Soil Carbon Stock and Particle Size Fractions in the Central Amazon Predicted from Remotely Sensed Relief, Multispectral and Radar Data

Marcos B. Ceddia [1],*, Andréa S. Gomes [1], Gustavo M. Vasques [2] and Érika F. M. Pinheiro [1]

[1] Departamento de Solos, Instituto de Agronomia, Universidade Federal Rural do Rio de Janeiro, BR 465, km 7, Seropédica 23890-000, Brazil; andrea_zooufc@yahoo.com.br (A.S.G.); erika.solos@gmail.com (É.F.M.P.)
[2] Embrapa Solos, Rua Jardim Botânico, 1024, Rio de Janeiro 22460-000, Brazil; gustavo.vasques@embrapa.br
* Correspondence: marcosceddia@gmail.com; Tel.: +55-21-3787-3772/99239-5942

Abstract: Soils from the remote areas of the Amazon Rainforest in Brazil are poorly mapped due to the presence of dense forest and lack of access routes. The use of covariates derived from multispectral and radar remote sensors allows mapping large areas and has the potential to improve the accuracy of soil attribute maps. The objectives of this study were to: (a) evaluate the addition of relief, and vegetation covariates derived from multispectral images with distinct spatial and spectral resolutions (Landsat 8 and RapidEye) and L-band radar (ALOS PALSAR) for the prediction of soil organic carbon stock (CS) and particle size fractions; and (b) evaluate the performance of four geostatistical methods to map these soil properties. Overall, the results show that, even under forest coverage, the Normalized Difference Vegetation Index (NDVI) and ALOS PALSAR backscattering coefficient improved the accuracy of CS and subsurface clay content predictions. The NDVI derived from RapidEye sensor improved the prediction of CS using isotopic cokriging, while the NDVI derived from Landsat 8 and backscattering coefficient were selected to predict clay content at the subsurface using regression kriging (RK). The relative improvement of applying cokriging and RK over ordinary kriging were lower than 10%, indicating that further analyses are necessary to connect soil proxies (vegetation and relief types) with soil attributes.

1. Introduction

The Brazilian Legal Amazon covers 5,217,423 km^2 (59% of the Brazilian territory) and encompasses a great diversity of vegetation and soils. However, despite its strategic importance, the knowledge about its soils is relatively poor and outdated [1]. The RADAMBRASIL project is the main source of data used to develop models and maps of soil attributes in the Brazilian Amazon. However, the soil maps were published at a coarse scale and the soil data available are very sparse across the Amazon, with ~0.7–3.5 soil profiles per 10,000 km^2 [2]. The main reason for the low density of soil data is the presence of the dense Amazon Rainforest, which acts as a barrier for the pedologists and source of many natural threats (tropical diseases, predators), combined with limited access routes, with most of the territory only accessible by boats or airplane.

Digital soil mapping allows for the creation of maps of soil attributes by exploring the potential of ancillary environmental data such as passive optical and active microwave instruments, and digital elevation models. The covariates derived from digital elevation models have been exhaustively explored in traditional soil surveys and digital soil mapping projects to derive soil–landscape correlations and prediction models, since the topographic maps can often be used as proxies of the

spatial distribution of hydrologic, geomorphologic and biologic factors controlling soil formation [3]. Topography controls the flow of water, solutes and sediments, and these processes affect the catenary soil development, allowing to characterize soils and their properties in relation to their topographic spatial patterns [4].

Multispectral and especially radar remote sensing data are relatively less explored in digital soil mapping, due to costs, the more complex processing required, and to their less direct connection with soil formation processes [5]. These data can be used in various ways, including to: (a) support the segmentation of the landscape into more homogeneous soil–landscape units; (b) predict soil properties using physically-based or empirical prediction methods; and (c) facilitate mapping in large inaccessible areas, where extensive field surveys are impractical. In sparsely vegetated areas, some soil properties (mineralogy, particle size distribution, iron, moisture, organic carbon, carbonate and salinity) have been successfully predicted using spaceborne, airborne and *in situ* measurements from multispectral and microwave sensors [5,6]. However, in densely vegetated areas, the use of remote sensing images for mapping soils suffers from the fact that the vegetation covers much of the soil making it necessary to search for indirect evidences of soil variation inferred from the vegetation. In these cases, the efficiency of using remote sensing to map soil properties is even more dependent on indirect relations between soil attributes and soil formation proxies (vegetation, topography and drainage patterns, for example) [5,7].

While visible and infrared sensors only measure surface characteristics, radar can provide spectral information beyond the vegetative cover and the soil surface (L-band or lower frequencies). In the case of synthetic aperture radar (SAR), the backscattering signal is used to retrieve target properties, including the aboveground biomass and soil moisture [8]. The backscattering coefficient is calculated based on radar sensor configuration (observation frequency, polarization and incident angle) and surface characteristics (roughness and dielectric properties). As the soil water content is the main factor influencing its dielectric constant, the soil moisture content is the most frequently retrieved attribute [8]. Singh and Kathpalia [9] developed a modeling approach based on a genetic algorithm to retrieve simultaneously soil moisture, roughness and texture from the dielectric constant derived from ERS-2 SAR backscatter data. Despite the agreement of the predictions with the field observations, there were problems with the retrieval of the input variables of the model.

The relationship between soil and vegetation has the potential to improve the prediction of soil attributes. However, more detailed information on the vegetation cover may be needed, such as plant functional types and Ellenberg indicator values [5]. Such an approach could be useful for the Amazon basin, which is commonly densely vegetated. Ceddia et al. [2] presented a qualitative model explaining the relationships between soil, relief and vegetation types in the same study area of the current study, in the Urucu River basin, Central Amazon, Brazil. They found that regions with higher slope and lower compound topographic index (CTI) presented higher clay contents and soil organic carbon stocks (CS). Under the conditions of the study, and considering that there is no water deficit during the year, the main vegetation type classified as Upland Dense Forest and the input of carbon from the vegetation, including that from deep roots, explained the higher values for both clay and CS. On the other hand, regions with higher CTI and lower slope (river plains and flatlands) presented both lower clay contents and CS. In these regions the soils are imperfectly drained and the vegetation types are classified as Flooded Lowland Open Tropical Rainforest, and Upland Open Tropical Rainforest, with species that are more adapted to soil aeration restrictions. The CS are lower because the inputs from leaves, trunks and roots are lower and concentrated at the surface layers. The authors showed that the use of CTI improved the prediction of CS using heterotopic cokriging.

Given the potential of remote sensing data to aid digital soil mapping in relatively remote and densely vegetated regions, we hypothesize that the predictions of CS and particle size fractions (PSFs) in the Central Amazon can be improved by adding relief and remotely sensed multispectral and radar covariates. Thus, the objectives were to: (a) evaluate the addition of relief and vegetation covariates derived from multispectral images with distinct spatial and spectral resolutions

(Landsat 8 and RapidEye) and L-band radar (ALOS PALSAR) for the prediction of soil CS and PSFs; and (b) evaluate the performance of four geostatistical methods to map these soil properties, including ordinary kriging (OK), isotopic cokriging (ICOK), heterotopic cokriging (HCOK), and regression kriging (RK).

2. Materials and Methods

This study extends the previous study reported by Ceddia et al. [2]. It uses the same study site and soil data used by the referred authors, but extends the analysis by including: the prediction of PSFs (clay, silt and sand contents), besides CS; other geostatistical methods; and covariates derived from multispectral, and radar remote sensing imagery to help predicting soil CS and PSFs.

2.1. Study Area

The study site is located in the central region of the Amazonas state, in the Urucu River basin, in the municipality of Coari, Brazil (Figure 1). It has an elongated shape of about 50 km by 8 m buffering an oil duct where access was possible, and covers about 7967 ha. The site is located about 640 km from Manaus (the state capital) and can be accessed only by boat or airplane. The climate is equatorial (Af, according to Köppen classification), with the temperature of the coldest month higher than 20 °C, mean annual precipitation of 2500 mm, and no pronounced dry period. According to Brasil [10], the soils of the region were formed from Tertiary and Quaternary sediments of the *Içá* Formation, which covers an area of 563,264 km^2 (36% of the Amazonas state). They are composed of very fine sandstone, claystone and siltstone. The Holocene alluvium of the Quaternary deposits are related to the current Amazonian drainage networks. In the study area, the main soils include Argissolos (Ultisols) and Cambissolos (Inceptisols) [11,12], most of them with low sum of bases, high aluminum and medium to high sand contents. Some soils in the area present aquic characteristics, especially those near the Urucu River floodplain.

According to Brasil [10], the main vegetation types along the study site are Upland Dense Tropical Rainforest, Flooded Lowland Open Tropical Rainforest and Upland Open Tropical Rainforest. The Upland Dense Tropical Rainforest is commonly found in areas with relief forms called *terra firme*, which are upland areas on flat-topped terrain located in river interfluves. The soils in these relief forms are well drained, with no water deficits for plants and have higher clay contents. The Flooded Lowland Open Tropical Rainforests are observed in relief forms called *várzeas*, which are the river floodplains, while the Upland Open Tropical Rainforests are commonly found in imperfectly drained flatlands. The soils of the study site were surveyed between 2008 and 2009 and the main soil classes and their respective numbers of soil profiles (94 in total) are shown in Table 1.

Table 1. Number of soil profiles (N) and frequency of soil suborders in the visited sites.

| SiBCS	Soil Taxonomy	N	Frequency (%)
Argissolos Amarelos	Oxyaquic Hapludults	26	28
Argissolos Vermelhos	Typic Hapludults	2	2
Argissolos Vermelho-Amarelos	Typic Hapludults	21	22
Argissolos Acinzentados	Typic Endoaquults	3	3
Cambissolos Háplicos	Typic Dystrudepts	38	41
Espodossolos Humilúvicos	Humods	1	1
Neossolos Quartzarênicos	Quartzipsamments	2	2
Planossolos Háplicos	Aquults	1	1
Total	-	94	100

SiBCS, Brazilian Soil Classification System.
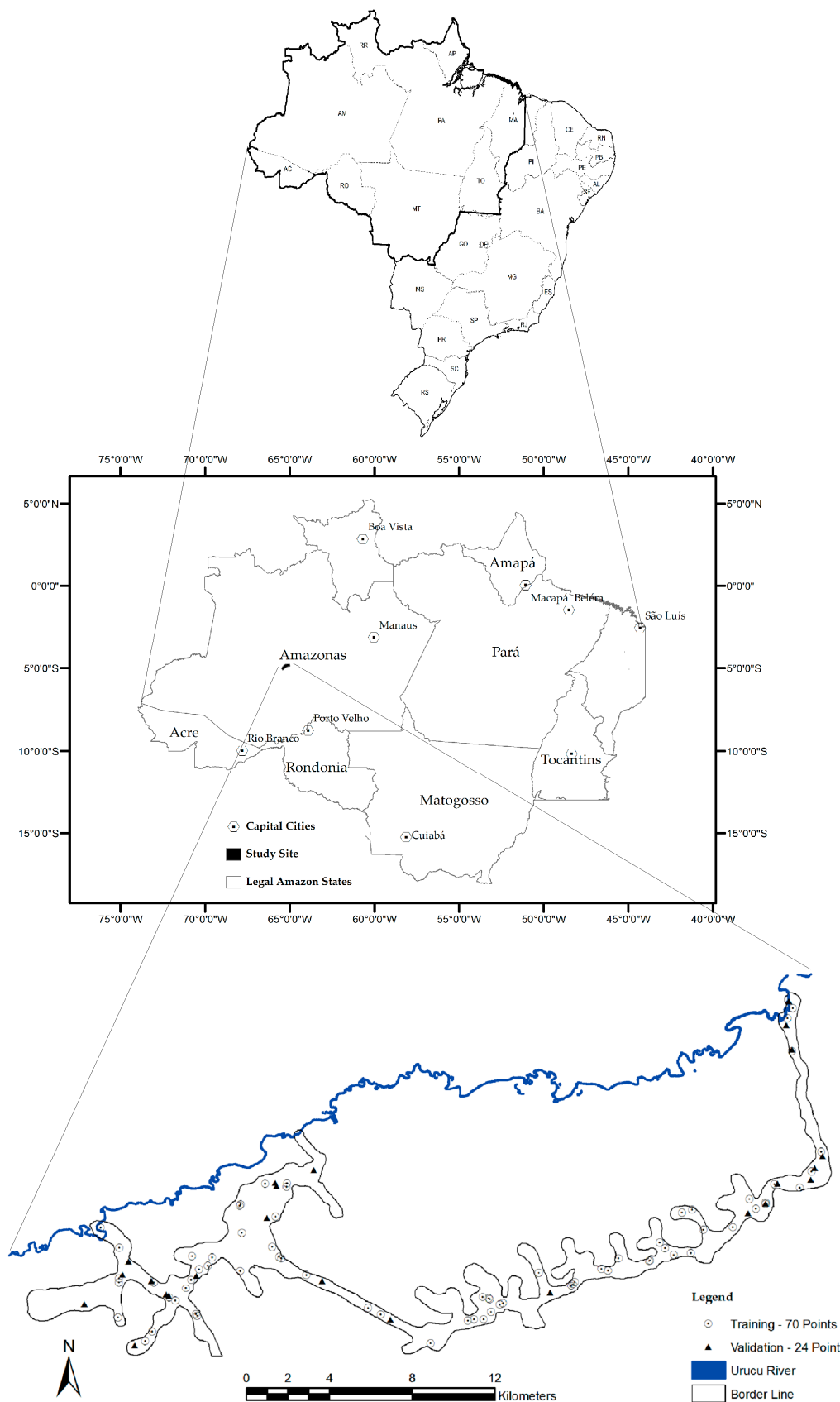
Figure 1. Location of the study area in the Central Amazon, Brazil (upper map), and the study site showing the location of the soil profiles (bottom map).

2.2. Soil Data

In each soil profile soil samples were collected following the soil horizon stratification, and analyzed for organic carbon (OC) by wet combustion [13], bulk density (BD) from Kopeck rings (4.2-cm high and 4.0-cm in diameter), and PSFs—sand, silt and clay contents—by the Pipette method [14]. Soil CS were calculated for the 0–30 and 0–100-cm layers as the depth-weighted sum of the CS across all horizons with portions lying within the desired depth interval (Equation (1)). These two layers are the most used in soil OC studies [15].

$$CS = \sum_{i=1}^{n} OC_i \times BD_i \times T_i \tag{1}$$

where *CS* is the OC stock (kg·m^{-2}) at the desired layer (0–30 or 0–100 cm), OC_i is the OC content (g·kg^{-1}) at horizon *i*, BD_i is the soil bulk density (Mg·m^{-3}) at horizon *i*, T_i is the thickness (m) of the portion of horizon *i* that lies within the desired layer, and *n* is the number of horizons that have a portion within the desired layer.

Similarly, the sand, silt, and clay contents of the surface (Surf) and subsurface (Sub) layers were calculated as the depth-weighted average across the A and AB horizons for the Surf layer, and the BA and B horizons for the Sub layer (Equation (2)). The BC and C horizons were not included.

$$PSF_{Surf/Sub} = \sum_{i=1}^{n} PSF_i \times T_i \bigg/ \sum_{i=1}^{n} T_i \tag{2}$$

where $PSF_{Surf/Sub}$ is the particle size fraction content (g·kg^{-1}) at the desired layer (surface or subsurface), PSF_i is the PSF content (g·kg^{-1}) at horizon *i*, T_i is the thickness (m) of the portion of horizon *i* that lies within the desired layer, and *n* is the number of horizons that have a portion within the desired layer.

2.3. Remote Sensing Covariate Data

Remotely sensed relief, multispectral, and radar imagery were assembled as independent variables, or covariates, for the prediction of soil CS at 0–30 and 0–100 cm, and sand, silt and clay contents at the surface and subsurface. Relief and multispectral data are commonly used in digital soil mapping studies, however, radar imagery is hardly ever included, although it has shown potential to predict soils and soil properties [7,16].

A topographic map with 2-m contour line intervals and a stream network map of the study area were derived at the 1:5000 scale from radar images (X-band, 0.5-m spatial resolution, HH polarization) using interferometry. A digital terrain model (DTM, 0.5-m resolution) with enforced drainage network and filled sinks was derived from the topographic map using the ANUDEM algorithm [17], implemented in the Topo-to-Raster tool in ArcGIS version 9.3 (Esri, Redlands, CA, USA) [18]. Four terrain attributes, namely slope gradient, slope aspect, curvatures, and CTI [19] were calculated from the DTM using the Spatial Analyst tools in ArcGIS.

Three orthorectified RapidEye (RE) satellite images from 16 June 2014, with 5-m resolution (resampled from the original 6.5-m resolution), were obtained from the Brazilian Ministry of Environment, mosaicked, and then used to derive vegetation indices from three bands, including band 3 at 0.630–0.685 μm (red), band 4 at 0.690–0.730 μm (red edge), and band 5 at 0.760–0.850 μm (near infrared, NIR). In addition, a Landsat 8 Operational Land Imager (OLI) image from 31 October 2015 was obtained from the Brazilian National Institute for Space Research. Accordingly, two Landsat 8 OLI sensor bands were used, including band 4 at 0.636–0.673 μm (red), and band 5 at 0.851–0.879 μm (NIR), with 30-m spatial resolution. The RE images and the Landsat 8 OLI image were radiometrically and atmospherically corrected using the 6S model (Second Simulation of Satellite Signal in the Solar Spectrum), originally developed for the simulation of radiance at the satellite level by Vermote et al. [20], and adapted for atmospheric correction by Antunes et al. [21]. As initial conditions, the tropical atmosphere, and the continental aerosol model were used.

Four vegetation indices were derived from the multispectral imagery, including the Normalized Difference Vegetation Index (NDVI) [22], the Enhance Vegetation Index (EVI) [23], the Soil-Adjusted Vegetation Index (SAVI) [24], and, for the RE images only, the Normalized Red Edge Vegetation Index (*NDVI_Ed*) (Equation (3)).

$$NDVI\_Ed = (NIR - Red\ edge)/(NIR + Red\ edge) \tag{3}$$

where *NIR* is the near infrared band, and *Red edge* is the red edge band.

Along with multispectral images, three L-band microwave (23.6-cm wavelength) radar images from the Advanced Land Observing Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) sensor from 6 July 2009 (two images), and 23 July 2009 (one image), with 12.5-m spatial resolution, were acquired and mosaicked. These images were downloaded from Alaska Satellite Facility Distributed Active Archive Center (ASF DAAC) [25]. The images were acquired in ascending orbit, with off-nadir angles of 38.8°, in Level 1.5 Fine Beam Dual (FBD) format, containing two bands with HH, and HV polarization, respectively. According to Shimada [26], the ALOS PALSAR's long operating wavelength in L-band is suitable for tropical forest monitoring due to its high sensitivity to forest structure and moisture characteristics. The ALOS PALSAR images were used to derive the backscattering coefficients ($\sigma°$) of the two polarizations (HH and HV), which indicate the amount of microwave energy that is reflected by the target and returns to the sensor antenna per unit area, in decibels (dB). This was performed using MapReady version 3.2.1 (Alaska Satellite Facility, Fairbanks, AK, USA) [27].

All remotely sensed covariate rasters were assembled in a Geographic Information System and their values extracted to the field soil data in ArcGIS, thus deriving the database used to build the prediction models.

2.4. Prediction Methods

Four geostatistical methods were compared to predict soil CS and PSFs, including ordinary kriging (OK), isotopic cokriging (ICOK), heterotopic cokriging (HCOK), and regression kriging (RK). Ordinary kriging is a univariate method that uses the primary variable (CS or PSF) measured at sampled locations to predict the same primary variable at unsampled locations. In OK, the mean is taken as a constant but unknown value, and its stationarity is assumed only within a local neighborhood centered at the location being predicted. The OK predictor is a best linear unbiased predictor and is written as a linear combination, or weighted average, of the values at the sampled points, with weights adding up to 1 (Equation (4), [28]). The unknown local mean *m(u)* is filtered from the linear estimator by forcing the kriging weights to sum to 1.

$$Z_{OK}^*(u) = \sum_{\alpha=1}^{n_\alpha(u)} \lambda_\alpha^{OK}(u)[Z_1(u)], \text{ with } \sum_{\alpha=1}^{n_\alpha(u)} \lambda_\alpha^{OK} = 1 \tag{4}$$

Cokriging is a bivariate extension of kriging in which a secondary variable (that is, relief, multispectral or radar covariates—Z_v) is incorporated for the prediction of the target variable (CS or PSFs—Z_1) at unsampled locations by accounting for the spatial correlation between them (Equation (5), [27]).

$$Z_{COK}^*(u) = \sum_{\alpha=1}^{n_\alpha(u)} \lambda_\alpha^{COK}(u)[Z_1(u_\alpha)] + \sum_{v=1}^{n_v(u)} \lambda_v^{COK}(u)[Z_v(u_v)], \text{ with } \sum_{\alpha=1}^{n_\alpha(u)} \lambda_\alpha^{COK}(u) = 1 \text{ and } \sum_{v=1}^{n_v(u)} \lambda_V^{COK}(u) = 0 \tag{5}$$

Cokriging does not require that the secondary variable is available at all locations to make predictions. When the target and secondary variables have been measured at the same sampling locations, this is a case of isotopic cokriging. Otherwise, when they have been measured at different locations, heterotopic cokriging can be used [29]. Since the influence of the secondary variable on

predicting the target one depends on: (i) the correlation between them; (ii) their spatial continuity; and (iii) their sampling density and spatial configuration [30], these two forms of cokriging (ICOK and HCOK) were compared.

Regression kriging combines regression (prediction) methods to model the global spatial trend, or external drift, with geostatistical methods to model (interpolate) the residuals from the regression [31]. In RK, multiple linear regression (MLR) with stepwise variable selection ($p < 0.05$) models were first derived in R version 3.1.1 [32] for the target soil variables (CS and PSFs) as a function of relief covariates only, or relief, multispectral and radar covariates together (all variables). For each variable, the model with the highest adjusted R^2 was used to make predictions at all pixels in the study area. Then, the MLR model residuals, derived at the sampled locations, were interpolated across the study area using OK. The final prediction maps were derived as the sum of the MLR predictions and interpolated residuals across the study area. For all geostatistical methods, the auto- and cross-variograms were fitted manually using the spherical model.

In order to run and compare the accuracy of the prediction methods, the whole dataset containing 94 observations was randomly split into training (70 observations) and validation (24 observations) sets (Figure 1). The training observations were used for model development, while the validation observations were set aside for model validation and comparison using the mean error (ME) and root mean square error (RMSE) (Equations (6) and (7)). The ME is used to check the bias of the predictions, whereas the RMSE measures the accuracy of the predictions. The relative improvement (RI, in %; Equation (8)) in the RMSE was derived to compare the ICOK, HCOK and RK against OK.

$$ME = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \tag{6}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \tag{7}$$

$$RI = \frac{RMSE_{OK} - RMSE_{COK \, or \, RK}}{RMSE_{OK}} \times 100 \tag{8}$$

where O_i and P_i are the n observed and predicted values, respectively, OK is ordinary kriging, COK is cokriging, and RK is regression kriging.

3. Results

3.1. Descriptive Statistics and Linear Regression Models

The descriptive statistics of the whole data set (94 observations), as well as the training (70 observations) and validation (24 observations) data sets are presented in Table 2. The different data sets have similar descriptive statistics, with the mean values of all attributes statistically equal between the training and validation data sets, according to a Student's *t*-test at 0.05 significance. The main differences between data sets are observed in the minimum and maximum values of the attributes. Overall, the data sets are adequate for the development of predictive models and their validation, and represent the total set of data collected in the field.

The soil CS at 0–100 cm (CS100) varied from 3.26 to 11.93 kg·m^{-2}, with an average value of 7.38 kg·m^{-2} (Table 2). About 47% of this stock was in the 0–30-cm layer, with an average of 3.44 kg·m^{-2}. The CS obtained in this study, compared with those observed by different studies in the same region [33–36], is 14% to 28% lower at 0–30 cm, and 5% to 17% lower at 0–100 cm. The differences occur because this study is constrained to a single geologic formation in the Central Amazon, whereas the other studies are not. Moreover, the bulk density data used to calculate stocks also differed among the studies. The soil textural classification was loam at the surface, and clay loam at the subsurface, reflecting the constitution of the parent material of the study site, which is mainly composed of very fine sandstone and ferruginous claystone and siltstone. We call attention to the relatively high silt

values of the region, both at the surface (mean: 371 g·kg^{-1}; maximum: 707 g·kg^{-1}) and subsurface (mean: 333 g·kg^{-1}; maximum: 584 g·kg^{-1}) layers, as these amounts of silt are not frequently found in Brazilian soils.

Table 2. Descriptive statistics of target soil variables.

Variables	Set	N	Min	Max	Mean	Median	SD	Skewness	Kurtosis
Sand Surf (g·kg^{-1})	W	94	182	883	456	435	153	0.450	−0.324
Sand Surf (g·kg^{-1})	T	70	182	883	454 [a]	434	154	0.445	−0.077
Sand Surf (g·kg^{-1})	V	24	249	777	463 [a]	437	151	0.577	−0.508
Sand Sub (g·kg^{-1})	W	94	66	855	317	288	145	0.709	0.431
Sand Sub (g·kg^{-1})	T	70	66	855	317 [a]	290	147	0.805	1.060
Sand Sub (g·kg^{-1})	V	24	145	600	317 [a]	259	138	0.501	−1.279
Clay Surf (g·kg^{-1})	W	94	34	352	173	167	77	0.304	−0.538
Clay Surf (g·kg^{-1})	T	70	34	352	171 [a]	165	75	0.435	−0.191
Clay Surf (g·kg^{-1})	V	24	37	335	180 [a]	208	83	−0.035	−1.147
Clay Sub (g·kg^{-1})	W	94	13	531	327	334	104	−0.452	0.145
Clay Sub (g·kg^{-1})	T	70	13	531	321 [a]	332	102	−0.414	0.524
Clay Sub (g·kg^{-1})	V	24	121	478	346 [a]	388	106	−0.665	−0.472
Silt Surf (g·kg^{-1})	W	94	58	707	371	354	134	0.164	−0.376
Silt Surf (g·kg^{-1})	T	70	58	707	375 [a]	359	138	0.102	−0.217
Silt Surf (g·kg^{-1})	V	24	158	627	357 [a]	339	119	0.388	−0.473
Silt Sub (g·kg^{-1})	W	94	85	584	333	338	108	−0.087	−0.096
Silt Sub (g·kg^{-1})	T	70	85	584	336 [a]	342	112	−0.097	−0.122
Silt Sub (g·kg^{-1})	V	24	118	529	325 [a]	321	94	−0.158	−0.036
CS30 (kg·m^{-2})	W	94	1.54	6.44	3.44	3.22	1.11	0.606	−0.071
CS30 (kg·m^{-2})	T	70	1.54	6.44	3.47 [a]	3.21	1.18	0.582	−0.273
CS30 (kg·m^{-2})	V	24	1.86	5.40	3.34 [a]	3.32	0.86	0.428	0.220
CS100 (kg·m^{-2})	W	94	3.26	11.93	7.38	7.54	2.02	0.042	−0.408
CS100 (kg·m^{-2})	T	70	3.64	11.93	7.46 [a]	7.68	2.09	0.059	−0.479
CS100 (kg·m^{-2})	V	24	3.26	10.64	7.16 [a]	7.19	1.75	−0.213	−0.252

W, whole dataset; T, training dataset; V, validation dataset; Surf, surface layer; Sub, subsurface layer; CS30, carbon stock at 0–30 cm; CS100, carbon stock at 0–100 cm; N, number of observations; Min, minimum; Max, maximum; SD, standard deviation. [a] Equal means between the training and validation data sets for each soil attribute, respectively, according to a Student's *t*-test at 0.05 significance.

The regression models developed for soil CS and PSFs are presented in Table 3, except for sand and silt at the subsurface, which had no significant predictor selected among relief, multispectral and radar covariates. The regression models to predict silt at the subsurface, CS at 0–100 cm, and clay at the subsurface presented the highest adjusted R^2, though all regression models had low adjusted R^2, which means that the proportion of the variance in the soil variables that is predictable from the covariates is low. These poor correlations among variables also affected cokriging, as shown later. Nevertheless, the inclusion of covariates derived from Landsat 8 OLI and ALOS PALSAR increased the R^2, mainly for silt at the surface, clay at the subsurface, and CS at 0–30 cm (CS30).

The covariates derived from RapidEye, despite having a higher spatial resolution (5 m), were not selected in any model to predict CS or PSFs. This indicates that better resolution does not necessarily improve or grant significant soil–landscape correlations. The NDVI derived from Landsat 8 OLI (NDVI_L8) was selected in all significant PSF models, and the radar backscatter coefficient was selected in all significant models but for sand at the surface. The selection of relief covariates varied among target variables, with the slope selected to predict sand and silt at the surface, elevation selected only to predict CS, and the CTI selected to predict clay at both layers and CS100.

Table 3. Linear regression models using relief only, and all remote sensing variables.

Target	Covariates	Regression Model	R^2_{adj}
Sand Surf	Relief	-	-
	All variables	$924.43 + 5.83 \times \text{Slope} - 604.31 \times \text{NDVI\_L8}$	0.12
Sand Sub	Relief	-	-
	All variables	-	-
Silt Surf	Relief	$445.91 - 9.40 \times \text{Slope}$	0.09
	All variables	$130.01 + 490.94 \times \text{NDVI\_L8} - 9.68 \times \text{Slope} +$ $11.59 \times \text{BackHH} - 18.58 \times \text{BackHV}$	0.22
Silt Sub	Relief	-	-
	All variables	-	-
Clay Surf	Relief	$310.71 - 14.69 \times \text{CTI} - 75.86 \times \text{CurvP}$	0.15
	All variables	$157.64 + 173.17 \times \text{NDVI\_L8} - 0.10 \times \text{Asp} - 48.16 \times \text{CurvC}$ $- 18.56 \times \text{CTI} - 61.89 \times \text{CurvP} - 61.89 \times \text{BackHH}$	0.19
Clay Sub	Relief	$431.57 + 0.17 \times \text{Asp} - 15.75 \times \text{CTI}$	0.07
	All variables	$50.85 + 337.20 \times \text{NDVI\_L8} - 15.86 \times \text{CTI} - 14.03 \times \text{BackHH}$	0.15
CS30	Relief	$0.07 + 0.05 \times \text{Elev}$	0.09
	All variables	$-0.68 + 0.05 \times \text{Elev} - 0.12 \times \text{BackHH}$	0.11
CS100	Relief	$4.14 - 0.32 \times \text{CTI} + 0.10 \times \text{Elev}$	0.15
	All variables	$3.02 - 0.38 \times \text{CTI} + 0.09 \times \text{Elev} - 0.25 \times \text{BackHH}$	0.19

Surf, surface layer; Sub, subsurface layer; CS30, carbon stock at 0–30 cm; CS100, carbon stock at 0–100 cm; NDVI_L8, Normalized Difference Vegetation Index derived from Landsat 8 Operational Land Imager; BackHH and BackHV, backscattering coefficients derived from ALOS PALSAR using HH and HV polarizations, respectively; CTI, compound topographic index; CurvP, profile curvature; Asp, aspect; CurvC, combined (standard) curvature; Elev, elevation.

3.2. Geostatistical Analysis

The results of the spatial dependence analysis of the target soil variables, residues of the regression models, and remote sensing covariates individually, as well as of the spatial cross-dependence analysis between target soil and remote sensing variables are presented in Table 4 and Figures 2 and 3.

The target soil variables that showed spatial autocorrelation were sand at the surface, silt at the surface and subsurface, clay at the subsurface, and CS100 (Table 4; Figure 2). The CS100 showed the highest range of spatial autocorrelation (8000 m), followed by the clay content at the subsurface (5500 m). Silt at the surface layer had the lowest range (2000 m) and the lowest proportion of the nugget variance over the total variance (that is, nugget-to-sill ratio) among the soil variables.

Ordinary cokriging (ICOK or HCOK) is more demanding than OK inasmuch as two auto-variograms (of the primary and secondary variable, respectively), as well as their cross-variogram, need to be known so that a model of co-regionalization can be defined. Considering these requisites, only two variables (silt at the subsurface and CS100) could be modeled by ICOK and HCOK. The candidate secondary variables that were spatially autocorrelated and could be used as covariates included elevation, slope, CTI, NDVI_RE, NDVI_Ed, BackHH and BackHV, and only a few of them were spatially correlated to silt or CS100 (Table 4; Figure 2). The cross-variograms show that silt at the subsurface is negatively correlated with elevation and slope, and positively correlated with CTI, whereas CS100 had the opposite behavior in relation to these covariates. Indeed, areas with higher elevation and slope presented higher CS and lower silt contents, whereas areas with higher CTI presented lower CS and higher silt contents in the study area. All cross-variograms presented lower nugget-to-sill ratio compared to the original soil variables, meaning that the spatial continuity is stronger when the secondary variables are considered for cokriging.

Table 4. Fitted auto- and cross-variogram parameters (spherical model).

Variables	C_0	C_1	Sill	Range (m)	$\frac{C_0}{C_0 + C_1} \times 100$ (%)
Original variables					
Sand Surf	14,656	9314	23,970	2256	61.1
Silt Surf	5000	14,000	19,000	2000	26.3
Silt Sub	6971	6042	13,013	3256	53.6
Clay Sub	8200	3500	11,700	5500	70.1
CS100	3.70	1.30	5.0	8000	74.0
Regression residuals					
Sand Surf	15,000	6870	21,870	2000	68.6
Silt Surf	10,000	4100	14,100	3000	70.9
Clay Sub	7000	2000	9000	8000	77.8
CS100	3.20	2.70	5.90	8200	54.2
Silt Sub $\times$ secondary variables					
Elevation	15	35	50	3256	30.0
Slope	16	9	25	3256	64.0
CTI	2.5	0.8	3.0	3256	75.8
NDVI_RE	0.001	0.002	0.003	3256	20.0
NDVI_Ed	0.0012	0.0005	0.0017	3256	72.7
BackHH	2.5	2.2	4.7	3256	53.2
BackHV	3.7	0.9	4.6	3256	80.4
Silt Sub $\times$ Elevation	−50.0	−109.0	−159.0	3256	31.4
Silt Sub $\times$ Slope	0.0	−130.0	−130.0	3256	0.0
Silt Sub $\times$ CTI	0.0	41.0	41.0	3256	0.0
CS100 $\times$ secondary variables					
Elevation	20	35	55	8000	36.4
Slope	19	6	25	8000	76.0
CTI	2.5	0.8	3.3	8000	75.8
NDVI_RE	0.001	0.002	0.000	8000	33.3
NDVI_Ed	0.0012	0.0005	0.0000	8000	72.7
BackHH	3.0	1.8	4.8	8000	62.5
BackHV	3.8	0.9	4.7	8000	80.9
CS100 $\times$ Elevation	0.9	5.1	6.0	8000	15.0
CS100 $\times$ Slope	0.0	3.5	3.5	8000	0.0
CS100 $\times$ CTI	−0.80	−0.40	−1.20	8000	66.7
CS100 $\times$ NDVI_RE	0.0000	0.0098	0.0098	8000	0.0

C_0, nugget variance; C_1, sill variance; Surf, surface layer; Sub, subsurface layer; CS100, carbon stock at 0–100 cm; NDVI_RE and NDVI_Ed, Normalized Difference Vegetation Index and Normalized Red Edge Vegetation Index derived from RapidEye, respectively; BackHH and BackHV, backscattering coefficients derived from ALOS PALSAR using HH and HV polarizations, respectively; CTI, compound topographic index.

The NDVI_RE was the only multispectral index that could be used to map CS using cokriging. In this case, areas with higher NDVI_RE had higher CS. The lack of other multispectral indices that present cross spatial dependence with CS or PSFs was unexpected because the soil types are to some extent related to vegetation types in the study area [2]. Moreover, the vegetation indices derived from RapidEye images did not cross-vary with the soil variables, which suggests that the higher spatial resolution of the indices did not help cokriging, except for CS100.

Regression kriging requires that the regression residuals are interpolated using OK. To do so, the regression residuals need to present spatial autocorrelation. From all soil attributes with fitted regression models, only sand and silt contents at the surface, clay content at the subsurface, and CS100 presented autocorrelated residuals (Table 4; Figure 3), which could be interpolated and added back to the regression predictions. Clay at the surface and CS30 could not be predicted using RK.
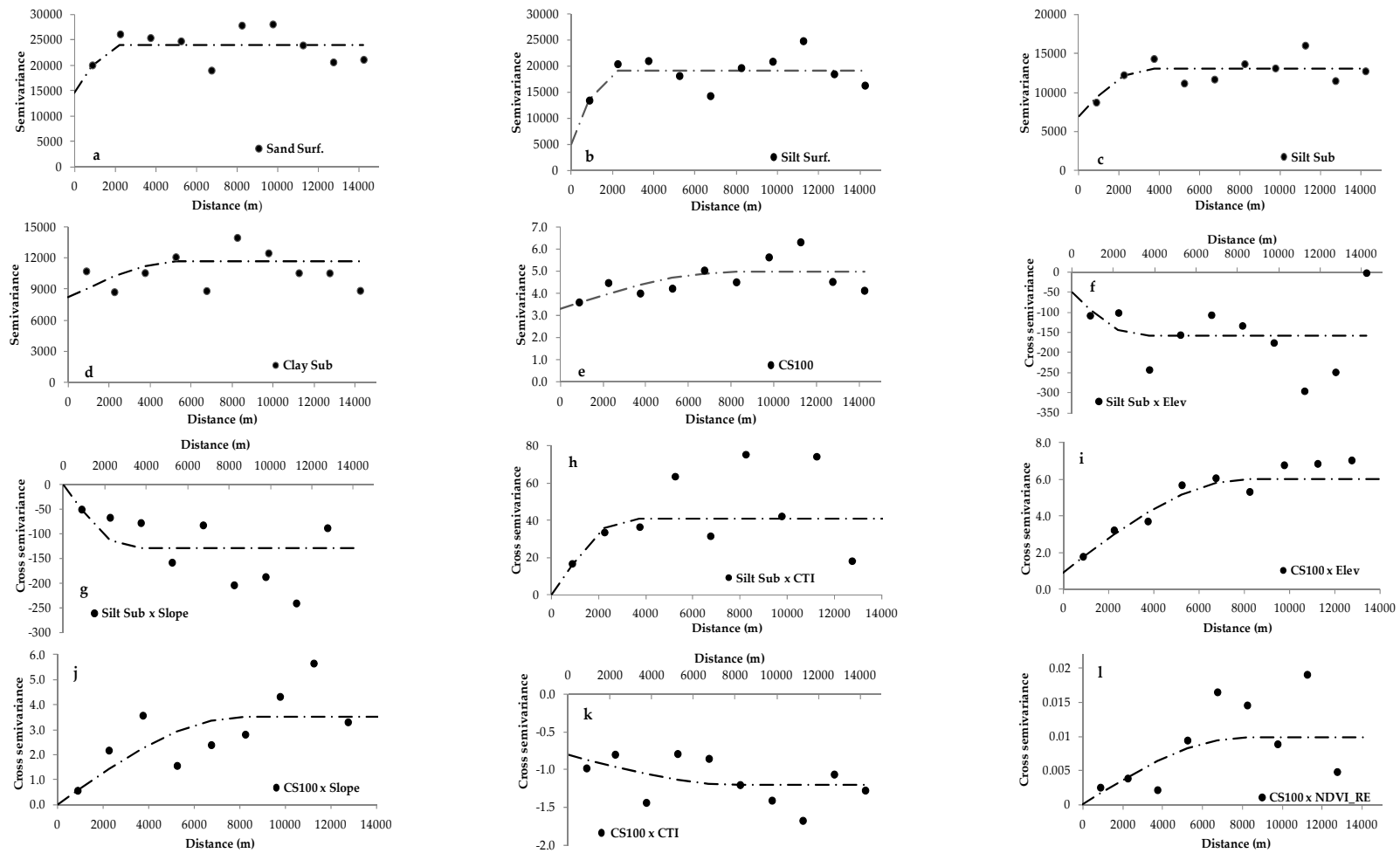
Figure 2. Experimental and fitted auto- and cross-variograms: (**a**) sand content at the surface; (**b**) silt content at the surface; (**c**) silt content at the subsurface; (**d**) clay content at the subsurface; (**e**) carbon stock at 0–100 cm (CS100); (**f**) silt content at the subsurface and elevation (Elev); (**g**) silt content at the subsurface and slope; (**h**) silt content at the subsurface and compound topographic index (CTI); (**i**) CS100 and elevation; (**j**) CS100 and slope; (**k**) CS100 and CTI; and (**l**) CS100 and Normalized Difference Vegetation Index derived from RapidEye (NDVI_RE).
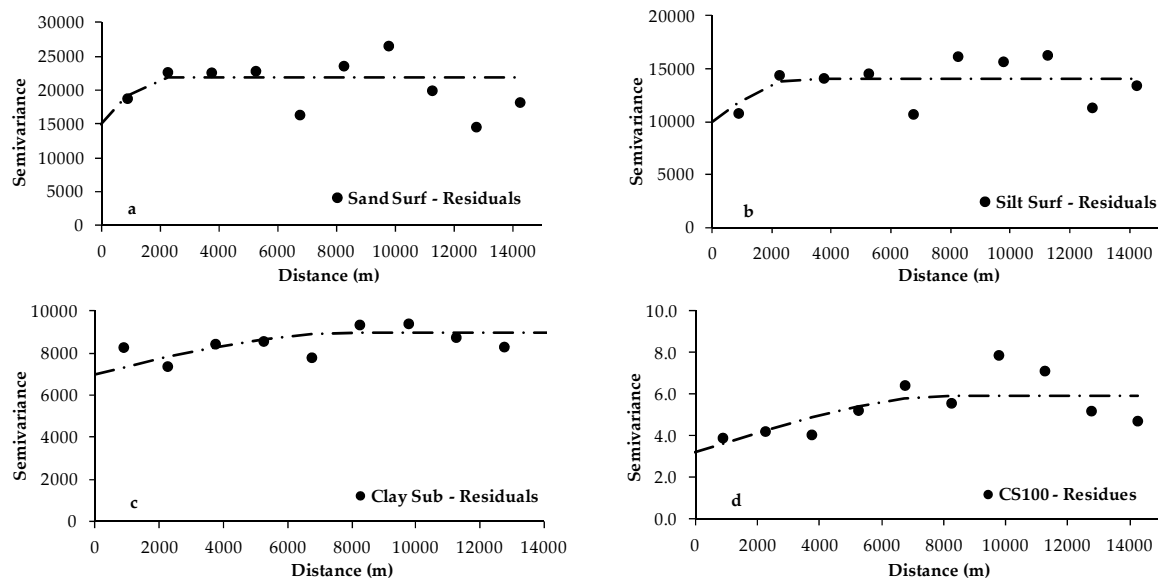
Figure 3. Experimental and fitted variograms of the regression residuals: (**a**) sand content at the surface; (**b**) silt content at the surface; (**c**) clay content at the subsurface; and (**d**) carbon stock at 0−100 cm.

In summary, considering the requirements of the kriging methods, most of the soil variables could be mapped by OK and RK, but only a couple could be cokriged. In OK, the variables not mapped included sand at the subsurface, clay at the surface and CS30, which did not show spatial autocorrelation. In RK, the variables not mapped included sand and silt at the subsurface, which lacked significant correlations with the covariates to derive a regression model, and clay at the surface and CS30, whose regression residuals were not spatially autocorrelated. Finally, only silt at the subsurface and CS100 could be mapped by ICOK and HCOK.

The final predicted maps of the soil attributes are shown in Figure 4. Among the prediction methods compared to map CS100, the best results (that is, with the lowest RMSE of external validation) were obtained by ICOK using slope as covariate, followed by ICOK using CTI, and then HCOK using NDVI from RapidEye (Table 5). For clay at the subsurface, RK outperformed OK. In this case, the regression model selected CTI, NDVI from Landsat 8 OLI, and the backscattering coefficient in HH polarization, contributing to improve both the ME and RMSE. These results show that cokriging (isotopic or heterotopic) has potential to improve the accuracy of soil attribute predictions using relief and multispectral remote sensing data. However, the gain in accuracy from applying these methods need to be evaluated from a cost-efficiency perspective, when compared to the simpler and cheaper OK method. Thus, the relative improvement (RI) of adding covariates to help in the predictions (as in ICOK, HCOK and RK) over OK was computed (Table 5). The RI shows that the RMSE of CS100 and clay at the subsurface were only modestly improved by up to 5% by ICOK, and 8% by RK. In comparison, OK derived the best predictions for sand at the surface, and silt at the surface and subsurface. Clay at the surface, sand at the subsurface, and CS30 were not mapped.

Table 5. External validation errors, with the best model indicated in bold for each target soil variable, respectively.

Variables	ME	RMSE	RI (%)
Ordinary kriging			
CS100 (kg·m^{-2})	0.08	1.79	-
Sand Surf (g·kg^{-1})	4.40	**198.80**	-
Silt Surf (g·kg^{-1})	6.26	**139.61**	-
Silt Sub (g·kg^{-1})	−27.38	**80.02**	-
Clay Sub (g·kg^{-1})	−33.43	127.96	-
Isotopic cokriging			
CS100/Slope (kg·m^{-2})	0.24	**1.70**	**5.0**
CS100/Elevation (kg·m^{-2})	0.24	1.75	2.2
CS100/NDVI_RE (kg·m^{-2})	0.20	1.78	0.5
CS100/CTI (kg·m^{-2})	0.26	1.73	3.4
Silt Sub/Slope (g·kg^{-1})	−17.22	83.40	−4.2
Silt Sub/Elevation (g·kg^{-1})	−16.62	82.15	−2.7
Silt Sub/NDVI_RE (g·kg^{-1})	−19.42	84.82	−6.0
Silt Sub/CTI (g·kg^{-1})	−19.16	81.39	−1.7
Heterotopic cokriging			
CS100/Slope (kg·m^{-2})	−0.03	2.11	−17.8
CS100/Elevation (kg·m^{-2})	0.27	1.91	−6.7
CS100/NDVI_RE (kg·m^{-2})	0.28	1.74	2.8
CS100/CTI (kg·m^{-2})	0.16	1.89	−5.6
Silt Sub/Slope (g·kg^{-1})	−7.40	98.78	−23.4
Silt Sub/Elevation (g·kg^{-1})	−19.32	81.59	−2.0
Silt Sub/NDVI_RE (g·kg^{-1})	−18.19	83.17	−3.9
Silt Sub/CTI (g·kg^{-1})	−29.34	83.72	−4.6
Regression kriging			
CS51 (kg·m^{-2})	0.51	2.52	−42
Sand Surf (g·kg^{-1})	−84.46	212.83	−7.1
Silt Surf (g·kg^{-1})	134.80	224.69	−61
Clay Sub (g·kg^{-1})	−32.32	**117.54**	**8.2**

ME, mean error; RMSE, root mean square error; RI, relative improvement.
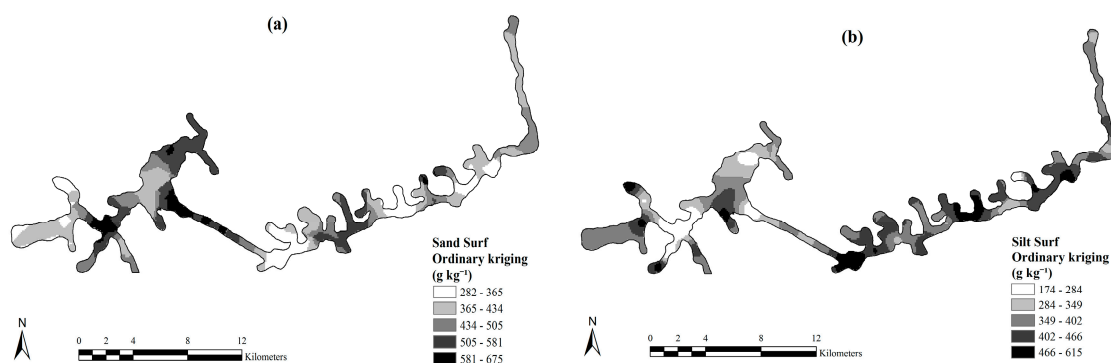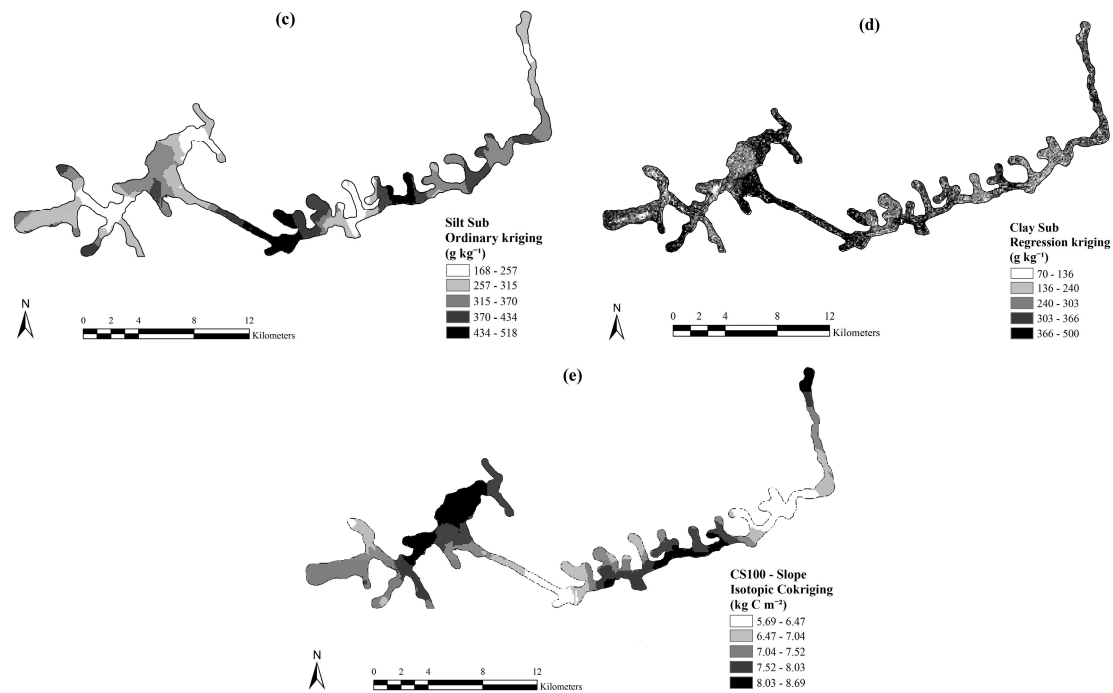

Figure 4. *Cont.*

Figure 4. Final predicted maps of soil attributes derived from their best prediction methods: (**a**) sand content at the surface, (**b**) silt content at the surface, and (**c**) silt content at the subsurface predicted by ordinary kriging; (**d**) clay content at the subsurface predicted by regression kriging; and (**e**) carbon stock at 0–100 cm (CS100) predicted by isotopic cokriging using slope as covariate.

4. Discussion

4.1. Potential of Using Multispectral and Radar Data as Covariates

This study offers new insights on the use of covariates derived from relief, multispectral and radar remote sensing data to improve the accuracy of soil attribute predictions from geostatistical methods under native Amazon forest. Considering the regression models selected to predict soil attributes, the covariates derived from relief, multispectral and radar data together explained only a low proportion (<25%) of the variance of the soil attributes, respectively. Nonetheless, the addition of NDVI (from Landsat 8 OLI) and backscattering coefficients (from ALOS PALSAR) increased substantially the prediction power of some models, compared to the others using only relief covariates (Table 3). This is the case of the models for silt at the surface (an increase from 9% to 22%) and clay at the subsurface (an increase from 7% to 15%). The clay content at the subsurface was better predicted by RK than OK, with the regression model selecting relief, multispectral, and radar variables. Although this was the only soil attribute better predicted by RK, the preference of OK over RK is not unusual, and has been reported elsewhere [37–40]. Thus, there is potential to improve soil attribute predictions by adding remote sensing covariates.

The clay content at the subsurface is higher when the value of NDVI is higher and both the backscattering coefficient and CTI are lower, respectively. The NDVI indicates the "greenness" of the land cover, or the vegetation health, and is related to the aboveground biomass. According to Ceddia et al. [2], in the study site the densest vegetation type is Upland Dense Forest, whose NDVI are the highest in the study area. This vegetation type occurs more frequently in areas of low CTI on soils with relatively high clay contents. In the opposite direction, regions with the highest CTI and lowest clay contents occur closer to the river floodplain, where the plant species are more adapted to soil aeration restrictions. The vegetation types in these regions are classified as Flooded Lowland Open Tropical Rainforest and Upland Open Tropical Rainforest, which have lower density and smaller trunk diameters (thus, lower NDVI). Some authors have found similar results, showing that some

soil attributes or types affect the NDVI. For example, Lozano-Garcia et al. [40] derived NDVI values from the Advanced Very High Resolution Radiometer (AVHRR) sensor over 45 fields representing 8 soil associations in the state of Indiana (USA), finding higher NDVI values in forested areas with soils without restrictive layers, and higher clay content and water-holding capacity. Farrar et al. [41] examined the extent to which differences in the rate of soil moisture generation, as a function of soil type or locality, accounts for the NDVI in semiarid Botswana. According to the authors, moisture availability is not the only soil characteristic determining the amount of vegetation growth. Instead, a group of soil attributes including moisture, porosity, nutrient availability, profile characteristics, and chemical attributes play a role on vegetation growth and rain-use efficiency. Although the latter example [41] refers to an environment completely different from the Amazon region, it emphasizes the complexity of factors directly or indirectly affecting the water dynamics in soils and the response of vegetation and NDVI.

The backscattering coefficient in HH polarization was selected in most regression models of PSFs and CS. The results found by Shimada [26] can explain the predominance of the HH polarization in the regression models. The author developed a polarimetric-calibration method for ALOS PALSAR using time series imagery acquired over dense forest in the Brazilian Amazon (Rio Branco, Acre), and found that the HH signal of PALSAR penetrates the forest canopy deeper and returns from the bottom of the forest stronger than the VV signal. The author argued that the total received power from the HH polarization could be decomposed into, or originated from, the double-bounce (or penetrating) signal, which accounts for about 20% of the signal, and the volume of the canopy, which accounts for the remaining 80%.

Although NDVI and radar backscattering coefficients showed higher potential to be used as covariates to improve the prediction of soil attributes, the prediction power of the models remained low. The causes may be related to the low sensitivity of these covariates to differentiate the soil patterns under native rainforest. In other words, the range of the NDVI values may not have been sufficient to discretize the different vegetation types and their associated soil attributes (CS and PSFs). Theoretical modeling results proved that large aboveground biomass in dense forests can decrease backscatter, due to the saturation of the free-space (or not attenuation) [42]. This decrease in the backscatter can have strong implications for the use of L-band radar data for mapping, such as severe underprediction of soil water content and aboveground biomass [8,42]. This is especially important because the soil water content and dynamics play an important role to connect soil formation factors (vegetation and topography) with soil formation processes and soil attributes. However, according to Mermoz et al. [42], the L-band does not entirely lose sensitivity at large biomass values, suggesting that much progress can be made by refining our understanding of radar backscattering behavior in the L-band.

The NDVI derived from RapidEye was selected as covariate to cokrige soil CS100 using HCOK (though the best CS100 predictor was ICOK using slope as covariate). Its higher spatial resolution (5 m) seems to be important to find the spatial cross-continuity between CS and NDVI. Rocchini [43] points out that sensors with higher spatial resolution capture higher levels of detail, consequently better representing the spatial variation within a short distance. On the other hand, coarser resolution data (for example, Landsat 8 OLI with 30-m resolution) tend to have mixed-pixel problems and hence less sensitivity to spatial complexity at short distances. However, in the opposite direction, recent research has shown that higher spatial resolution in covariates does not necessarily improve prediction quality [44].

4.2. Performance of the Different Prediction Methods

The performance of the prediction methods applied in this study changed according to the soil attribute being mapped. Ordinary kriging presented the best performance to predict sand content at the surface and silt both at the surface and subsurface, whereas the clay content at the subsurface was better predicted using RK. The only soil variables that could be cokriged were CS100 and silt at the

subsurface. This is unfortunate because CS100 was best predicted by ICOK and the cokriged silt maps were practically as accurate as the OK one. This suggests that the other soil variables could have been better predicted by cokriging but only a few covariates had the necessary spatial configurations to meet the assumptions of this method (for example, NDVI derived from RapidEye). These results are in accordance with other studies [45,46] showing that there is no single best prediction method for all soil attributes. Cokriging and RK, despite having potential to improve the predictions, are more complex methods, and thus, the extra effort for their implementation must be taken into consideration. The RI (Table 5) gives an idea of how much improvement can be obtained from applying cokriging or RK, relative to OK, which takes the least effort. Kravchenko and Robertson [47], mapping total soil carbon, evaluated the RI of applying RK compared to OK. The authors found that the combination of the primary variable having strong spatial continuity (nugget-to-sill ratio < 30%), and the primary and secondary variables being relatively weakly correlated (R^2 < 0.40) appeared to be particularly unfavorable to obtain any improvement from RK. Besides, when the R^2 was smaller than 0.30, the maximum improvement from RK did not exceed 10%. The same reasoning from Kravchenko and Robertson [47] can be considered to explain the poor performance of RK in this study (RI < 10%), since the R^2 of all regression models were less than 0.30. The exception was the clay content at the subsurface, which had a R^2 of 0.15 and a nugget-to-sill ratio of 70.1%, but still was better predicted by RK than OK. Its high nugget-to-sill ratio can be caused by measurement errors and by the limited number and spatial configuration of the training observations, which might not have captured the structural variance at shorter distances. The weak correlation between the clay content at the subsurface and the covariates from the regression models can be a consequence of the reduction of the effect of topography and surface processes as the soil depth increases. For example, within a relatively short distance, some soils in the study site have textural gradients (intensive increase of clay content with depth), while others do not. The textural gradient depends on the translocation of clay, which, in turn, depends on both the water dynamics and the degree of clay flocculation or dispersion within the soil profile.

5. Conclusions

This work hypothesized that the use of relief, multispectral and radar remote sensing data can improve the accuracy of soil attribute predictions under Amazon native forest. Although the improvements were modest and limited to a couple of soil attributes, there is potential to use remotely sensed multispectral and radar data in the Amazon to improve the accuracy of soil attribute predictions. However, further research is necessary to better connect soil formation proxy variables derived from remote sensing to the spatial distribution of soil attributes in the Amazon, including the more commonly used vegetation indices and relief variables, but especially those still little explored like L-band radar data. The Amazon Rainforest offers harsh conditions for both soil data collection in the field (due to its remoteness, access limitations, risks, etc.), and data correlation between field soil and remote sensing data (due to the dense land cover that hinders direct soil observation from above, with the models relying on indirect correlations among remote sensing reflectance data, vegetation/land cover, and soil attributes).

The source and type (multispectral vs. radar) of the remote sensing data influenced which covariates were selected by the prediction models, with the NDVI derived from Landsat 8 OLI and the ALOS PALSAR backscattering coefficients most frequently selected. On the other hand, among multispectral and radar covariates, only RapidEye NDVI presented spatial cross-continuity to predict CS100 by cokriging. This may be related to the spatial resolution of the data, because only fine-resolution (multispectral and relief) covariates could be used for cokriging, contrary to the coarser-resolution radar (12.5 m) and Landsat 8 data (30 m). This is important because ICOK and HCOK presented lower RMSE to predict CS100, thus it may be worth testing other fine-resolution data in the search for a better improvement than those found in this study that reached a maximum of 8%.

References

1. Departamento Nacional de Produção Mineral. *Projeto RADAMBRASIL*; Departamento Nacional de Produção Mineral: Rio de Janeiro, Brazil, 1973–1987.
2. Ceddia, M.B.; Villela, A.L.O.; Pinheiro, E.F.M.; Wendroth, O. Spatial variability of soil carbon stock in the Urucu river basin, Central Amazon-Brazil. *Sci. Total Environ.* **2015**, *526*, 58–69. [CrossRef] [PubMed]
3. Moore, I.D.; Grayson, B.R.; Ladson, A.R. Digital terrain modeling: A review of hydrological, geomorphological, and biological applications. *Hydrol. Process.* **1991**, *5*, 3–30. [CrossRef]
4. Sumfleth, K.; Duttmann, R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. *Ecol. Indic.* **2008**, *8*, 485–501. [CrossRef]
5. Mulder, V.L.; Bruin, S.; Schaepman, M.E.; Mayr, T.R. The use of remote sensing in soil and terrain mapping—A review. *Geoderma* **2011**, *162*, 1–19. [CrossRef]
6. Shabou, M.; Mougenot, B.; Chabaane, Z.L.; Walter, C.; Boulet, G.; Aissa, N.B.; Zribi, M. Soil clay content mapping using a time series of Landsat TM data in semi-arid lands. *Remote Sens.* **2015**, *7*, 6059–6078. [CrossRef]
7. McBratney, A.B.; Mendonça Santos, M.L.; Minasny, B. On digital soil mapping. *Geoderma* **2003**, *117*, 3–52. [CrossRef]
8. Dubois, P.C.; van Zyl, J.; Engman, T. Measuring soil moisture with imaging radar. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 915–926. [CrossRef]
9. Singh, D.; Kathpalia, A. An efficient modeling with GA approach to retrieve soil texture, moisture and roughness from ERS-2 SAR data. *Prog. Electromagn. Res.* **2007**, *77*, 121–136. [CrossRef]
10. Departamento Nacional de Produção Mineral. *Projeto RADAMBRASIL. Folha SB. 20 Purus: Geologia, geomorfologia, pedologia, vegetação e uso potencial da terra*; Departamento Nacional de Produção Mineral: Rio de Janeiro, Brazil, 1978.
11. Embrapa. *Sistema Brasileiro de Classificação de Solos*, 3rd ed.; Embrapa: Brasília, Brazil, 2013.
12. Soil Survey Staff. *Keys to Soil Taxonomy*, 12th ed.Natural Resources Conservation Service, United States Department of Agriculture: Washington, DC, USA, 2014.
13. Walkley, A.; Black, I.A. An examination of the Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci.* **1934**, *37*, 29–38. [CrossRef]
14. Embrapa. *Manual de Métodos de Análise de Solo*, 2nd ed.Centro Nacional de Pesquisa de Solos: Rio de Janeiro, Brazil, 1997.
15. Batjes, N.H. Effects of mapped variation in soil conditions on estimates of soil carbon and nitrogen stocks for South America. *Geoderma* **2000**, *97*, 135–144. [CrossRef]
16. Metternicht, G.I.; Zinck, J.A. Remote sensing of soil salinity: Potentials and constraints. *Remote Sens. Environ.* **2003**, *85*, 1–20. [CrossRef]
17. Hutchinson, M.F. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* **1989**, *106*, 211–232. [CrossRef]
18. *ArcGIS*, version 9.3; ESRI (Environmental Systems Research Institute): Redlands, CA, USA, 2006.
19. Gessler, P.E.; Moore, I.D.; McKenzie, N.J.; Ryan, P.J. Soil-landscape modeling and spatial prediction of soil attributes. *Int. J. Geogr. Inf. Sci.* **1995**, *9*, 421–432. [CrossRef]
20. Vermote, E.F.D.; Tanré, D.; Deuzé, J.L.; Herman, M.; Morcrete, J.J. Second simulation of the satellite signal in the solar spectrum, 6S: An overview. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 675–686. [CrossRef]

21. Antunes, M.A.H.; Debiasi, P.; Costa, A.R.; Gleriani, J.M. Correção atmosférica de imagens ALOS/AVNIR-2 utilizando o modelo 6S. *Rev. Bras. Cartogr.* **2012**, *64*, 531–539.

22. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the great plains with ERTS. In *Earth Resources Technology Satellite-1 Symposium*; NASA: Washington, DC, USA, 1973; Volume 1, pp. 309–317.

23. Huete, A.; Justice, C.; Liu, H. Development of vegetation and soil indices for MODIS-EOS. *Remote Sens. Environ.* **1994**, *29*, 224–234. [CrossRef]

24. Huete, A. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]

25. Alaska Satellite Facility Distributed Active Archive Center (ASF DAAC). Available online: https://www.asf.alaska.edu/sar-data/palsar/download-data/ (accessed on 23 March 2016).

26. Shimada, M. Model-Based polarimetric SAR calibration method using forest and surface-scattering targets. *IEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1712–1733. [CrossRef]

27. Alaska Satellite Facility. MapReady 3.2.1. Available online: http://www.asf.alaska.edu/sardatacenter/softwaretools (accessed on 15 March 2016).

28. Goovaerts, P. *Geostatistics for Natural Resource Evaluation*; Oxford University Press: New York, NY, USA, 1997; Volume 27, pp. 315–334.

29. Wackernagel, H. *Multivariate Geostatistics: An Introduction with Applications*, 3rd ed.; Springer: New York, NY, USA, 2003; p. 387.

30. Simbahan, G.C.; Dobermann, A.; Goovaerts, P.; Ping, J.; Haddix, M.L. Fine-resolution mapping of soil organic carbon based on multivariate secondary data. *Geoderma* **2006**, *132*, 471–489. [CrossRef]

31. Grunwald, S. *Environmental Soil-Landscape Modeling-Geographic Information Technologies and Pedometrics*; CRC Press: New York, NY, USA, 2006; pp. 154–196.

32. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria, Version 3.1.1; 2014; Available online: http://www.R-project.org (accessed on 24 January 2017).

33. Moraes, J.L.; Cerri, C.C.; Melillo, J.M.; Kicklighter, D.; Neill, C.; Skole, D.L.; Steudler, P.A. Soil carbon stocks of the Brazilian Amazon Basin. *Soil Sci. Soc. Am. J.* **1995**, *59*, 244–247. [CrossRef]

34. Batjes, N.H.; Dijkshoorn, J.A. Carbon and nitrogen stocks in the soils of the Amazon Region. *Geoderma* **1999**, *89*, 273–286. [CrossRef]

35. Cerri, C.C.; Bernoux, M.; Arrouays, D.; Feigl, B.; Piccolo, M.C. Carbon pools in soils of the Brazilian Amazon. In *Global Climate Change and Tropical Ecosystems*; Lal, R., Ed.; CRC Press: Boca Raton, FL, USA, 2000; pp. 33–50.

36. Batjes, N.H. Organic carbon stocks in the soils of Brazil. *Soil Use Manag.* **2005**, *21*, 22–24. [CrossRef]

37. Vasques, G.M.; Grunwald, S.; Comerford, N.B.; Sickman, J.O. Regional modelling of soil carbon at multiple depths within a subtropical watershed. *Geoderma* **2010**, *156*, 326–336. [CrossRef]

38. Zhu, Q.; Lin, H.S. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. *Pedosphere* **2010**, *20*, 594–606. [CrossRef]

39. Vasques, G.M.; Coelho, M.R.; Dart, R.O.; Oliveira, R.P.; Teixeira, W.G. Mapping soil carbon, particle-size fractions, and water retention in tropical dry forest in Brazil. *Pesqui. Agropecu. Bras.* **2016**, *51*, 1371–1385. [CrossRef]

40. Lozano-Garcia, D.F.; Fernandez, R.N.; Johannsen, C.J. Assessment of regional biomass-soil relationships using vegetation indexes. *IEEE Trans. Geosci. Remote Sens.* **1991**, *29*, 331–338. [CrossRef]

41. Farrar, J.F.; Minchin, P.E.H.; Thorpe, M.R. Carbon import into barley roots: Stimulation by galactose. *J. Exp. Bot.* **1994**, *45*, 17–22. [CrossRef]

42. Mermoz, S.; Rejou-Mechain, M.; Villard, L.; Toan, T.L.; Rossi, V.; Gourlet-Fleury, S. Decrease of L-band SAR backscatter with biomass of dense forests. *Remote Sens. Environ.* **2015**, *159*, 307–317. [CrossRef]

43. Rocchini, D. Effects of spatial and spectral resolution in estimating ecosystem alpha-diversity by satellite imagery. *Remote Sens. Environ.* **2007**, *111*, 423–434. [CrossRef]

44. Samuel-Rosa, A.; Heuvelink, G.B.M.; Vasques, G.M.; Anjos, L.H.C. Do more detailed environmental covariates deliver more accurate soil maps? *Geoderma* **2015**, *243*, 214–227. [CrossRef]

45. Odeh, I.O.A.; McBratney, A.B.; Chittleborough, D.J. Further results on prediction of soil properties from terrain attributes: Heterotpic cokriging and regression-kriging. *Geoderma* **1995**, *67*, 215–225. [CrossRef]

46. Hengl, T.; Heuvelink, G.B.M.; Stein, A. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* **2004**, *120*, 75–93. [CrossRef]

47. Kravchenko, A.N.; Robertson, G.P. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? *Agron. J.* **2007**, *99*, 12–17. [CrossRef]