values in the interpolations. Especially in reference to  $K_{\text{Sat}} \times \text{Mac}$  ( $r^2 = 0.91$ ), this statistic indicated an excellent

linear fit in the cross-validation cokriging data (Fig. 6). The lowest performance can be attributed to cokriging where the covariate of land use was used. In addition to the extremely random standard of the map generated by cokriging, the  $K_{\text{Sat}} \times \text{Land Use}$  cross-validation (Fig. 6e) showed a poor fit between the estimated and observed values. On the other hand, the results obtained by  $K_{\text{Sat}} \times \text{Mac}$  (Fig. 6c) during cokriging proved to be the most efficient mapping method, with a high performance in statistical indicators and cross-validation.

Saturated soil hydraulic conductivity is the most important soil hydrology attribute for soil management, water yield and baseflow behavior in watersheds, sediment transport, and for the basic information of initial hydrogeology exploration. Several applications require a good map of  $K_{\text{Sat}}$ , such as for identifying areas with higher potential for groundwater recharge (Pinto et al. 2016; Alvarenga et al. 2012), to characterize the effect of land use on the watershed streamflow behavior by means of a distributed hydrological model (Alvarenga et al. 2016) or based on measurements in the field (Price 2012; Muñoz-Villers and McDonnell 2013; Pinto et al. 2019), to understand the water balance in watersheds (Mello et al.

2019; Salemi et al. 2013; Fleischbein et al. 2006) and for use as a support for storm hydraulic structure design (Chappell et al. 2017).

## 5 Conclusions

Among univariate geostatistical techniques, the use of the robust Cressie and Hawkins estimator on untransformed data presented a better performance than using the classical estimator of Matheron on transformed data. The cluster multivariate method presented a clearer separation of the covariates for the  $K_{\text{Sat}}$  cokriging process, which can be considered as the most efficient method for selection of covariates compared to classic techniques.

The cokriging technique demonstrated a greater performance than univariate techniques when robust cokriging was applied using macroporosity and bulk density as covariates, demonstrated by the quality of the  $K_{\text{Sat}}$  map increased estimates. Land use covariate, although showing an influence on  $K_{\text{Sat}}$ , was not successful in subsequent geostatistical multivariate processes for the studied watershed.

The  $K_{\text{Sat}}$  robust cokriging map can be used to improve land planning in this headwater watershed, with the aim of increasing the water infiltrability and preserving the uses and management strategies that have shown a positive impact on the hydrology of the watershed.

**Acknowledgments** The authors wish to thank Brazilian National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel, Brazil (CAPES), Finance Code 001, for scholarships.

**Funding Information** The study is financially supported by Brazilian National Council for Scientific and Technological Development (CNPq).

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

- Adhikary SK, Muttill N, Yilmaz AG (2017) Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments. *Hydrol Process* 31:2143–2161
- Ahuja L, Ma L, Green T (2010) Effective soil properties of heterogeneous areas for modeling infiltration and redistribution. *Soil Sci Soc Am J* 74:1469–1482
- Alvarenga CC, Mello CR, Mello JM, Silva AM, Curi N (2012) Índice de qualidade do solo associado à recarga de água subterrânea (IQS RA) na Bacia Hidrográfica do Alto Rio Grande, MG. *Rev Bras Ci Solo* 36:1608–1619
- Alvarenga LA, Mello CR, Colombo A, Cuartas LA, Bowling LC (2016) Assessment of land cover change on the hydrology of a Brazilian headwater watershed using the distributed hydrology-soil-vegetation model. *Catena* 143:7–17
- Alvarez R, Steinbach HS (2009) A review of the effects of tillage systems on some soil physical properties, water content, nitrate availability and crops yield in the Argentine Pampas. *Soil Tillage Res* 104:1–15
- Archer N, Bonell M, Coles N, Macdonald AM, Auton CA, Stevenson R (2013) Soil characteristics and landcover relationships on soil hydraulic conductivity at a hillslope scale: a view towards local flood management. *J Hydrol* 497:208–222
- Baiamonte G, Bagarello V, D’Asaro F, Palmeri V (2017) Factors influencing point measurement of near-surface saturated soil hydraulic conductivity in a small Sicilian basin. *Land Degrad Dev* 28:970–982
- Becker R, Gebremichael M, Märker M (2018) Impact of soil surface and subsurface properties on soil saturated hydraulic conductivity in the semi-arid Walnut Gulch Experimental Watershed, Arizona, USA. *Geoderma* 322:112–120
- Beskow S, Timm LC, Tavares VEQ, Caldeira TL, Aquino LS (2016) Potential of the LASH model for water resources management in data-scarce basins: A case study of the Fragata River basin, Southern Brazil. *Hydrol Sci J* 61:2567–2578
- Bitencourt DGB, Barros W, Timm LC, She D, Penning LH, Parfitt JMB, Reichardt K (2016) Multivariate and geostatistical analyses to evaluate lowland soil levelling effects on physico-chemical properties. *Soil Tillage Res* 156:63–73
- Blake GR, Hartge KH (1986) Bulk density. In: Klute A (ed) *Methods of soil analysis*. Part 1, 2nd edn. Agronomy Monograph, ASA-SSSA, Madison, pp 363–382
- Boadu FK (2000) Hydraulic conductivity of soils from grain size distribution: new models. *JGGE* 126:739–746
- Cambardella CA, Moorman TB, Novack JM, Parkin TB, Karlen DL, Turco RF, Knopka AE (1994) Field-scale variability of soil properties in Central Iowa soils. *Soil Sci Soc Am J* 58:1240–1248
- Chappell NA, Jones TD, Tych W, Krishnaswamy J (2017) Role of rain-storm intensity underestimated by data-derived flood models: emerging global evidence from subsurface-dominated watersheds. *Environ Modelling Soft* 88:1–9
- Cressie N, Hawkins DM (1980) Robust estimation of the variogram: I. *Math Geol* 12:115–125
- Elsenbeer H, Cassel K, Castro J (1992) Spatial analysis of soil hydraulic conductivity in a tropical rain forest catchment. *Water Resour Res* 28:3201–3214
- Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) (1999) *Manual de métodos de análises de solo*. Embrapa Solos, Rio de Janeiro
- Empresa Brasileira De Pesquisa Agropecuária (EMBRAPA) (2006) *Sistema brasileiro de classificação de solos*. Embrapa Solos, Rio de Janeiro
- Everitt B, Hothorn T (2011) *An introduction to applied multivariate analysis with R*. Springer, New York
- Fleischbein K, Wilcke W, Valarezo C, Zech W, Knoblich K (2006) Water budgets of three small catchments under montane forest in Ecuador: experimental and modeling approach. *Hydrol Process* 20:2491–2507
- Genton MG (1998) Highly robust variogram estimation. *Math Geol* 30:213–221
- Germer S, Neill C, Krusche AV, Elsenbeer H (2010) Influence of land-use change on near-surface hydrological processes: Undisturbed forest to pasture. *J Hydrol* 380:473–480
- Godoy VA, Zuquette LV, Gómez-Hernández JJ (2018) Stochastic analysis of three-dimensional hydraulic conductivity upscaling in a heterogeneous tropical soil. *Comput Geosci* 100:17s4–17s1187
- Godoy VA, Zuquette LV, Gómez-Hernández JJ (2019) Spatial variability of hydraulic conductivity and solute transport parameters and their spatial correlations to soil properties. *Geoderma* 339:59–69

- Goovaerts P (1997) Geostatistics for natural resources evaluation. Applied Geostatistics Series. Oxford University Press, Oxford
- Hu W, Dongli S, Ming'an S, Kwok PC, Bingcheng S (2015) Effects of initial soil water content and saturated hydraulic conductivity variability on small watershed runoff simulation using LISEM. *Hydrol Sci J* 60:1137–1154
- Instituto Brasileiro de Geografia e Estatística - IBGE (1986) Geologia, geomorfologia, pedologia, vegetação, uso potencial da terra. IBGE, Rio de Janeiro
- Jenny H (1941) Factors of soil formation: a system of quantitative pedology. Dover Publications, New York
- Johnston PR, Kilpatrick D, Li CY (2001) The importance of anisotropy in modeling ST segment shift in subendocardial ischaemia. *IEEE T Bio-Med Eng* 48:1366–1376
- Klute A (1986) Water retention: laboratory methods. In: Klute A (ed) *Methods of soil analysis. Part 1, 2nd edn.* Agronomy Monograph, ASA-SSSA, Madison, pp. 635–662
- Klute A, Dirksen C (1986) Hydraulic conductivity and diffusivity: laboratory methods. In: Klute A (ed) *Methods of Soil Analysis. Part 1, 2nd edn.* Agronomy Monograph, ASA-SSSA, Madison, pp. 687–734
- Kuinchtner A, Buriol GA (2001) Clima do estado do Rio Grande do Sul segundo a classificação climática de Köppen e Thornthwaite. *Disciplinarum Scientia* 2:171–182
- Kurnianto S, Selker J, Kauffman JB, Murdiyarto D, James PJP (2018) The influence of land-cover changes on the variability of saturated hydraulic conductivity in tropical peatlands. *Mitig Adapt Strat Gl* 24:535–555
- Lark RM (2000) A comparison of some robust estimators of the variogram for use in soil survey. *Eur J Soil Sci* 51:137–157
- Lark RM (2003) Two robust estimators of the cross-variogram for multivariate geostatistical analysis of soil properties. *Eur J Soil Sci* 54:187–202
- Lebrez H, Bárdossy A (2017) Estimation of the variogram using Kendall's tau for a robust geostatistical interpolation. *J Hydrol Eng* 22:38–46
- Libohova Z, Schoeneberger P, Bowling LC, Owens PR, Wysocki D, Wills S, Williams CO, Seybold C (2018) Soil systems for upscaling saturated hydraulic conductivity for hydrological modeling in the critical zone. *Vadose Zone J* 17:170051
- Lin H (2006) Temporal stability of soil moisture spatial pattern and subsurface preferential flow pathways in the Shale Hills catchment. *Vadose Zone J* 5:317–340
- Liu Z, Ma D, Hu W, Li X (2018) Land use dependent variation of soil water infiltration characteristics and their scale-specific controls. *Soil Tillage Res* 178:139–149
- Ma Y, Li X, Guo L, Lin HS (2017) Hydopedology: interactions between pedologic and hydrologic processes across spatiotemporal scales. *Earth Sci Rev* 171:181–195
- Mello CR, Ávila LF, Lin H, Terra MCNS, Chappell NA (2019) Water balance in a neotropical forest catchment of southeastern Brazil. *Catena* 173:9–21
- Mojena R (1977) Hierarchical grouping methods and stopping rules: an evaluation. *Comput J* 20:359–363
- Muñoz-Villers LE, McDonnell JJ (2013) Land use change effects on runoff generation in a humid tropical montane cloud forest region. *Hydrol Earth Syst Sci* 17:3543–3560
- Papanicolaou AN, Elhakeem M, Wilson CG, Burras CL, West LT, Lin H, Clark B, Oneal BE (2015) Spatial variability of saturated hydraulic conductivity at the hillslope scale: Understanding the role of land management and erosional effect. *Geoderma* 243–244:58–68
- Picciafuoco T, Morbidelli R, Flammini A, Saltalippi C, Corradini C, Strauss P, Blöschl G (2019) A pedotransfer function for field-scale saturated hydraulic conductivity of a small watershed. *Vadose Zone J* 12:1–15
- Pinto LC, Mello CR, Owens PR, Norton LD, Curi N (2015) Role of inceptisols in the hydrology of mountainous catchments in south-eastern Brazil. *J Hydrol Eng* 21:e05015017
- Pinto LC, Mello CR, Norton LD, Owens PR, Curi N (2016) Spatial prediction of soil-water transmissivity based on fuzzy logic in a Brazilian headwater watershed. *Catena* 143:26–34
- Pinto LC, Mello CR, Norton LD, Curi N (2019) Land-use influence on the soil hydrology: an approach in upper Grande River basin, Southeast Brazil. *Ciênc Agrotec* 43:e015619
- Price K, Jackson CR, Parker AJ (2010) Variation of surficial soil hydraulic properties across land uses in the southern Blue Ridge Mountains, North Carolina, USA. *J Hydrol* 383:256–268
- Price JN, Hiiesalu I, Gerhold P, Pärtel M (2012) Small-scale grassland assembly patterns differ above and below the soil surface. *Ecology* 93:1290–1296
- Qiao JB, Zhu YJ, Jia X, Huang LM, Shao MA (2018) Estimating the spatial relationships between soil hydraulic properties and soil physical properties in the critical zone (0–100 m) on the Loess Plateau, China: a state-space modeling approach. *Catena* 160:385–393
- Ramos MC, Cots-Folch R, Martínez-Casasnovas JA (2007) Effects of land terracing on soil properties in the Priorat region in Northeastern Spain: a multivariate analysis. *Geoderma* 142:251–261
- Reichardt K, Timm LC (2020) *Soil, plant and atmosphere: concepts, processes and applications.* Springer, Basel
- Salemi LF, Groppo JD, Trevisan R, Moraes JM, Ferraz SFB, Villani JP, Duarte-Neto PJ, Martinelli LA (2013) Land-use change in the Atlantic rainforest region: consequences for the hydrology of small catchments. *J Hydrol* 499:100–109
- She D, Dongdong L, Yingying L, Yi L, Cuilan Q, Fang C (2014) Profile characteristics of temporal stability of soil water storage in two land uses. *Arab J Geosci* 7: 21–34
- She D, Qian C, Timm LC, Beskow S, Wei H, Caldeira TL, Oliveira LM (2017) Multi-scale correlations between soil hydraulic properties and associated factors along a Brazilian watershed transect. *Geoderma* 286:15–24
- Simonson RW (1959) Outline of a generalized theory of soil genesis I. *Soil Sci Soc Am J* 23:152–156
- Song X, Chen X, Ye M, Dai X, Hammond G, Zachara JM (2019) Delineating facies spatial distribution by integrating ensemble data assimilation and indicator Geostatistics with level-set transformation. *Water Resour Res* 55:2652–2671
- Wallin J, Bolin D (2015) Geostatistical modelling using non-gaussian matérn fields. *Scand J Stat* 42:872–890
- Wang Z, Shi W (2018) Robust variogram estimation combined with isometric log-ratio transformation for improved accuracy of soil particle-size fraction mapping. *Geoderma* 324:56–66
- Wang Y, Shao M, Liu Z, Horton R (2013) Regional-scale variation and distribution patterns of soil saturated hydraulic conductivities in surface and subsurface layers in the loessial soils of China. *J Hydrol* 487:13–23
- Webster R, Oliver M (2007) *Geostatistics for environmental scientists,* second edn. Wiley, Chichester
- Wilding LP, Drees LR (1983) *Spatial variability and pedology.* In: Wilding LP, Drees LR (eds) *Pedogenesis and soil taxonomy: concepts and interactions.* Elsevier, New York, pp 83–116
- Wood MD, King NE (1977) Relation between earthquakes, weather, and soil tilt. *Science* 197:154–156

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.