

Spatial prediction of soil infiltration using functional geostatistics

Diego Leonardo Cortes-D, Jesus Hernan Camacho-Tamayo*, Ramon Giraldo

Universidad Nacional de Colombia, Department of Civil Engineering and Agricultural, Colombia

* Corresponding author: jhcamachot@unal.edu.co

ABSTRACT

The infiltration of water into the soil is a necessary parameter for irrigation systems design. Characterizing its spatial behavior allows a site-specific management of water according to soil conditions and crop requirements. The aim of this study is to establish the spatial distribution of infiltration in an Andisol by means of two geostatistical approaches: on the one hand by means of functional kriging, taking as input infiltration curves (obtained after a smoothing stage), and on the other hand by using classical ordinary kriging on the parameters of the Kostiakov and Phillip models. The comparison between these methodologies is carried out taking as a criterion the sum of squared errors of a leave-one-out cross-validation analysis. The results show a high correlation between observations and predictions (R^2 values around 99%), which indicates that the use of functional geostatistics in this context could be a good alternative. Moreover, from a descriptive point of view, we can point out that the contour maps of basic infiltration (BI), cumulative infiltration (Ci), saturated hydraulic conductivity (Ks), and sorptivity (S) obtained with the observed data, as well as the predictions by functional geostatistics, show a very similar behavior, which empirically validates the use of this methodology.

KEYWORDS

functional data; functional kriging; Kostiakov; Philip; spatial variability

Received: 29 October 2016

Accepted: 19 June 2018

Published online: 27 August 2018

Cortes-D, D. L., Camacho-Tamayo, J.-H., Giraldo, R. (2018): Spatial prediction of soil infiltration using functional geostatistics. *AUC Geographica* 53(2), 149–155
<https://doi.org/10.14712/23361980.2018.15>

© 2018 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>).

1. Introduction

Soil management does not usually consider the particularities of the terrain or its physical properties; thus it is important to perform an evaluation based on the concepts of sustainability (Martins et al. 2010). In efficient and sustainable agricultural production systems, it is important to understand the behavior of water within the soil in order to be able to use this resource reasonably.

Infiltration is a hydrodynamic attribute that concerns the movement of water and that is closely related to the capillarity processes and forces associated with the adhesion and cohesion of soil particles (Orjuela-Matta et al. 2010). Moreover, infiltration is a basic parameter that must be accounted for when designing and implementing irrigation systems (Chowdary et al. 2006; Machiwal et al. 2006). Other attributes associated with soil water movement include hydraulic conductivity and sorptivity (S). Hydraulic conductivity is associated with the resistance of soil pores relative to the soil itself, is used to solve drainage and soil conservation problems, and depends not only on soil properties but also on the soil water content (Jačka et al. 2016). Sorptivity explains the movement of water in the soil as an effect of its matric potential, since S is estimated as a measure of the ability of a porous medium to absorb or desorb fluid by capillarity (Moret-Fernández et al. 2017).

Understanding soil behavior and its rate of infiltration over time could help control soil erosion or other degradation processes, establish the availability of water for crops and the movement of substances in the soil, and be used to create strategies for watershed management (Orjuela-Matta et al. 2010).

In general, the study of soil infiltration begins with experimental trials in a finite group of sites within an area of interest. Next, the experimental data are fit to theoretical models, such as those of Kostiakov, Lewis, Horton, or Philip, among others, and the empirical parameters of the models are determined. This process concludes with a calculation of the univariate descriptive measures (localization and dispersion), obtaining distribution graphs (histograms and box plots), conducting multivariate analyses (correlation, classification, and principle components), and performing univariate geostatistical analyses (variogram estimation, kriging prediction, and creation of maps). The spatial behavior can be described (Camacho-Tamayo et al. 2013) and the management zones can be established for agricultural production by using these tools (Cucunubá-Melo et al. 2011).

Geostatistics (Isaaks, Srivastava 1989) is a field of statistics which is used in practice for predicting a variable for non-sampled locations of a region. Specifically, several variations of kriging and cokriging methods (Cressie 1993) are used to fulfill this task. These tools are considered when for each sampled

site we have data for one or several properties (for example, hydraulic conductivity, or sorptivity). However, in many practical situations a large number of measurements are recorded continuously (depth, time, wavelength, etc.). In these cases, instead of univariate or multivariate data, we have (after a smoothing step) curves or functions (functional data). The extension of the classical univariate and multivariate geostatistical methods to the context of functional data has given rise a new field known as functional geostatistics (Giraldo et al. 2010). In particular, kriging and cokriging methods have been extended to in order to deal with functional data (Giraldo et al. 2009; Giraldo et al. 2010). These tools allow predicting curves rather than data of one or several variables. The application of functional geostatistics starts with the application of non-parametric smoothing techniques (e.g., kernel or B-spline regression) to convert the discrete data at each sampling site (i.e. infiltration) into continuous functions. After that, functional kriging (Giraldo et al. 2010) can be used for predicting curves (instead of scalar values) at sites for which information is lacking.

In this study, functional geostatistics is applied to infiltration curves in order to evaluate their predictive capacity. A cross-validation analysis is performed to estimate prediction errors and compare them with those obtained using standard univariate kriging methods that are used to analyze infiltration data (Giraldo et al. 2010). The infiltration curve of each site (obtained after smoothing the discrete data) is temporally removed and predicted by using functional kriging and classical ordinary kriging (in this case, the parameters of the Kostiakov and Philip models are predicted, and then they are used to construct the predicted curve).

This paper is organized as follows: Section 2 gives an overview of the data analyzed, the Kostiakov and Phillip models, and the theory of functional geostatistics. Then in Section 3, a comparison between functional and classical univariate kriging methods (based on infiltration data) is carried out.

2. Material and methods

2.1 Description of study site

This study was performed at the Marengo Agricultural Center (Centro Agropecuario Marengo, CAM by its initials in Spanish) located in the municipality of Mosquera (Cundinamarca, Colombia) at 4°42' N and 74°12' W and at an altitude of 2,543 m. The mean annual temperature in this zone is 13.1 °C, and the mean annual rainfall is 665 mm (weather records of the National University of Colombia). The area described, according to temperature and precipitation characteristics, is classified according to Holdrige as a life zone of low montane dry forest (bs-MB) and the climate as cold dry (FS).

The soils of the CAM have been formed from three types of parent materials, namely lacustrine clays, volcanic ash, and alluvial sediments. The type of relief corresponds to terraces and floodplains, specifically terrace plains and overflow plains (Bolívar, Ordóñez 2014).

2.2 In-situ infiltration measurement

An evenly spaced, rigid grid with 75 recording sites was created over four plots with mainly kikuyu grass (*Pennisetum clandestinum*). These plots covered an area of 16 ha and were located using a GPS Garmin Etrex 20 receiver in real time via satellite. At each site, a double-ring infiltrometer was used to measure water infiltration over a period of 150 minutes (at 1, 2, 3, 4, 5, 10, 15, 30, 45, 60, 90, 120, and 150 minutes) using variable loads and while ensuring that the difference between the maximum and minimum reading was never greater than 100 mm.

2.3 Determination of infiltration characteristics

The results obtained from the field were fit to two theoretical infiltration models, the Kostiakov (Equation 1) and the Philip (Equation 4) models. The parameters for these models were estimated by using the R software (R Development Core Team 2011) and then were used to calculate S , K_s , B_p , and C_i and to carry out conventional statistical and geostatistical analyses.

Kostiakov model

$$I(t) = a \cdot t^b \quad (1)$$

where

$I(t)$ is the cumulative infiltration content (cm) at time t and a and b are empirically-derived fit coefficients.

$$-600(b - 1) = t_{base} \quad (2)$$

$$B_i = 60 \cdot a \cdot b \cdot t_{base}^{b-1} \quad (3)$$

with B_i as the basic infiltration (cm h⁻¹).

Philip model

$$I(t) = S \cdot t^{1/2} + K_s \cdot t \quad (4)$$

where S depends on the initial soil moisture conditions (cm h^{-1/2}) and K_s is the saturated hydraulic conductivity of the soil (cm h⁻¹).

2.4 Functional geostatistics

Functional geostatistics (Giraldo et al. 2010) offers methods for spatial prediction of curves. Usually, the discrete data (for example infiltration records) at each sampling site are previously recorded in continuous curves after smoothing them by using basis functions (Ramsay, Silverman 2005). Afterward, they

are used for predicting a whole curve at unsampled sites. A cross-validation analysis was carried out. Each of the 75 curves was temporarily removed and predicted based on the remaining 74 by using functional kriging (Giraldo 2009). Thus observed and predicted curves (by using functional kriging) were obtained for each of the 75 sites. The comparison between them (cross-validation) was used for evaluating the quality of this approach and also to compare with the results obtained by means of classical geostatistics.

The functional kriging predictor is defined as follows (Giraldo 2009):

$$\hat{\chi}(s_0) = \sum_{i=1}^n \lambda_i \chi(s_i); \lambda_1, \dots, \lambda_n \in \mathbb{R} \quad (5)$$

where $\hat{\chi}(s_0)$ is the predicted curve at site s_0 , $\chi(s_i)$ corresponds to the curve observed at site s_i , and $i = 1, 2, \dots, n$ and $\lambda_i = 1, \dots, n$, are the weighting parameters that indicate the contributions of each observed curve to the predicted curve.

The λ_i parameters are estimated by solving the following system of equations (Giraldo et al. 2010):

$$\begin{pmatrix} \int_T \gamma_t (\|s_1 - s_1\|) dt & \dots & \int_T \gamma_t (\|s_1 - s_n\|) dt & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \int_T \gamma_t (\|s_n - s_1\|) dt & \dots & \int_T \gamma_t (\|s_n - s_n\|) dt & 1 \\ 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ -\mu \end{pmatrix} \quad (6)$$

where the integrals correspond to the trace-variogram function (Giraldo et al. 2010) evaluated for both the distances between observation sites (left matrix) and the distances between the observation site and the prediction site (right vector). The trace-variogram function is estimated by the method of moments (Giraldo et al. 2010):

$$\hat{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{i,j \in N(h)} \int_T (\chi_t(s_i) - \chi_t(s_j))^2 dt \quad (7)$$

where $\chi_t(s_i)$ is the curve at site i and $N(h) = \{(s_i, s_j): \|s_i - s_j\| = h\}$ is the number of pairs of sites separated by a distance h . Once the trace-variogram function for a sequence of K values is estimated, a parametric semivariance model (spherical, Gaussian, exponential, etc.) is fit to the scatterplot in order to make an estimation at any possible distance between sites.

The results achieved with functional geostatistics were compared with those obtained by using conventional geostatistical methods (estimating the semivariogram and using classical ordinary kriging). For this process, the variables B_p , C_p , S , and K_s from the theoretical infiltration models were considered. For each

case (each variable), the semivariance function was estimated (Cressie 1993):

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{n(h)} [z(x_i) - z(x_i+h)]^2 \quad (8)$$

where $z(x_i)$ is the value of each variable at a site i , $z(x_i + h)$ is the value of the same variable at a point with a distance h away from the previous point, and $N(h)$ is the number of pairs of data separated by the distance h . This empirical semivariogram is fit using one of the above-mentioned theoretical semivariance models. Various criteria are considered for selecting the best model, including the coefficient of determination (R^2), the least sum of squared errors (LSSE), and the cross-validation correlation (CVC). The shared parameters among the theoretical semivariance models include the nugget (C_0), which is a discontinuity in the semivariogram at the origin, the variance of the process (C), and the reach (r), which is the distance until there is a spatial correlation. The nugget-variance ratio, $C/(C_0 + C)$, is also often used as a criterion for model selection. This parameter establishes the degree of spatial dependence (DSD) expressed by the studied attribute. Cambardella (1994) states that if the DSD is greater than 75% the dependence is strong, if the DSD is between 25% and 75% the dependence is moderate, and if the DSD is less than 25% the dependence is weak.

3. Results and discussion

In Figure 2, we can see the fit of the data to a decreasing potential curve, which describes the infiltration rate. The strong change in the first moments of the curve confirms that the soil was in a state of water deficit at the time of the test. Given the trend of the curves, the results suggest that at some points the soil did not reach a condition close to saturation, after being subjected to a constant application of water, for a period of 150 minutes, especially for those sites that they registered a high infiltration rate. This behavior was observed for all the infiltration tests performed and reported by different authors (Machiwal et al. 2006; Orjuela-Matta et al. 2010; Latorre et al. 2015).

To provide an example of a functional geostatistical analysis of collected data, a curve was predicted for a site located at coordinates 984475 (W) and 1009685 (N) (Figure 1), as shown in Figure 2. To perform this prediction, the trace-semivariogram function was estimated using Equation 7. By using this function and Equation 6, the λ_i parameters for the functional kriging predictor, as defined in Equation 5, were estimated. Additionally, based on the predicted curve, the parameters for the theoretical Kostiakov and Philip models were estimated (Table 1).

The predicted curve is consistent with the behavior patterns of the observations. After 60 minutes,

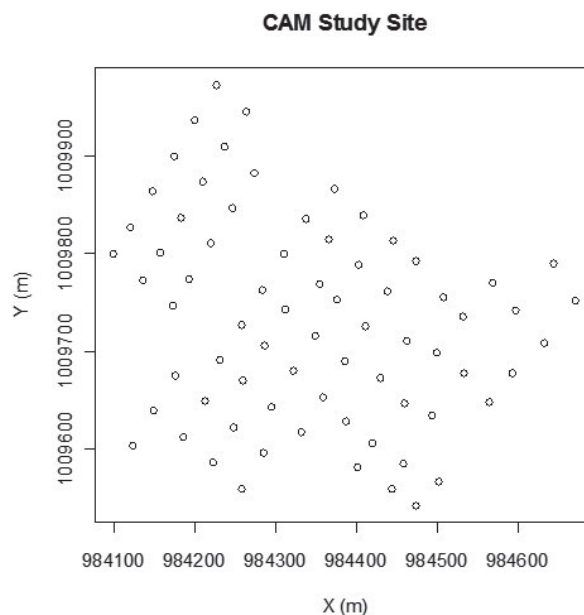


Fig. 1 Distribution of recording sites and identification of random sites where the cumulative infiltration (C_i) curve was predicted.

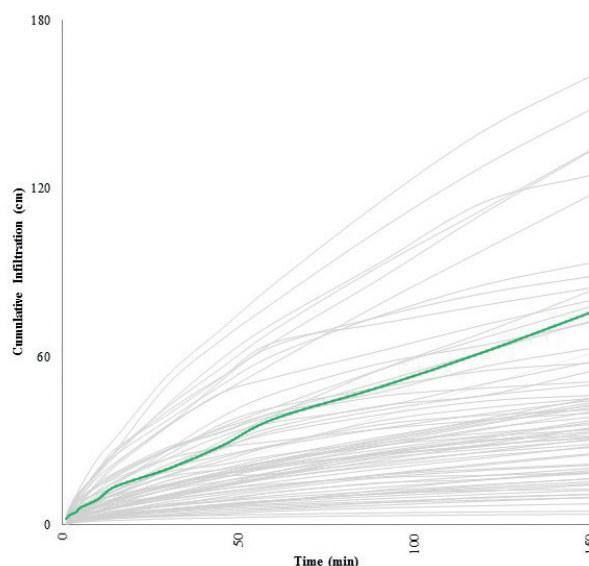


Fig. 2 Cumulative infiltration curve for the site at coordinates of 984475 (W) and 1009685 (N).

Tab. 1 Parameters estimated for Kostiakov and Philip models for the location with coordinates 984475 (W) and 1009685 (N).

Parameter	Value
a	1.53
b	0.78
B_i (cm h^{-1})	23.66
C_i (cm)	75.64
S ($\text{cm h}^{-1/2}$)	15.25
K_s (cm h^{-1})	20.48

the water entry rate into the soil, or infiltration velocity, became constant, due to the increased matric potential of the soil. According to the observations

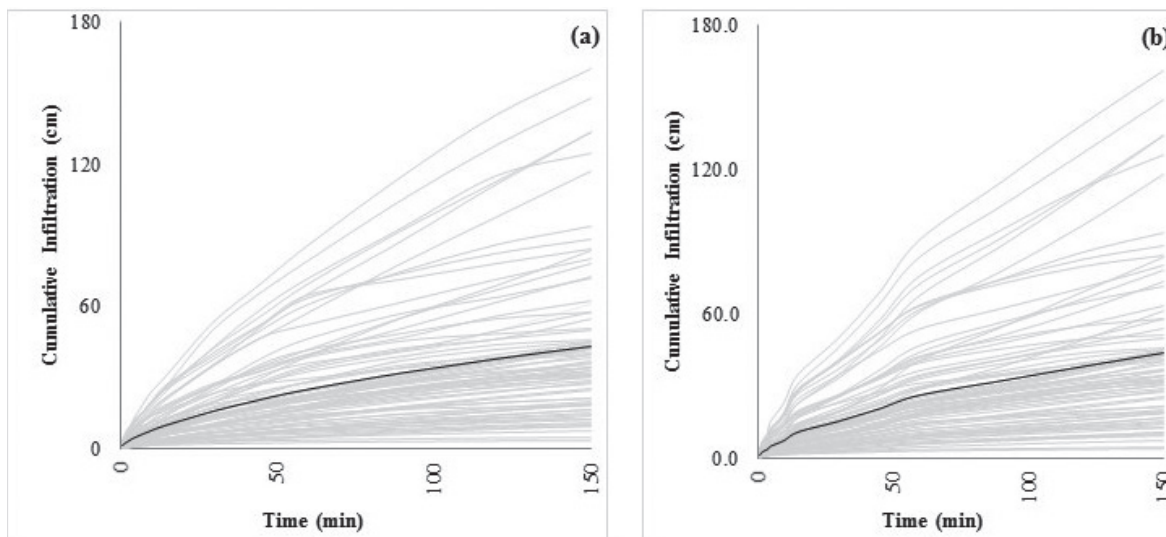


Fig. 3 (a) Cumulative infiltration curves recorded at the 75 sites. (b) Predicted curves at each site using functional kriging.

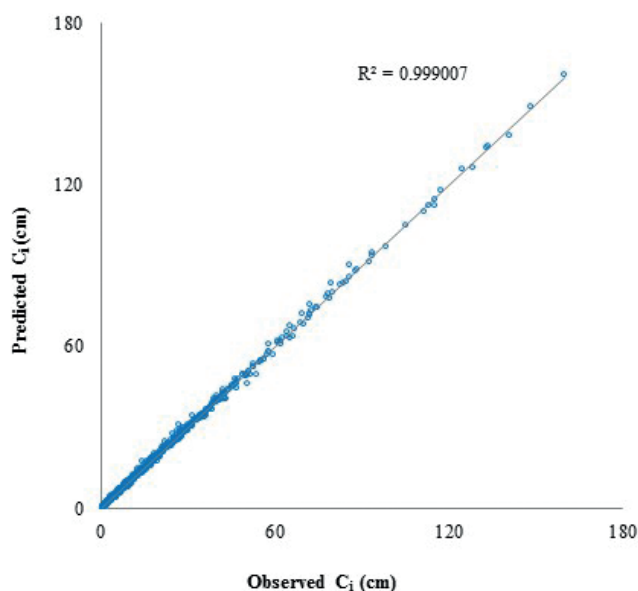


Fig. 4 Functional cross-validation of the amount of infiltration (C_i).

of Montenegro and Malagón (1990), the infiltration velocity at the unsampled site was high ($12.7\text{--}25.4\text{ cm h}^{-1}$), potentially due to a greater quantity of micropores (Richmon, Rillo 2006) relative to the sites with lower infiltration, which resulted in a high K_s value.

The recorded cumulative infiltration curves and the predictions at each site are shown in Figure 3. As an indicator of goodness-of-fit, a simple linear regression (Figure 4) was estimated to compare observed and predicted values. This plot shows a high correlation between them. The R^2 (around 99%) confirms that the method used allows obtaining good predictions.

We compared observed and predicted data of C_i , B_i , K_s and S , respectively, by means of descriptive measures (Table 2), estimated parameters of the semivariogram models (Table 3), and maps of spatial distribution (Figure 5).

Tab. 2 Descriptive measures for the Kostiakov and Philip parameters estimated from the observed (Obs) and predicted (Pre) data.

Model	Parameter	Mean	Median	CV (%)	Min	Max
Kostiakov	a_{Obs}	2.09	1.59	70.87	0.34	6.44
	a_{Pre}	2.09	1.60	70.82	0.34	6.43
	b_{Obs}	0.61	0.60	16.08	0.36	0.86
	b_{Pre}	0.61	0.60	16.08	0.36	0.86
	B_{iObs}	10.65	7.51	97.78	0.33	45.16
	B_{iPre}	10.62	7.49	97.78	0.33	45.09
Philip	C_{iObs}	45.08	35.35	77.82	3.90	160.00
	C_{iPre}	45.19	35.35	77.96	3.91	160.97
	S_{Obs}	17.50	14.01	67.29	3.37	53.46
	S_{Pre}	18.03	14.25	67.12	3.40	53.44
Philip	K_{sObs}	5.14	3.59	138.19	-7.65	36.16
	K_{sPre}	5.12	3.58	138.40	-7.68	36.02

In both cases (Kostiakov and Phillip models), the descriptive values in Table 2 are very similar, indicating a good performance of the functional kriging results.

Tab. 3 Parameters of the theoretical semivariogram model fit by observed (Obs) and predicted (Pre) basic infiltration values (B_i), cumulative infiltration (C_i), sorptivity (S), and saturated hydraulic conductivity (K_s).

Parameter	Model	C_0	$C_0 + C$	Range (m)	CVC	DSD
B_{iObs}	Spherical	35.5	103.8	247.9	0.90	0.66
B_{iPre}	Spherical	35.3	103.8	248.0	0.90	0.66
C_{iObs}	Exponential	393	1122	290.7	0.91	0.65
C_{iPre}	Exponential	404	1133	294.3	0.91	0.64
S_{Obs}	Exponential	0.10	124.9	94.20	0.79	1.00
S_{Pre}	Exponential	0.10	131.2	100.5	0.78	1.00
K_{sObs}	Exponential	0.10	53.20	268.5	0.82	0.99
K_{sPre}	Exponential	0.10	52.90	267.6	0.82	0.99

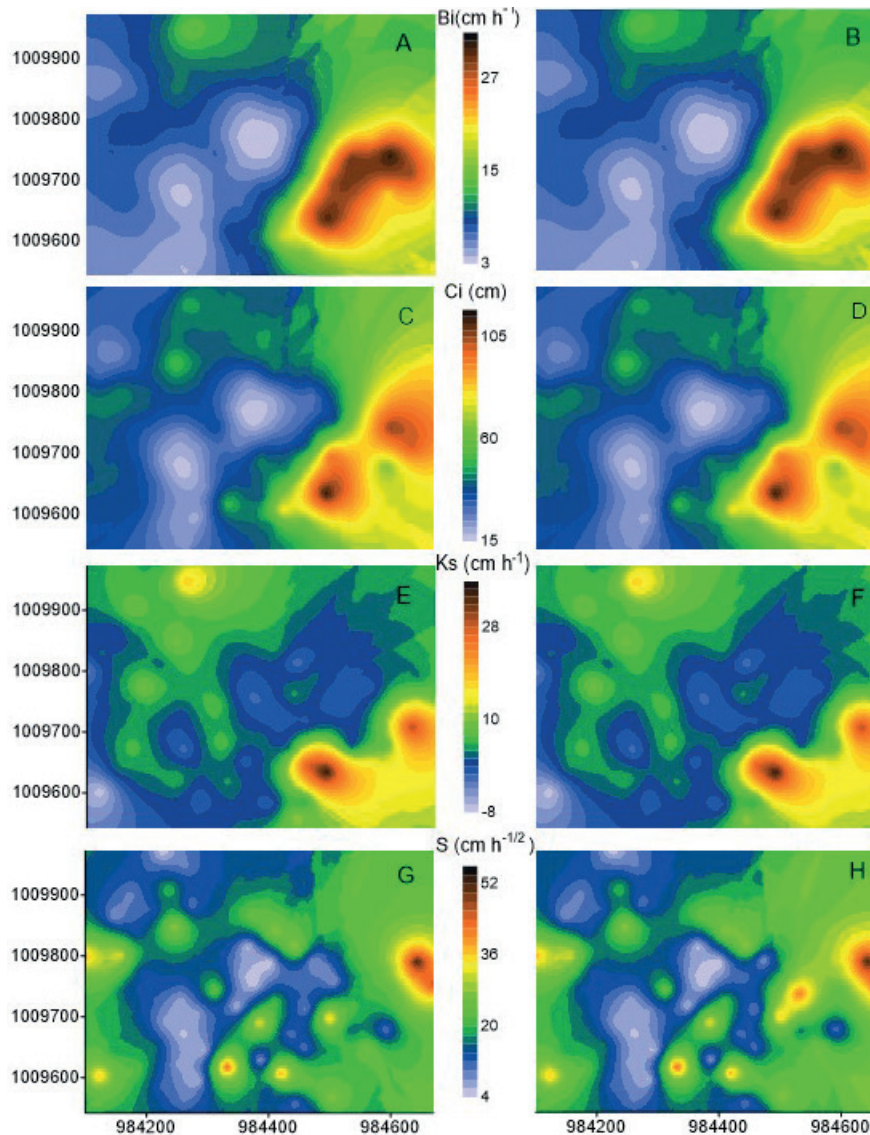


Fig. 5 Contour maps for observed (left) and predicted (right) data for (A–B) basic infiltration (B_i); (C–D) cumulative infiltration (C_i); (E–F) saturated hydraulic conductivity (K_s); and (G–H) Sorptivity (S).

The mean of parameter B_i (Table 2) indicates a moderately rapid infiltration velocity (Montenegro, Malagón 1990). However, at the location of the study, B_i varies from moderately slow to very fast. Some high infiltration velocities represent relatively dry soils due to the hydrological stress that they are subjected to. Additionally, high CV (greater than 60%) values indicate general areas or sites where infiltration behavior is very different from the other sites (Rodríguez-Vásquez et al. 2008; Martins et al. 2010).

The estimations of the semivariogram parameters (C_o , $C_o + C$ and Range, respectively) were similar (Table 3), which again indicates a good performance of the functional predictor used in this study. In both cases, spherical and exponential models were fit. In this sense, it is important to mention that some authors have reported problems for fitting semivariogram models to these properties (Rodríguez-Vásquez et al. 2008).

According to the results shown in Table 3, we can conclude that there is a moderate spatial dependence for C_i and B_i , because the values of DSD are greater than 25% and lower than 75% (Cambradella et al. 1974). This behavior is similar to that reported by Martins et al. (2010). In addition, the K_s and S show C_o values close to zero, which indicates a strong spatial dependence. A similar result was reported by Rodríguez-Vásquez et al. (2008) for the same type of soil.

The contour maps reveal high spatial variability in terms of the soil properties studied (Figure 5). The maps generated through the estimations obtained by means of functional kriging have a pattern very similar to those obtained from the recorded data (only minimal differences can be identified in each case). We can see in these maps that there is a direct relationship between the parameters (zones with high values of B_i have also high values of C_i). This is also evidenced with the parameters of the Phillip model.

We also note in these maps that there is direct relationship between K_s and B_i and C_i (zones with high values of K_s also have high B_i and C_i values). High values of K_s result from the passage of time and the saturation of soil pores with water, which cause the rate of infiltration to reach a constant value (B_i) that is similar to K_s (Gil, 2002), and consequently the soil can drain greater amounts of water and achieve greater C_i values.

4. Conclusions

The cross-validation analysis showed a high correlation between the results obtained with observed and predicted data (R^2 around 99%), which supported the use of functional geostatistics. Traditional statistical analyses demonstrated the reliability of applying FG to predict infiltration curves, given that the descriptive measures of the parameters for the theoretical infiltration model were similar for the observed data and predictions, even when the parameters displayed high spatial variability.

The estimations of the semivariogram parameters and the contours maps were also similar. The contour maps for B_i , C_i , K_s , and S for the recorded and predicted data demonstrate the wide range of these parameters. However, the maps also showed the same behavioral patterns, which confirmed the congruence between the studied models and the efficacy of predicting these parameters via FG. This was established even when the parameters displayed high spatial variability. All the analyses indicate that functional geostatistics can be a useful tool in the study of the spatial variability of soil properties.

References

- Bolívar, A., Ordóñez, N. (2014). Levantamiento Agrológico del Centro Agropecuario Marengo. Instituto Geográfico Agustín Codazzi. Bogotá, Colombia.
- Camacho-Tamayo, J. H., Rubiano Sanabria, Y., Santana, L. M. (2013): Management units based on the physical properties of an Oxisol. *Journal of Soil Science and Plant Nutrition* 13, 767–785. <https://doi.org/10.4067/S0718-95162013005000061>
- Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., Konopka, A. E. (1994): Field-scale variability of soil properties in Central Iowa Soils. *Soil Science Society of America Journal* 58, 1501–1511. <https://doi.org/10.2136/sssaj1994.03615995005800050033x>
- Chowdary, V. M., Damodhara Rao, M., Jaiswal, C. S. (2006): Study of infiltration process under different experimental conditions. *Agricultural Water Management* 83, 69–78. <https://doi.org/10.1016/j.agwat.2005.09.001>
- Cressie, N. (1993): *Statistics for Spatial Data*. John Wiley and Sons.
- Cucunubá-Melo, J. R., Alvarez Herrera, J. G., Camacho-Tamayo, J. H. (2011): Identification of agronomic management units based on physical attributes of soil. *Journal of Soil Science* 11, 87–99. <https://doi.org/10.4067/S0718-95162011000100008>
- Gil, R. C. (2002). *El comportamiento físico-funcional de los suelos*. Buenos Aires: Instituto de Suelos. Inta Castelar, 19.
- Giraldo, R. (2009). *Geostatistical Analysis of Functional Data*. Barcelona. PhD Thesis submitted to Universitat Politècnica de Catalunya
- Giraldo, R., Delicado, P., Mateu, J. (2010): Continuous time-varying kriging for spatial prediction of functional data: an environmental application. *Journal of Agricultural, Biological, and Environmental Statistics* 15, 66–82. <https://doi.org/10.1007/s13253-009-0012-z>
- Jačka, L., Pavlásek, J., Pech, P., Kuráž, V. (2016): Assessment of evaluation methods using infiltration data measured in heterogeneous mountain soils. *Geoderma* 276, 74–83. <https://doi.org/10.1016/j.geoderma.2016.04.023>
- Isaaks, E., Srivastava, M. (1989): *Applied Geostatistics*. Oxford University Press.
- Latorre, B., Peña, C., Lassabatere, L., Angulo-Jaramillo, R., Moret-Fernández, D. (2015). Estimate of soil hydraulic properties from disc infiltrometer three-dimensional infiltration curve. Numerical analysis and field application. *Journal of Hydrology* 527, 1–12. <https://doi.org/10.1016/j.jhydrol.2015.04.015>
- Machiwal, D., Jha, M. K., Mal, B. C. (2006): Modelling infiltration and quantifying spatial soil variability in a wasteland of Kharagpur, India. *Biosystems Engineering* 95, 569–582. <https://doi.org/10.1016/j.biosystemseng.2006.08.007>
- Martins, A. L., Moura, E., Camacho-Tamayo, K. H. (2010): Variabilidad espacial de la infiltración y su relación con algunas propiedades físicas. *Ingeniería e Investigación* 30, 116–123.
- Montenegro, H., Malagón, D. (1990): *Propiedades Físicas de los Suelos*. Colombia: Instituto Geográfico Agustín Codazzi.
- Moret-Fernández, D., Latorre, B., Angulo-Martínez, M. (2017). Comparison of different methods to estimate the soil sorptivity from an upward infiltration curve. *CATENA* 155, 86–92. <https://doi.org/10.1016/j.catena.2017.03.009>
- Orjuela-Matta, H. M., Rubiano-Sanabria, Y., Camacho-Tamayo, J. H. (2010): Comportamiento de la infiltración en un Oxisol. *Revista U.D.C.A Actualidad and Divulgación Científica* 13, 31–39.
- Ramsay, J., Silverman, B. (2005): *Functional Data Analysis, second edition*. New York: Springer. <https://doi.org/10.1002/0470013192.bsa239>
- Richmond, P., Rillo, S. (2006): Evaluación del efecto de la compactación por el rodado de máquinas sobre algunas propiedades físicas del suelo y el cultivo de trigo en siembra directa. *Argentina: INPOFOS Informaciones agronómicas*, 32.
- Rodríguez-Vásquez, A. F., Aristizábal-Castillo, A. M., Camacho-Tamayo, J. H. (2008): Variabilidad espacial de los modelos de infiltración de Philip y Kostikov en un suelo ándico. *Engenharia. Agrícola* 28, 64–75. <https://doi.org/10.1590/S0100-69162008000100007>