Geoderma 213 (2014) 46–56

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty

A.H. Cambule [a,*], D.G. Rossiter [b], J.J. Stoorvogel [c], E.M.A. Smaling [b]

[a] Universidade Eduardo Mondlane, Faculdade de Agronomia e Engenharia Florestal, C.P. 257, Maputo, Mozambique
[b] University of Twente, Faculty of Geoinformation Science and Earth Observation (ITC), P.O. Box 6, 7500 AA Enschede, The Netherlands
[c] Wageningen University, Land Dynamics Group, P.O. Box 47, 6700 AA Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 14 March 2012
Received in revised form 25 June 2013
Accepted 21 July 2013
Available online 26 August 2013

Keywords:
Soil organic carbon stocks
Kriging
Regional trend
Digital soil mapping

ABSTRACT

Many areas in sub-Saharan African are data-poor and poorly accessible. The estimation of soil organic carbon (SOC) stocks in these areas will have to rely on the limited available secondary data coupled with restricted field sampling. We assessed the total SOC stock, its spatial variation and the causes of this variation in Limpopo National Park (LNP), a data-poor and poorly accessible area in southwestern Mozambique. During a field survey, A-horizon thickness was measured and soil samples were taken for the determination of SOC concentrations. SOC concentrations were multiplied by soil bulk density and A-horizon thickness to estimate SOC stocks. Spatial distribution was assessed through: i) a measure-and-multiply approach to assess average SOC stocks by landscape unit, and ii) a soil-landscape model that used soil forming factors to interpolate SOC stocks from observations to a grid covering the area by ordinary (OK) and universal (UK) kriging. Predictions were validated by both independent and leave-one-out cross validations. The total SOC stock of the LNP was obtained by i) calculating an area-weighted average from the means of the landscape units and by ii) summing the cells of the interpolated grid. Uncertainty was evaluated by the mean standard error for the measure-and-multiply approach and by the mean kriging prediction standard deviation for the soil-landscape model approach. The reliability of the estimates of total stocks was assessed by the uncertainty of the input data and its effect on estimates. The mean SOC stock from all sample points is 1.59 kg m^{-2}; landscape unit averages are 1.13–2.46 kg m^{-2}. Covariables explained 45% (soil) and 17% (coordinates) of SOC stock variation. Predictions from spatial models averaged 1.65 kg m^{-2} and are within the ranges reported for similar soils in southern Africa. The validation root mean square error of prediction (RMSEP) was about 30% of the mean predictions for both OK and UK. Uncertainty is high (coefficient of variation of about 40%) due to short-range spatial structure combined with sparse sampling. The range of total SOC stock of the 10,410 km^{-2} study area was estimated at 15,579–17,908 Gg. However, 90% confidence limits of the total stocks estimated are narrower (5–15%) for the measure-and-multiply model and wider (66–70%) for the soil-landscape model. The spatial distribution is rather homogenous, suggesting levels are mainly determined by regional climate.

1. Introduction

Soil organic carbon (SOC) drives natural soil fertility and is a common indicator of livelihoods and ecosystem functions. It has been a focus of attention in the context of both agricultural development and carbon sequestration. Under the various United Nations protocols, there is an increasing need for accurate estimates of SOC stocks at national and sub-national scale to aid policy makers in making land use and management decisions (Milne et al., 2007). Estimates of current SOC stocks and their spatial variation are the starting point for the estimation of the carbon sink capacity and SOC sequestration. The focus of the study determines the type of data required. In the case of climate change, estimates of total SOC stocks are important for mitigation purposes. However, when carbon payments are considered, the spatial distribution of stocks and their respective change become important (Antle et al., 2007).

Techniques for estimating SOC stocks have been grouped into two categories (Mishra et al., 2010; Thompson and Kolka, 2005): (1) the measure-and-multiply approach and (2) the soil-landscape modeling approach. In the measure-and-multiply approach the study area is stratified. Point measurements per stratum are averaged and multiplied by the area of each stratum of maps that stratify (Guo et al., 2006; Tan et al., 2009; Thompson and Kolka, 2005). Soil survey maps and field observations are primary resources to estimate SOC stocks with the measure and multiply approach that has been applied from regional (Amichev and Galbraith, 2004; Batjes, 2008; Tan et al., 2004; Thompson and Kolka, 2005) to global (Batjes et al., 2007) scales. The approach has the advantage of being simple, though it is not exempt of several limitations like potentially high within-stratum SOC variability

* Corresponding author. Tel.: +258 843285400.
E-mail address: armindo.cambule@uem.mz (A.H. Cambule).

0016-7061/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geoderma.2013.07.015

(Mishra et al., 2010; Thompson and Kolka, 2005). The soil landscape modeling approach analyzes the spatial variability of SOC stocks with respect to variations in environmental covariables such as topography, land use or climate (Mishra et al., 2010). A model is built based on the various environmental covariables covering the entire study area plus limited number of field observations of SOC stocks, and is used to make predictions over a grid across the study area (Gessler et al., 2000; Thompson et al., 2001). These are then summed to an area total. Examples of use of this approach are many, e.g., Ungaro et al. (2010) and Ziadat (2005), though many have successfully been applied to small areas (<100 ha) and using of digital elevation models as the covariate, e.g. Florinsky et al. (2002), Bhatti et al. (1991) and Gessler et al. (2000).

The soil-landscape approach may result in a lower estimation error at each prediction location, due to the use of complete spatial coverages of secondary information, i.e., the environmental covariables. The measure-and-multiply approach has the advantage of simplicity, although within-stratum variability (heterogeneous strata) limits precision (Aubry and Debouzie, 2000; Mishra et al., 2010; Thompson and Kolka, 2005). Further, the soil-landscape approach produces a grid map of SOC stocks whereas the measure-and-multiply approach produces a chloropleth map with an average value per stratum. It is not clear a priori which method gives lower estimation errors for total stocks.

In 2001, Mozambique declared an area known as "Coutada 16" (hunting zone) the Limpopo National Park (LNP), which forms part of a trans-frontier park with South Africa and Zimbabwe. The LNP provides ecosystem services and supports the livelihoods of about 20,000 people living within its boundaries. The formation of LNP and the planned relocation of the communities within the park will result in major land use changes, both in terms of vegetation and wildlife (Ministerio do Turismo, 2003). These changes are expected to affect SOC stocks in and around the LNP, including in resettlement areas where SOC stocks are a major contributor to soil fertility. Any change cannot be assessed without a proper baseline, i.e., present-day stocks. Therefore, the aim of this study was to quantify the total SOC stock and its spatial variation in the Limpopo National Park, and the probable causes of any variation. Further, we wanted to compare the various approaches to estimating SOC stocks.

2. Study area

The study area of 10,410 km^{-2} covers most of Limpopo National Park, which is one part of the study area of the "Competing Claims on Natural Resources" project (Giller et al., 2008), centered on the trans-frontier national parks of the Mozambique–Zimbabwe–South Africa border. LNP is located in Mozambique (Fig. 1) between 22° 25′ and 24° 10′ S and 31° 18′ and 32° 38′ E. Altitudes range from about 50 to about 500 m above sea level (Stalmans et al., 2004). It has a warm arid climate (BWh, Köppen classification) with a dry winter and mean annual temperature exceeding 18 °C (Peel et al., 2007). Absolute maximum temperatures (between November and February) increase northwards to above 40 °C. Annual rainfall decrease northwards from above 500 mm in the southeast to about 350 mm at the extreme north (Ministerio do Turismo, 2003; Stalmans et al., 2004).

The dominant lithology is the extensive Quaternary aeolian sand cover along the NNW–SSE spine of the park. Tertiary sedimentary rocks (limestones, sandstone) are found close to the drainage lines where the sand mantle has been exposed. Rhyolite rocks from the Karroo formation are located along the western border while alluvium

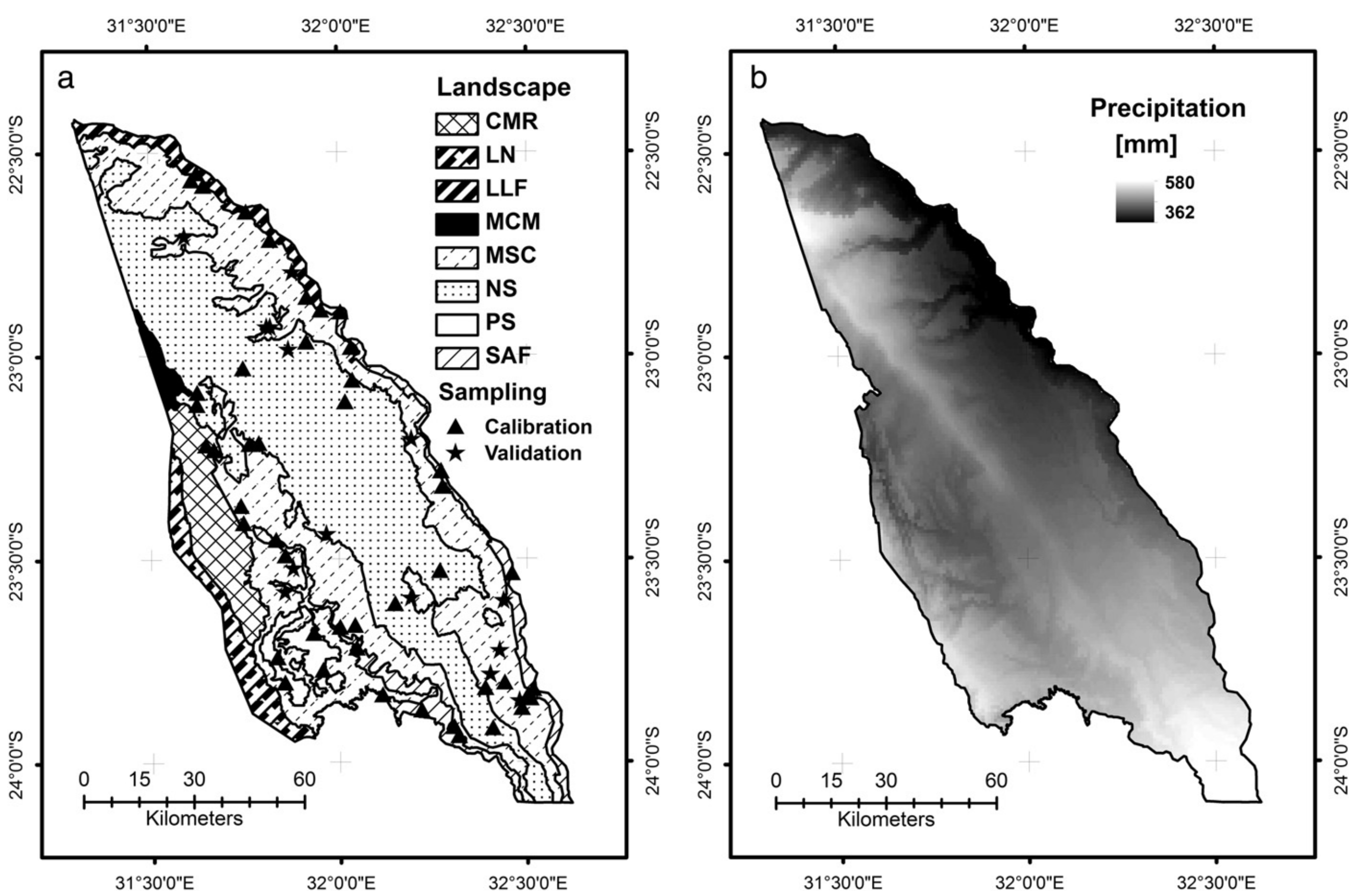

Fig. 1. (a) LNP landscapes and sampling points; (b) annual precipitation. Landscapes after Stalmans et al. (2004); precipitation after Hijmans et al. (2011).

lies along the main drainage lines. Soils derived from aeolian sands range from shallow to deep and are sandy, those derived from rhyolite are shallow and clayey, those derived from sedimentary rocks are deep, structured and clayey and those derived from alluvium materials are clayey (Manninen et al., 2008; Rutten et al., 2008; Stalmans et al., 2004).

Very little has been published about the soil distribution in the LNP. The national reconnaissance soil map at 1:1,000,000 scale shows LNP covered by five soil units from five major soil grouping (FAO and UNESCO, 1997; INGC et al., 2003; INIA, 1995); the Arenosols/Haplic Luvisols and Ferralic Arenosols on the Quaternary Aeolian sands, the Eutric Leptosols over the Karroo formation, and the Calcaric Cambisols and Eutric Fluvisols along the main drainage lines. The LNP was classified into ten landscape units by Stalmans et al. (2004) on the basis of plant community composition, their environmental determinants and distribution. For the present study, two small units (<1% of the area) were merged into the larger ones, resulting in eight units (Fig. 1a): Combretum/Mopane Rugged Veld (CMR), Limpopo Levubu Floodplains (LLF), Limpopo north (LN), Mixed Combretum/Mopane woodland (MCM), Mopane Shrubveld on Calcrete (MSC), Nwambia sandveld (NS), Pumbe Sandveld (PS), *Salvadora angustifolia* floodplains (SAF). Given the rocky nature of the LN and MCM units, we assumed their SOC contents to be zero. Most landscapes in LNP are dominated by plant communities with *Colophospermum mopane*. However, this mopane vegetation is not present in the Aeolian sands (PS) and in wetter (LLF and SAF) landscapes along the major drainage lines. The park has only a few improved roads, and access is quite difficult, especially off-road, due to dense vegetation, rough ground, and large wild animals.

3. Methodology

A summary of the methodology follows; later in the section we explain each step in detail. We assessed SOC stocks for sampling points, its variation across the LNP by landscape, and the total SOC stock. First SOC concentrations were converted to SOC stocks at the sampling points using the field measured A-horizon thickness and estimated soil bulk density. To estimate the SOC stocks distribution across the LNP, we have followed two approaches: (a) the measure-and-multiply method where mean stocks are calculated per landscape unit, and (b) the soil-landscape approach where stocks are estimated over a grid using spatial models derived using auxiliary information and limited field sampling. Total stocks were calculated by summing up (a) the estimated stocks of the landscape units and (b) the estimates at each grid cell. In addition, total stocks were estimated based on calculated naïve and spatial means converted to LNP area size. We also assessed the uncertainty of estimates of stocks' spatial distribution by calculating the standard error (SD of the mean) and kriging prediction standard deviation, respectively for the measure-and-multiply and soil-landscape approaches. Uncertainties of estimates of total stocks were obtained by calculating the standard error and mean kriging prediction standard deviation plus the 90% confidence interval. Finally we assessed the reliability of the estimates of total stocks by assessing the uncertainty of the input data and its effect on estimates of total stocks. The results from the various methods are then compared based on their width of confidence interval (Janssen and Heuberger, 1995; Smith et al., 1997; Wösten et al., 2001). Statistical analysis was performed in the R environment for statistical computing (R Development Core Team, 2006).

3.1. Field sampling and laboratory analysis

Field sampling was carried out during the same field campaign as for studies reported earlier (Cambule et al., 2012), following a stratified random design focusing on accessible areas. This design provides a statistically-valid sample with high operational efficiency (De Gruijter et al., 2006). The landscape units described by Stalmans et al. (2004) were used as strata, because they express the integrated effect of soil-forming processes. Soil samples were collected from the entire field-identified A-horizon of variable depth (which was recorded), in a total of 60 clusters. The number of cluster per stratum was proportional to stratum area. A cluster was defined as two orthogonal and mid-way intersecting transects of 720 and 360 m, respectively, along which seven sampling points, were placed 180 m apart. To capture the maximum variation, the longer transect was oriented perpendicular to the slope at the midpoint. A GPS was used to navigate to the coordinates of planned observations. Field positional adjustments were made, wherever necessary, to ensure sample representativeness (for example, avoiding anthills). The field-identified A-horizon was chosen as the sample volume because it is where most biological activities take place and therefore most of the soil carbon is stored (Gessler et al., 2000). In order to ignore very short-range SOC variation, five subsamples from the four corners of a 90×90 m support area plus the center were thoroughly mixed into a composite sample, from which a portion was taken for analysis. This support corresponds to medium-resolution remotely sensed data (e.g., Landsat and SRTM) and to agricultural fields of smallholder farmers.

In the laboratory the samples were all scanned with a near-infrared (NIR) spectrometer (Shepherd and Walsh, 2002) and a sub-set was analyzed for SOC concentrations and particle size (van Reeuwijk, 2002). The sub-set was used to build a Partial Least Square Regression (PLSR) model to translate NIR measurements into SOC concentration which was later used to compute SOC concentration at all observation points. Details on the establishment of the PLSR model are described by Cambule et al. (2012). The PLSR model explained 83% of variation in laboratory-measured SOC concentration.

3.2. Assessing SOC stocks

3.2.1. SOC stocks at the sampling points

To assess SOC stocks at sampling points, SOC concentrations (Table 1) were converted into SOC stocks using the field measured A-horizon thickness and soil bulk density (BD), estimated as 1.44 ± 0.02 g·cm^{-3}. This estimate is based on measurements ($n = 14$) by COBA Consultores (1982) around the confluence area between the Singuedzi and Elephant Rivers and Nhantumbo et al. (2009) on the sandy soils in the extensive NS landscape unit. This average was used instead of estimating BD at each point by a pedotransfer function (PTF) from the measured clay content and SOC concentration, because there is no calibrated PTF for the area and the use PTF developed elsewhere is not appropriate even under similar ecological conditions (Gijsman et al., 2002). The average BD used is consistent with ranges reported in the literature by EUROCONSULT (1989) for the sandy loam to sandy clay loam soil textural classes found in LNP (1.4–1.65 g·cm^{-3}). Our practice is consistent with that of Williams et al. (2008) in the miombo woodlands of central Mozambique, who justified the use of a single value of BD (1.29 cm^{-3}) because of the low variability of BD from 28 composite topsoil samples. Despite the similarity in soil textural classes, their study site is located in a much wetter climate than ours (annual precipitation of about 700 mm vs. 450 mm) as depicted by the much richer miombo vegetation, so the soils with higher organic matter are expected to have lower BD.

Table 1
Summary statistics of SOC concentration, A-horizon thickness and SOC stocks.

SOC	Unit	N	Min	1st Qu.	Med.	Mean	3rd Qu.	Max
SOC concentration*	%	399	0.00	0.61	0.88	0.93	1.20	2.68
A-horizon thickness	cm	399	0.0	10.0	13.0	13.3	17.0	26.0
SOC stock	kg m^{-2}	399	0.00	0.95	1.47	1.59	2.10	5.59
SOC stock, clusters	kg m^{-2}	59	0.51	1.09	1.48	1.62	2.02	3.91

* Source: Cambule et al. (2012).

3.2.2. SOC stocks spatial distribution

3.2.2.1. The measure-and-multiply approach. In this approach, we interpreted the spatial distribution of SOC stocks across the LNP to be a function of the sampling strata, i.e., the landscape map of Stalmans et al. (2004). In this approach an average of each landscape unit was calculated based SOC stocks data from sampling points. The averages were computed by a single-factor ANOVA (R function 'lm'), followed by a pairwise means comparison with pooled standard deviation and the Holm correction for multiple comparisons (R function 'pairwise.t.test') to group and rank landscape units, thus showing the stocks spatial distribution across the LNP as a chloropleth map of single values (with uncertainty) per landscape unit.

3.2.2.2. The soil-landscape approach. Here, we interpreted the spatial distribution of SOC stocks across LNP as a function of soil-forming factors (McBratney et al., 2003). These were represented by explanatory variables derived from readily-available, full-coverage secondary information. Spatial models were developed to describe the variation in SOC stocks in relation to the soil-forming factors. These steps are now described in detail.

3.2.2.2.1. Soil-forming explanatory variables. We followed the framework for digital soil mapping described by McBratney et al. (2003). In the study area, SOC stocks are expected to be related to a number of soil forming factors including rainfall, vegetation, topography, parent material, and soil conditions. Selected secondary data corresponding to these soil forming factors (Table 2) include: (1) the mean annual precipitation on a 30 arc-second (approximately 1 km) resolution grid obtained from the WorldClim website (Hijmans et al., 2011), (2) multi-spectral satellite imagery (Landsat TM, 30 m resolution) for wet and dry seasons obtained at USGS website (www.usgs.gov, preprocessing: L1T level), (3) the 3 arc-second (approximately 90 m) resolution from Shuttle Radar Topographic Mission (SRTM) digital elevation model (DEM), obtained at JPL website (www.jpl.nasa.gov, preprocessing: research grade), (4) a 1:250,000 lithology map developed by the Geological Survey of Finland (Manninen et al., 2008; Rutten et al., 2008) and (5) the 1:1,000,000 scale landscape map of Stalmans et al. (2004) as an integrated soil forming factor. We also considered first- and second-order trend surfaces, which are surrogates for regional change in soil-forming factors.

3.2.2.2.2. Selection of explanatory variables for spatial models. To select the explanatory variables for model building, we first extracted their values at sampling points through map overlay. The SOC stock at these points was linearly regressed on the continuous explanatory variables, and as a one-way or multiway linear partitioning of variance for the categorical variables, both using R function 'lm'. Models were evaluated by ANOVA of the model compared to a null model, and by visual inspection of regression diagnostic plots (Fox, 1997). The highest adjusted goodness-of-fit of models with acceptable diagnostics was used to select explanatory covariables for model building (Moore et al., 1993).

3.2.2.2.3. Spatial structure and models. To assess the spatial structure and scale of SOC stocks variation, we first performed the within- and between-cluster ANOVA, then calculated the respective experimental variograms (Franklin and Mills, 2003; Oliver, 2001; Webster et al., 2006) for the residuals from linear models (obtained in previous section) and original values of SOC stocks. Variogram maps were prepared to visually detect any anisotropy, followed by automatic variogram model fitting using the weighted least square (WLS) method (Pebesma, 2004). In order to minimize irregularities caused by the small sample size and to avoid arbitrary decisions on variogram bin width we applied the residual maximum likelihood (REML) method to estimate sills directly to the variogram cloud starting from the WLS fit (Marchant and Lark, 2007). We also constructed variogram models of the residuals from the feature-space and trend surface models described in Section 3.2.2.2.2.

3.2.2.2.4. Spatial distribution of SOC stocks. The selected spatial models were used for spatial prediction on a 1 × 1 km grid and their results compared. This resolution was chosen as a compromise among the resolutions of the secondary data, and also to account for the practical support, given the scale of spatial variation as revealed by the within-cluster variograms.

3.2.2.2.5. Model validation. Spatial models were validated by leave-one-out cross-validation (LOOCV) as well as by independent validation. The latter was performed by randomly splitting the sample set (70% calibration and 30% validation) and fitting variograms based only on the calibration sample set. Differences between observed and predicted values were summarized as the root-mean squared error of prediction (RMSEP) and the bias of the estimation. Independent validation was compared with the internal measure of goodness-of-fit, i.e., the standard deviation (SD), to assess which model most closely estimates the true error (Goovaerts, 1999; McBratney et al., 2000).

3.2.3. Assessing total LNP SOC stocks and their uncertainty

The total stock from the measure-and-multiply approach was computed three ways: (1) summing the total SOC stocks of the landscape units (equivalent to landscape unit area weighted average), (2) calculating the naive mean of all observations and multiplying by LNP area and (3) calculating the spatial mean of all observations and multiplying by LNP area. The spatial mean (i.e., without stratification) was computed as the best linear unbiased estimate (BLUE) of the mean, taking into account the modeled spatial structure of the all-sample ordinary variogram (Aubry and Debouzie, 2000). This is the first step in kriging estimation by the Gstat package's 'krige' function (Pebesma, 2004).

The total stock from the soil-landscape approach was computed by summing the interpolation grid. Prediction uncertainty was expressed as 90% confidence intervals based on prediction standard deviations. For the measure-and-multiply approach these were calculated in each landscape unit from the standard errors of each unit's mean; for the soil-landscape approach by summing the grid cells' kriging standard deviation.

3.2.4. Assessing the reliability of total SOC stocks estimates

The sources of uncertainty affecting SOC estimates are (1) field measurement of A-horizon thickness, (2) laboratory analysis of SOC (3) spectral measurements of soil samples, (4) PLSR models used to predict SOC of samples measured by spectroscopy, (5) estimation of bulk density, and (6) sampling density. For each their reliability was discussed based on the technique followed and numerical measures of consistency. At sampling points we assessed uncertainty in measured A-horizon thickness by calculating the SD. We assessed the uncertainty of BD by assigning a range from the literature. Uncertainty in SOC concentration was taken from previously reported work by (Cambule et al., 2013). At spatial distribution level we assessed uncertainty of A-horizon thickness, SOC (concentration) by calculating the standard error and mean kriging prediction standard deviation for the measure-and-multiply and soil-landscape model approaches, respectively. We then assessed the reliability of estimates of total SOC stocks by checking whether their 90% confidence interval cover the effect of uncertainties from these three inputs.

Table 2
Summary statistics of explanatory variables.

Variable	unit	Min	Max	Range	Mean	SD
Elevation	m.a.s.l.	54	531	477	241	99
Flow accumulation	nr. pixels	0	50	50	4	8.2
NDVI wet season	–	−1.0	0.69	1.69	0.35	0.13
NDVI dry season	–	−0.34	0.56	0.91	0.11	0.08
Annual precipitation	mm	362	580	218	461	40

Source: Cambule et al. (2013).

4. Results and discussion

4.1. SOC stocks at sampling points

The histograms of SOC concentration, A-horizon thickness and that of derived SOC stocks at sampling points are shown in Fig. 2a–c. Whereas the histograms of concentrations and stocks are right-skewed, that of A-horizon thickness is approximately normally-distributed. All histograms span over a wide range (SOC concentration: 0–2.7%; A-horizon thickness: 0–25 cm and SOC stocks: 0–5.6 kg m^{-2}), for this semi-arid environment. SOC concentration and A-horizon thickness are negatively-correlated ($r = -0.44$, Fig. 3a); there is thus a compensation effect: total stock is less variable than concentration, because soils with lower concentrations tend to have thicker A-horizons, and vice-versa. This suggests that total stock is mostly controlled by the general climate and vegetation of the area, whereas A-horizon thickness and SOC concentrations vary with local site factors. Thus, the expected positive correlations between SOC stock and A-horizon ($r = +0.40$, Fig. 3b) and SOC concentration ($r = +0.57$, Fig. 3c) are only moderate. Both relations are poor for high SOC stocks.

The fitted variogram for A-horizon thickness (see parameters in Section 4.2.2) shows that its range of spatial dependence almost matches the cluster size. About two-thirds of the variance is spatially-dependent (structural sill vs. total sill). Given this good within-station spatial structure of A-horizon thickness (parameters in Section 4.2.2) as well as the substantial (30%) and mostly random within-cluster SOC variation revealed by one-way ANOVA, the low correlation can be attributed to the high short-range variability of SOC due to local factors such as animal activity and vegetation patches (Cambule et al., 2013).

4.2. Assessing SOC stocks spatial distribution

4.2.1. The measure-and-multiply approach

Summary statistics of A-horizon thickness, SOC concentration, and SOC stocks by landscape unit, as well as grouped boxplots of these, are shown in Table 3 and Fig. 4. Pairwise mean differences of A-horizon thickness from one-way ANOVA show that landscapes units CMR, LLF and MSC form one group, with thinner A-horizons, and NS, PS and SAF another group; overall explained variation is 19.5%. Thus, the sandier upland (PS and NS) and the low lying floodplains (SAF) soils tend to have thicker A-horizons (Fig. 4a) and correspondingly lower SOC concentrations (Fig. 4b).

Boxplots of SOC stocks by landscape unit (Fig. 4c) depict a rather lower landscape influence as compared to SOC concentration; only 13.3% vs. 33.9% variance explained by a one-way ANOVA. The resulting chloropleth map produced by reclassifying the map units with the mean SOC stock resembles the landscape units with same sharp boundaries (see further Section 4.2.2.4).

Pairwise mean differences from the ANOVA showed that the extensive NS has distinctly lower mean SOC stock than all others but the LLF; this latter however is not distinguishable from the others. Thus the fairly homogeneous distribution of stocks across LNP is explained by the negative correlation between SOC concentration and A-horizon thickness. That is, large SOC stocks may have either thick A-horizons or high SOC concentrations (and vice-versa), but rarely both. Within-landscape variation is, however, considerable as shown by the coefficients of variation (Table 3); from an overall CV of about 60%, CMR has the lowest and MSC the highest. This heterogeneity may be due to local differences in soil-forming processes that were not recognized or mappable by Stalmans et al. (2004). The differences in sample size per landscape unit also affect the computation: smaller sample sizes give less reliable statistics. The summary ANOVA table is shown in Table 4.

4.2.2. The soil-landscape approach

4.2.2.1. Explained SOC stocks variation. The proportion of SOC stocks variation of all observations explained by the *scorpan* covariables is about 13.5% (landscape units, i.e., the integrated soil forming factor), 9.5% (coordinates, i.e., geographic trend) and 45% (sampling clusters, i.e., soil factor). Other covariables explained lesser variation, so that a spatial model based on them would not be helpful. When cluster averages were considered, only the coordinates showed increased explanatory power, from 9.5% to about 17%. The obtained amounts of explained variation, from both numerical and categorical explanatory variables, are substantially lower than those obtained for SOC concentration Cambule et al. (2013). That is, the SOC stocks are less variable than SOC concentrations across the study area, which suggests that stocks are mostly in equilibrium with regional climate whereas concentrations vary more with local factors.

4.2.2.2. Spatial structure. The result of the within- vs. between-cluster ANOVA shows that the clusters explain about 45% of SOC stocks variation. This is much less between-cluster effect than found for SOC concentration (71.1%) by Cambule et al. (2013). The lower between-cluster stocks variation and its fairly homogeneous spatial distribution (as explained in Section 4.2.1) may be interpreted as the result of a

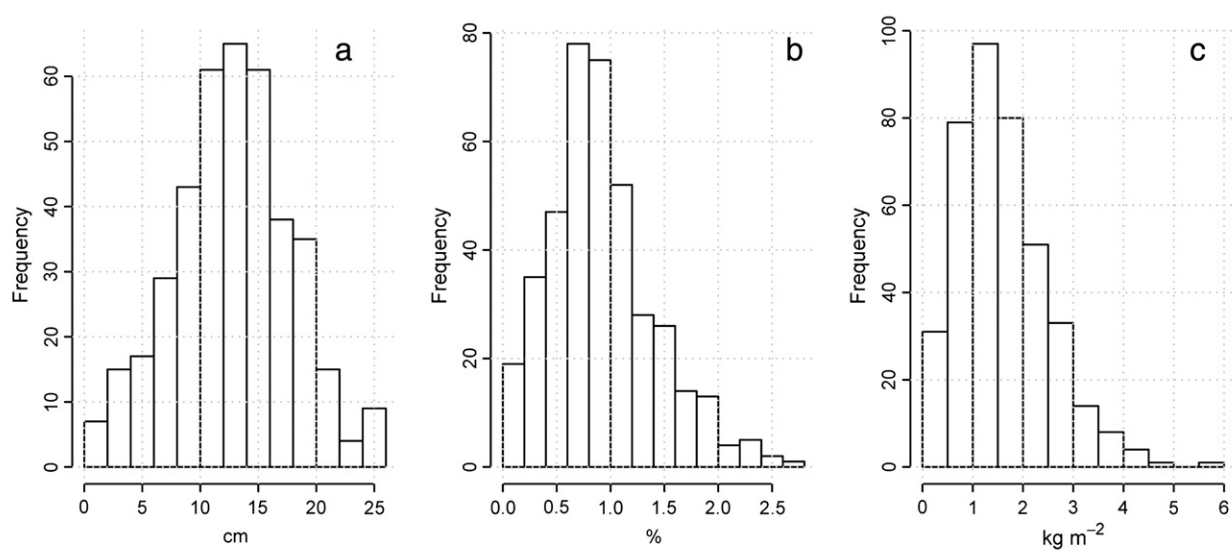

Fig. 2. Histograms of (a) A-horizon thickness, (b) SOC concentrations and (c) SOC stocks, all at sampling points.

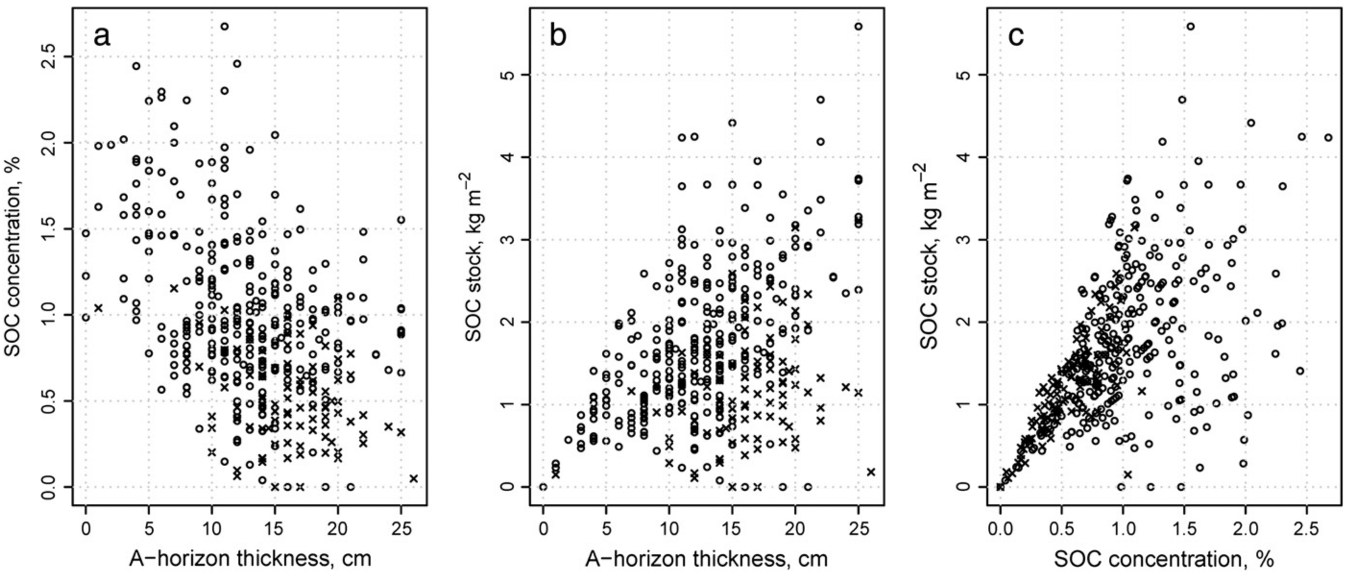

Fig. 3. (a) SOC concentration and (b) stocks as a function of A-horizon thickness; (c) relation between SOC stocks and concentration. Results for landscape unit NS shown with 'x' symbol.

general equilibrium of SOC stocks with regional climate, i.e., in this environment the net primary production is equal to the litter input to the soil, and the decomposition rate matches these.

Visual assessment of the within-cluster variogram (to 720 m) (Fig. 5a) suggested a very weak spatial dependency to about 300 m with total sill of about 0.44 (kg m^{-2})^2, quite close to the MSE of the within- vs. between-cluster ANOVA (0.446 kg m^{-2}), which is taken as the residual variance. The total sill is equivalent to about 0.4% of SOC concentration (Fig. 3c), about three times the root mean square error (RMSE) of SOC determination on the base of duplicate samples (0.13%). When the experimental variogram was fitted with a pentaspherical function using WLS fit, an unrealistic zero nugget resulted; this was not improved by the REML. The apparent visible spatial dependence could not be modeled; therefore a pure nugget (0.436 kg m^{-2}) variogram was fitted by WLS (Table 5). Consequently the within-cluster SOC stocks variation (about 55%) can be considered as spatially random, i.e., caused by local unmapable factors with high very short-range spatial variability (Janzen and Ellert, 2002; Mapa and Kumaragamage, 1996). Thus the nugget found in the long-range variogram represents a support of at least a cluster. Therefore the remainder of the analysis is based on the cluster averages.

The long-range (beyond cluster) ordinary experimental variogram showed local spatial dependence to a range of about 20 km (Fig. 5b). The WLS and REML fitted spherical variograms (separations <20 km) both showed reasonable structures, with spatial autocorrelation to about 11 km separation (Table 5). This range is about a quarter of the east–west dimension of the LNP, so from the present sample distribution predictions for most of the unsampled areas can only be made by the (spatial) mean, thus giving little insight into spatial variability of SOC stocks. The spherical model was selected based on the patchy structure of spatial variability exhibited by most soil properties.

The nugget effect is about 12% of total sill, indicating a high proportion of autocorrelation (Mapa and Kumaragamage, 1996). This low nugget is explained by the averaging effect of clusters, since when the variogram is based on all observations rather than cluster averages, the nugget effect raises to about 43% of the total variance. This is interpreted as the effect of uneven spatial distribution of the "organisms" soil-forming factor (e.g., woody- and non-woody vegetation, termites) as plant production in semi-arid regions depend on small differences on water availability, runoff, infiltration and storage, whose combination results in a very large variation in vegetation and soil properties over small areas (Janzen and Ellert, 2002; Martius et al., 2001; Tiessen and Santos, 1989; Wang et al., 2009); however when this is averaged over a cluster, this variation largely is averaged out.

The REML-fitted residual variogram (Table 5, Fig. 5c) for the residuals from a first-order polynomial (the best explanatory variable, representing "spatial position" or local trend) had a range of about 10 km and a nugget effect of 29% of the total sill. Higher-order trend surfaces resulted in lower adjusted goodness-of-fit, and had no obvious interpretation so were not considered. Despite the weak spatial model, a large proportion of spatial dependence spans across a substantial range (>10 times the cluster length; 720 m),

Table 3
Summary statistics by landscape of SOC concentrations, A-horizon thickness and SOC stocks.

SOC	Unit	CMR	LLF	MSC	NS	PS	SAF
Number of samples	–	14	22	197	98	15	52
SOC concentration*	%	1.90	0.91	1.06	0.51	0.92	0.93
SOC concentration SD	%	0.36	0.20	0.47	0.26	0.33	0.41
SOC concentration CV	%	18.9	22.0	44.3	51.0	35.9	44.1
Area size	km^2	689	264	4058	4514	253	637
A-horiz. thickness mean	cm	9.07	11.11	11.59	15.82	16.53	15.07
A-horiz. thickness SD mean	cm	0.86	0.75	0.36	0.41	0.82	0.66
A-horiz. Thickness CV	%	35.5	31.5	43.1	25.7	19.1	29.6
SOC stock mean	kg m^{-2}	2.46	1.50	1.62	1.13	2.02	2.05
SOC stock SD mean	kg m^{-2}	0.25	0.14	0.07	0.06	0.15	0.13
SOC stock CV	%	38.6	43.1	57.1	53.5	28.1	46.6
SOC total stocks	Gg	1695.9	395.4	6587.7	5081.8	510.3	1307.4

* Based on data from Cambule et al. (2012).

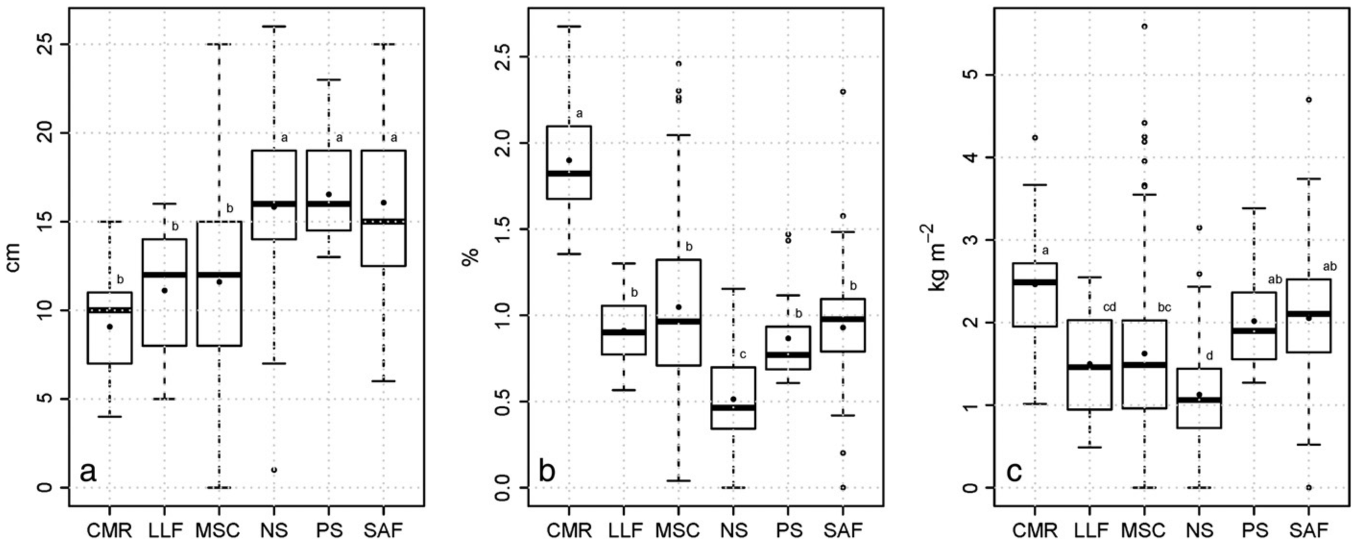

Fig. 4. Relation between landscape units and (a) A-horizon thickness, (b) SOC stocks and (c) concentration. Letters show significant differences in means by Duncan's new multiple range test, and the mean is shown by a dot.

although this range is short relative to the LNP dimensions; thus similarly to the ordinary variogram model, most of unsampled area is predicted by the trend surface, a marginal improvement over the spatial mean. None of beyond-cluster variogram maps showed anisotropy.

4.2.2.3. Selection of spatial prediction model. Based on the above result we had therefore the following options for the spatial model: ordinary kriging (OK) considering only the SOC stocks from sampling clusters (soil), and universal kriging (UK) with the soil-forming explanatory variable "spatial position" (the coordinates) determining the clear but weak trend. UK is a combination of the standard model of multiple linear regression and the geostatistical methods of ordinary kriging the (trend) residuals (McBratney et al., 2000). The fitted first-order polynomial (plane) represents the trend, whose coefficients (slope in each direction) indicated decrease towards the NNW. This is interpreted as the effect of decreased annual precipitation and longer dry season in this direction, moving away from the Indian Ocean.

4.2.2.4. SOC stocks distribution in LNP. Following the demonstration that within-cluster stocks variation has no spatial structure, cluster averages were used as "points" and therefore the block-kriging was implemented as "punctual" kriging over the practical support of 1 × 1 km grid, a similar size to the 720 × 720 m clusters. The summary statistics of the SOC stocks across the park by the soil-landscape modeling approach is shown in Table 6.

The two spatial models predicted similarly as the means differ by about 2%. However OK has larger extreme values, by about 14% and 7% in the lower and higher end, respectively. The maps (Fig. 6a and c) show clear hot and cold spots. These are unlikely to be true hot/cold spots, rather, the result of limited sampling density relative to the variogram range; thus areas between apparent hot/cold spots are predicted by the spatial mean, resulting in the "pock-marked" map. Both kriged maps show a smooth surface, by contrast to the chloropleth map from the measure-and-multiply approach (Fig. 6b). The UK map (Fig. 6c) shows a clear but weak NNW–SSE trend (especially in the higher predictions in the SE corner) and fits well the moderate drop-off in rainfall (Fig. 1b), although adjusted best-fit of linear model of stocks on annual precipitation were not as good as the trend surface. The precipitation surface also takes into account the modest elevation differences (approx. 150 m), which apparently do not improve the relation with SOC.

Similarly the uncertainty in the estimates (Fig. 7) is high (CV about 40%) and as is usual for kriging, is much lower near observation points; this effect is more pronounced in OK than UK. In the former, SD is as low as 20% of the mean prediction closer to sampling clusters, rapidly increasing to the maximum SD over most of the study area. The uncertainty of the UK estimates follows a similar pattern, however, with more gradual changes due to the trend surface. The high uncertainty is mainly due to the low sampling density relative to the short-range spatial variation.

4.2.2.5. Model validation. Validation statistics are presented in Table 6. The RMSEP determined by LOOCV is about 44 and 43% of the median of predicted SOC stocks by OK and UK, respectively and therefore poor. OK and UK RMSEP are respectively about 6 and 8% higher than their mean kriging SD, which is therefore a slightly over-optimistic estimation of the actual error.

Refitted spatial models for independent validation were constrained by the reduced number of point-pair within the effective range and the corresponding necessity to make wide bins for the experimental variograms. However, the resulting variograms showed a reasonable structure, with ranges similar to the all-cluster averages variograms.

Table 4
Summary of ANOVA by Duncan's new multiple range test for A-horizon thickness, SOC concentration and TOC stocks by landscape: $\alpha = 0.05$ and df = 397.

Variable	Landscape	Mean	Std. error	Group
A-horizon thickness (adj. $R^2 = 19.5$)	CMR	9.07	0.86	a
	LLF	11.11	0.75	a
	MSC	11.59	0.36	a
	NS	15.82	0.41	b
	PS	16.53	0.81	b
	SAF	16.06	0.66	b
SOC concentration (ad. $R^2 = 33.9$)	CMR	1.90	0.10	a
	LLF	0.91	0.04	b
	MSC	1.05	0.03	b
	NS	0.51	0.03	b
	PS	0.87	0.07	b
	SAF	0.93	0.06	c
TOC stocks (adj $R^2 = 13.3$)	CMR	2.46	0.25	a
	LLF	1.50	0.14	ab
	MSC	1.62	0.07	ab
	NS	1.13	0.06	bc
	PS	2.02	0.15	cd
	SAF	2.05	0.13	d

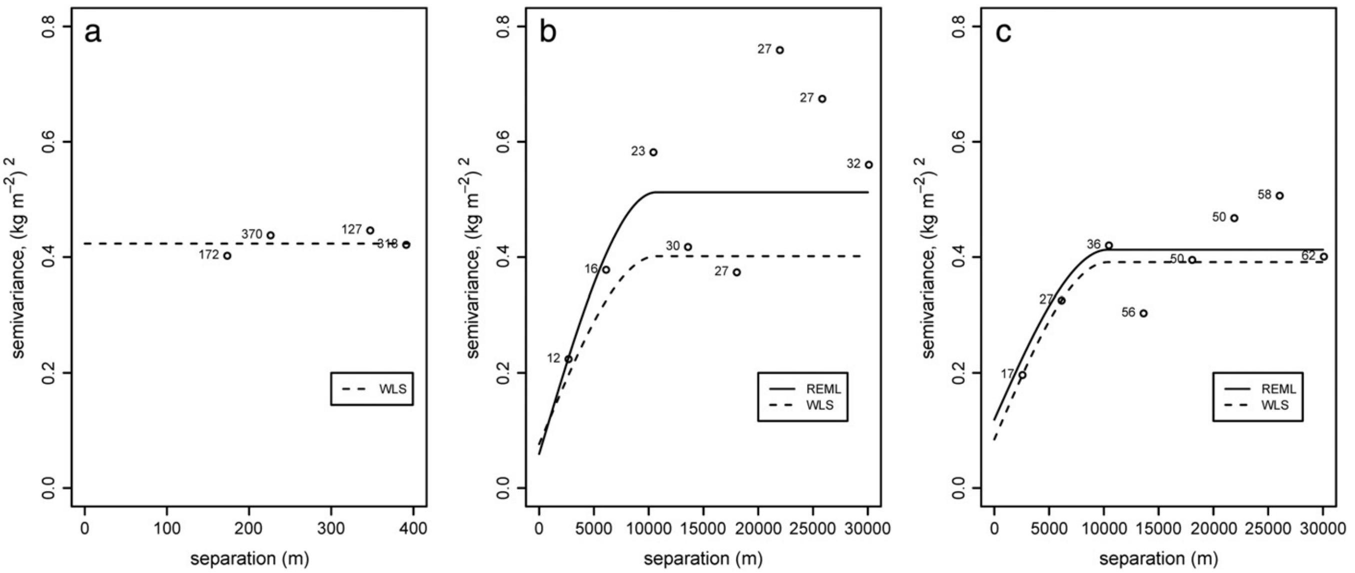

Fig. 5. Empirical and fitted variogam models (a) within-cluster, based on all sampling points, (b) between-cluster, based on cluster averages; (c) residuals from first-order trend surface, based on cluster averages.

Nugget was however set to 0.0 as REML fit resulted in an unrealistic negative value (Table 3). The zero nugget corresponds to the averaging effect of clustered samples.

The RMSEP of the independent validation of both spatial models are about 30% of the median predicted SOC stock. When comparing with mean kriging SD, RMSEP is, in both cases, about 35% lower, so kriging SD is a pessimistic measure of the actual prediction error. The models are also biased towards over-prediction, though UK is slightly less so.

There are many published studies which estimate SOC stocks; however, in most of these the proportion of the error relative to the range of the predictions is not discussed. Of those that do, our results appear to be slightly better than those obtained in large areas, and as good as those obtained from models applied to smaller areas. In large areas, Mendonça-Santos et al. (2010) estimated the SOC of Rio de Janeiro state (Brazil) in an area of about 44,000 km^{-2}, also following the *scorpan*-SSPFe framework. Their results show SOC stocks strongly correlated with landscape units and their final map is dominated by SOC stocks <2.5 kg m^{-2} and a RMSEP of about 1.2–1.4 kg m^{-2}, i.e., a CV of 50%. Similarly, Mishra et al. (2010) predicted SOC stocks for an area of about 650 00 km^{-2} in seven Midwestern states of the USA, following three methods: multiple linear regression, regression kriging and geographically-weighted regression, obtaining a proportion of RMSE to mean prediction of 103%, 69% and 68%, respectively. Scott et al. (2002) estimated SOC stocks for all of New Zealand based on soil moisture and temperature regimes and land use. Their findings also show a RMSE proportion of about 44% relative to the mean predictions in sandy soils, but much better (7%) for soils with low-activity clays.

In small areas the precision of estimates is somewhat better. For example, Minasny et al. (2006) estimated SOC stocks in the lower Namoi Valley (1 500 km^{-2}) in Australia following the *scorpan*-SSPFe framework. Their estimates were mostly in the range 2–9 kg m^{-2} (to 1 m depth) with CV of 30–140%. Simbahan et al. (2006) estimated SOC stocks in Nebraska for fields of about 50–65 ha using OK, kriging with external drift, regression kriging, and co-kriging, and estimated SOC stocks from 4 to 7 kg m^{-2} with RMSE from 1.1 to 1.3 kg m^{-2}; for a CV from 17 to 30%. Our estimates over a much larger (10,415 km^2) area are similar to these.

However, our soil-landscape modeling approach was not very precise: a RMSEP about 30% the prediction median for both OK and UK. This may be due to the more complex soil-landscape relations at larger areas of the already highly (spatially) variable soil carbon in the landscape (Janzen and Ellert, 2002), justifying the general less precise estimates also found in the literature. It may also be a result of the low sampling density. Nevertheless, in data-poor and poorly accessible areas like our study area, achieving comparable results to those from smaller areas shows that the approach is as promising as those based on more comprehensive sampling.

4.3. Total SOC stocks estimates and their uncertainty

The estimates of total SOC total stocks in LNP are presented in Table 7. The results reveal a difference of about 15% between applied methods, being that obtained by summing the landscape totals higher than that from the spatial mean.

However, all estimates of the area-normalized mean stock are in a narrow range, 1.59–1.62 kg m^{-2}, which is comparable to those reported in the literature for southern Africa. A review by Vågen et al. (2005) reports values for southern Africa savanna soils (to 30 cm depth) between 1 and 1.3 kg m^{-2} (sandy) and 1.44 to 2 kg m^{-2} (clay). Williams et al. (2008) studied the SOC stocks of the eastern miombo

Table 5
Nugget, structural sill and range of REML fitted variogram model parameters of SOC stocks.

Variogram type	Nugget (kg m^{-2})^2	Structural sill (kg m^{-2})^2	Range (m)
Ordinary, SOC stock (within-cluster) *	0.436	0	0
Ordinary, SOC stock (between-clusters)	0.059	0.453	10,692
Residual from 1st order trend, SOC stock	0.119	0.294	10,362
Ordinary, A-thickness (within-cluster)	6.089**	12.652**	788

* WLS fitted variogram.
** unit in cm^2.

Table 6
Summary statistics of kriging predictions, kriging prediction standard deviation (SD) and validation results of SOC stocks (kg m^{-2}).

Model	Prediction					Cross-validation		Validation	
	Min	Median	Mean	Max	SD	RMSEP	Bias	RMSEP	Bias
OK	0.71	1.63	1.62	3.53	0.68	0.72	0.01	0.51	−0.35
UK	0.81	1.59	1.59	3.29	0.64	0.69	0.01	0.49	−0.26

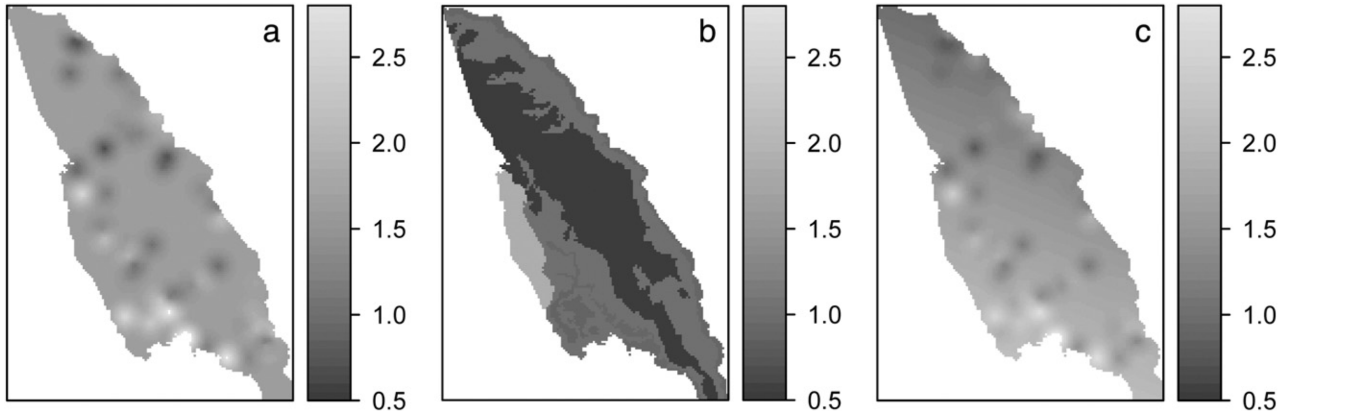

Fig. 6. SOC stocks spatial distribution as predicted by (a) ordinary kriging, (b) landscape unit mean and by (c) universal kriging.

woodlands soils in Mozambique (to 30 cm depth) and found stocks of about 1.8 to 14 kg m^{-2}. The values are relatively high and may be explained by high leaf litter from leguminous trees (*Brachystegia spiciformis*), typical of rich miombo vegetation. Ryan et al. (2011) estimated SOC stocks (to 50 cm depth) in the same vegetation type in Gorongoza District (central Mozambique) at 13.3 kg m^{-2}. Our study area is in a much drier environment with less dense vegetation and that of leguminous trees so our figures are much lower.

Our estimates of SOC stocks by the different methods differ by about 15%, while estimates of the mean SOC stocks only differ by about 2%. This is because of the relative area covered by the different landscape units, each with a different mean stock. This is also corroborated by the wide range between the extreme values of SD (93%), the naive mean having the least dispersion, in contrast to uncertainty estimates from prediction by the spatial models, of which OK is the worse. The naive mean does not account for spatial correlation of the clusters, and thus underestimates the variance, which is thus more realistic when estimated by the spatial mean.

The different approaches result in different confidence intervals. The spatially-explicit kriging-based methods of the soil-landscape approach have very wide intervals, 69% (OK) and 65% (UK) of the mean prediction. This is because each grid cell has an uncertainty, and these are not pooled, as in the measure-and-multiply or means approaches. The narrow interval for the naive mean is probably too optimistic, because it does not consider the spatial correlation between observations. Thus the confidence interval based on the spatial mean (15% of the mean prediction) is preferred if only the total stock, not accounting for spatial distribution either over a grid or by landscape unit, is wanted.

Despite the differences in confidence intervals and totals, when one is interested in total stocks the measure-and-multiply approach is sufficient given the similarity in the estimates of total stocks. It has also the advantage of not requiring a variogram, which may be difficult to model from a small sample. On the contrary, when interested on the spatial patterns of stocks, then the soil-landscape models are required.

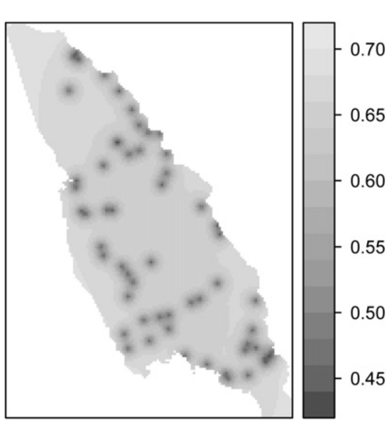

Fig. 7. Kriging prediction standard deviation by universal kriging.

4.4. Reliability of SOC stocks estimates and potential Improvements

This section discusses the effect of each input on estimates of total stocks, and how each could be improved.

The measurements of A-horizon thickness showed a normal distribution, with a precise estimation of the mean (SD–mean about 2% of the mean), narrow 95% confidence limits (4% of the mean) and few extreme values (5% trimmed mean within 95% confidence limits). The field method is simple and reproducible with an estimated precision of <0.5 cm in this landscape. There seems to be little room for further improvement, thus its impact on the reliability of estimates is minimal.

Laboratory results reported by Cambule et al. (2012), show that SOC concentration measured on 129 samples has an SD of the mean just 6% of the mean, a 95% confidence interval of 12% of the mean, and 5% trimmed mean within 95% confidence limits. Further, RMSE on twenty duplicate samples for quality control was 0.13% SOC, about 15% of the mean and similar to the expected precision for the Walkley–Black method. Again the room for improvement is small. Note however that this results do indirectly influence the PLSR predicted SOC concentration (see further on).

The spectrometer used to scan the soil samples has an internal validation test and its spectrum is calibrated before each scan to an internal gold reference. Spectra were only read when internal test was positive. Soil samples were uniformly prepared, and duplicates showed almost no difference in spectra. So it is expected that uncertainty derived from spectroscopy are minimal.

Laboratory-measured SOC concentrations were used to build a PLSR calibration model with which estimates were made for the remaining 283 samples. The summary statistics for the entire sample set showed that the SD of the mean was only 3% of the mean (down from the 6% for the laboratory sub-set), the 95% confidence limits were about 5% of the mean and that the 5% trimmed mean fell within the 95% confidence

Table 7
Mean SOC stocks, their uncertainty and total stocks, following different approaches.

Approach	Method	Mean (kg m^{-2})	SD (kg m^{-2})	TOC stock (Gg)	90% conf. limit (Gg)
Measure and multiply	Landscape	1.59	0.45	15,579	±1433
	Naive mean	1.59	0.05	17,828	±831
	Spatial mean	1.60	0.14	17,908	±2669
Soil-landscape modeling	OK	1.62	0.68	16,858	±11,663
	UK	1.59	0.64	16,545	±10,892

limits of original data. The calibration model had a RMSEP of 0.32% SOC, corresponding to 15% the mean. This uncertainty includes that from the lab analysis of the 129 samples. If the mean stocks would change by 3%, it would still be within the SD-error so it is not expected to affect the 90% confidence intervals. Therefore the estimates made based on spectroscopy can be considered reliable.

Bulk density was used to convert SOC from volume concentration to weight and therefore both uncertainty as well as reliability of SOC stock estimates is affected by the BD. In our study the spatial distribution of BD is not known, nor is there a known relation with landscape unit. Only a single value of soil bulk density (1.44 ± 0.02 g·cm^{-3}), derived from nearby measurements and checked against literature, was used. However, given the limited range of SOC concentration (0–2.68%, Table 2) and that of the textural classes (sandy loam to sandy clay loam) from the study area, it seems unlikely that BD could be outside the range 1.4–1.65 g·cm^{-3} (EUROCONSULT, 1989). This would correspond to maximum variation in estimated BD of about 15% which would affect the SOC stocks estimates by the same amount. This is more than the SD for the naive and spatial means (SD < 9%) and therefore one could expect that the 90% confidence intervals of total SOC stocks estimated by these methods would also be exceeded. This is then the most unreliable part of the estimate. However, this is a worst-case situation: from our estimate of 1.44 g·cm^{-3} to the upper limit 1.65 g·cm^{-3}, which would correspond to sandier soils across the entire LNP. Our estimate is based on not only somewhat heavier soils (sandy clay loams) near the confluence of the Singuedzi and Elephant Rivers, similar to the *S. angustifolia* Floodplains (SAF) map unit, but also from sandy soils from the NS map unit. Thus, it seems unlikely that the true mean BD as high as this upper limit, and so the reliability of our estimate may not be as poor as this worst-case.

Despite the quality control in measurements, the successful PLSR calibration model and the validation statistics for spatial prediction models, the kriged maps have high uncertainty away from sampling locations, and so the reliability of the kriged maps of SOC stock spatial distribution is questionable. This results from a combination of low sampling density and short spatial autocorrelation range relative to study area dimensions. Bulking within-cluster samples at the target grid resolution of 1 × 1 km would remove the high within-cluster variability. Thus future sampling can be more efficient: only one-seventh of the observations are needed, although some time must be taken in adequately covering the block with a composite sample.

With a known variogram, reducing uncertainty in SOC stocking mapping can be aided by the "optimal sampling scheme for isarithmic mapping" (OSSFIM) approach through which the target kriging prediction standard error can be achieved by either (1) reducing the sampling spacing or (2) making predictions over larger blocks (McBratney and Webster, 1981; McBratney et al., 1981). To lower the uncertainty to, e.g., a coefficient of variation of 30% or less (SD ≤ 0.48 kg m^{-2} on a mean of 1.59 kg m^{-2}), observations would have to be made at maximum 4 km interval for cover the whole area by block universal kriging on 1 km blocks (total of 700, Table 8). Similar SD can be obtained by predicting block averages on 5 × 5 km blocks based on 138 points spaced 9 km apart, i.e., spatial resolution is traded for efficiency. A blocks size smaller than 1 × 1 km has too much short-range variability to map without very intensive sampling; whether a 1 × 1 km, 5 × 5 km or larger block is needed depends on the minimum decision area for management. There is a limit to the efficiency gain, since ground must in any case be traversed. Saving time by bulking the within-cluster samples would make possible to reach further away sampling points.

Table 8
Optimal sample spacing for an uncertainty (CV) target of 30% (mean = 1.59 kg m^{-2}).

Block size (side, m)	Widest spacing (m)	Achieved SD (kg m^{-2})	Number of observation points required
1000	4000	0.4361	700
2000	5000	0.4475	448
4000	7000	0.4565	229
5000	9000	0.4730	138
7500	20,000	0.4325	28
10,000	20,000	0.3717	28

The sampling points in the regular grid derived from the OSSFIM approach can further be optimized by the spatial simulated annealing, which also allows the minimization of the kriging variance taking into account existing samples (van Groenigen et al., 1999); this would be a sound strategy for a second phase of sampling, starting from the current phase to achieve a target uncertainty. This would however result in a more spread of sampling points which would cost sampling time in exchange for map quality, though not realistic in poorly-accessible areas.

5. Conclusions

In the present study, we estimated the total SOC stocks in the LNP based on limited data collected from accessible areas and have made use of secondary information covering the entire area. The estimates followed both the "measure-and-multiply" and "soil-landscape modeling" approaches. In the former we used the per stratum mean, the naive mean and the spatial mean, while in the latter the ordinary and universal kriging methods, chosen based on the fact that sampling cluster and regional trend were the soil-forming factors that explained SOC stocks variation the most.

The mean SOC stocks obtained in all methods are very close however, the SD were distinct, with the soil-landscape modeling methods having at least four times as high SD as the maximum SD for the measure-and-multiply ones. The high uncertainty is mainly due to the short-range spatial variation, by the sparse sample, the weak trend, and poor correlation with covariables. It would be difficult to improve on this estimate without intensive sampling.

The total stocks, obtained by summing (1) the landscape averages and (2) the block-universal-kriged estimates of 1 × 1 km averages across the whole study area are similar, and also similar to average estimates in soils of similar texture from southern Africa. The high uncertainty of these estimates limit its use as a baseline, however they may be useful for many agricultural studies.

Acknowledgments

We thankfully acknowledge support from the LNP management for logistic support, The International Research and Education Fund (INREF) of the Wageningen University through the "Competing Claims on Natural Resources Programme" for funding. We would like also to thank the National Agrarian Research Institute (IIAM) for making available the NIR spectrometer (MPA) for soil sample spectral acquisition.

References

Amichev, B.Y., Galbraith, J.M., 2004. A revised methodology for estimation of forest soil carbon from spatial soils and forest inventory data sets. Environmental Management 33, S74–S86.
Antle, J.M., Stoorvogel, J.J., Valdivia, R.O., 2007. Assessing the economic impacts of agricultural carbon sequestration: terraces and agroforestry in the Peruvian Andes. Agriculture, Ecosystems & Environment 122 (4), 435–445.
Aubry, P., Debouzie, D., 2000. Geostatistical estimation variance for the spatial mean in two-dimensional systematic sampling. Ecology 81 (2), 543–553.
Batjes, N.H., 2008. Mapping soil carbon stocks of Central Africa using SOTER. Geoderma 146 (1–2), 58–65.
Batjes, N.H., Al-Adamat, R., Bhattacharyya, T., Bernoux, M., Cerri, C.E.P., Gicheru, P., Kamoni, P., Milne, E., Pal, D.K., Rawajfih, Z., 2007. Preparation of consistent soil data sets for modelling purposes: secondary SOTER data for four case study areas. Agriculture, Ecosystems & Environment 122 (1), 26–34.
Bhatti, A.U., Mulla, D.J., Frazier, B.E., 1991. Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and thematic mapper images. Remote Sensing of Environment 37 (3), 181–191.
Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A., 2012. Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique. Geoderma 183–184, 41–48.

Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., 2013. A methodology for digital soil mapping in poorly accessible areas. Geoderma 192, 341–351.
Consultores, C.O.B.A., 1982. Aproveitamento hidroagrícola de Massingir-Chinhangane. SERLI, Maputo.
De Gruijter, J.J., Brus, D.J., Bierkens, M.F.P., Knotters, M., 2006. Sampling for natural resource monitoring. Springer, Berlin.
EUROCONSULT, 1989. Agricultural Compendium for Rural Development in the Tropics and Subtropics, Third revised edition. Elsevier Scientific, Amsterdam (etc.).
FAO, Unesco, 1997. Mapa dos Solos do Mundo — Legenda revista. FAO, Roma.
Florinsky, I.V., Eilers, R.G., Manning, G.R., Fuller, L.G., 2002. Prediction of soil properties by digital terrain modelling. Environmental Modelling & Software 17 (3), 295–311.
Fox, J., 1997. Applied Regression, Linear Models, and Related Methods. Sage, Newbury Park.
Franklin, R.B., Mills, A.L., 2003. Multi-scale variation in spatial heterogeneity for microbial community structure in an eastern Virginia agricultural field. FEMS Microbiology Ecology 44, 335–346.
Gessler, P.E., Chadwick, O.A., Chamran, F., Althouse, L., Holmes, K., 2000. Modeling soil-landscape and ecosystem properties using terrain attributes. Soil Science Society of America Journal 64 (6), 2046–2056.
Gijsman, A.J., Jagtap, S.S., Jones, J.W., 2002. Wading through a swamp of complete confusion: how to choose a method for estimating soil water retention parameters for crop models. European Journal of Agronomy 18 (1–2), 77–106.
Giller, K.E., Leeuwis, C., Andersson, J.A., Andriesse, W., Brouwer, A., Frost, P., Hebinck, P., Heitkonig, I., van Ittersum, M.K., Koning, N., Ruben, R., Slingerland, M., Udo, H., Veldkamp, T., van de Vijver, C., van Wijk, M.T., Windmeijer, P., 2008. Competing claims on natural resources: what role for science? Ecology and Society 13 (2), 18.
Goovaerts, P., 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma 89 (1–2), 1–45.
Guo, Y., Amundson, R., Gong, P., Yu, Q., 2006. Quantity and spatial variability of soil carbon in the conterminous United States. Soil Science Society of America Journal 70 (2), 590–600.
Hijmans, R.J., Cameron, S., Parra, J., 2011. WorldClim — Global Climate Data. Berkley, CA.
INGC, UEM, Fews-NET MIND, 2003. Atlas for Disaster Preparedness and Response in the Limpopo basin. Instituto Nacional de Gestao de Calamidades (INGC).
INIA, 1995. Legenda da Carta Nacional de Solos — escala 1:1,000,000, Instituto Nacional de Investigação Agronómica, Maputo.
Janssen, P.H.M., Heuberger, P.S.C., 1995. Calibration of process-oriented models. Ecological Modelling 83 (1–2), 55–66.
Janzen, H.H., Ellert, B.H., 2002. Organic matter in the landscape. In: Lal, R. (Ed.), Encyclopedia of Soil Science. M. Dekker Inc., New York, pp. 905–909.
Manninen, T., Eerola, T., Makitie, H., Vuori, S., Luttinen, A., Senvano, A., Manhica, V., 2008. The Karoo volcanic rocks and related intrusions in southern and central Mozambique. Geological Survey of Finland, Special Paper 48. 211–250.
Mapa, R.B., Kumaragamage, D., 1996. Variability of soil properties in a tropical Alfisol used for shifting cultivation. Soil Technology 9 (3), 187–197.
Marchant, B.P., Lark, R.M., 2007. Robust estimation of the variogram by residual maximum likelihood. Geoderma 140 (1–2), 62–72.
Martius, C., Tiessen, H., Vlek, P.L.G., 2001. The management of organic matter in tropical soils: what are the priorities? Nutrient Cycling in Agroecosystems 61 (1), 1–6.
McBratney, A.B., Webster, R., 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables — II: program and examples. Computers & Geosciences 7 (4), 335–365.
McBratney, A.B., Webster, R., Burgess, T.M., 1981. The design of optimal sampling schemes for local estimation and mapping of regionalized variables — I: theory and method. Computers & Geosciences 7 (4), 331–334.
McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97 (3–4), 293–327.
McBratney, A.B., Mendonca Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1–2), 3–52.
Mendonça-Santos, M.L., Dart, R.O., Santos, H.G., Coelho, M.R., Berbara, R.L.L., Lumbreras, J.F., 2010. Digital soil mapping of topsoil carbon content of Rio de Janeiro state, Brazil. In: Boettinger, J.L., et al. (Ed.), Digital Soil Mapping, Progress in Soil Science, vol. 2. Springer Science+Business Media B.V., pp. 255–266. http://dx.doi.org/10.1007/978-90-481-8863-5_21.
Milne, E., Adamat, R.A., Batjes, N.H., Bernoux, M., Bhattacharyya, T., Cerri, C.C., Cerri, C.E.P., Coleman, K., Easter, M., Falloon, P., Feller, C., Gicheru, P., Kamoni, P., Killian, K., Pal, D.K., Paustian, K., Powlson, D.S., Rawajfih, Z., Sessay, M., Williams, S., Wokabi, S., 2007. National and sub-national assessments of soil organic carbon stocks and changes: the GEFSOC modelling system. Agriculture, Ecosystems & Environment 122 (1), 3–12.
Minasny, B., McBratney, A.B., Mendonça-Santos, M.L., Odeh, I.O.A., Guyon, B., 2006. Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. Australian Journal of Soil Research 44, 233–244.
Ministerio do Turismo, 2003. Limpopo National Park: Management and Development Plan. Ministério do turismo, Maputo.
Mishra, U., Lal, R., Liu, D., Van Meirvenne, M., 2010. Predicting the spatial variation of the soil organic carbon pool at a regional scale. Soil Science Society of America Journal 74 (3), 906–914.
Moore, I.D., P.E., Gessler, Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57, 443–452.
Nhantumbo, A., Ledin, S., Du Preez, C., 2009. Organic matter recovery in sandy soils under bush fallow in southern Mozambique. Nutrient Cycling in Agroecosystems 83 (2), 153–161.
Oliver, M.A., 2001. Determining the spatial scale of environmental properties using the variogram. In: Tate, N.J., Atkinson, P.M. (Eds.), Modelling Scale in Geographical Information Science. John Willey & Sons, Ltd., pp. 193–219.
Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences 30 (7), 683–691.
Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Koppen–Geiger climate classification. Hydrology and Earth System Sciences 11, 1633–1644.
R Development Core Team, 2006. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Rutten, R., Makitie, H., Vuori, S., Marques, J.M., 2008. Sedimentary rocks of the mapai formation in the Massingir–Mapai region, Gaza province, Mozambique. Geological Survey of Finland, Special Paper 48. 251–262.
Ryan, C.M., Williams, M., Grace, J., 2011. Above- and Belowground Carbon Stocks in a Miombo Woodland Landscape of Mozambique. Biotropica 43 (4), 423–432.
Scott, N.A., Tate, K.R., Giltrap, D.J., Tattersall Smith, C., Wilde, H.R., Newsome, P.J.F., Davis, M.R., 2002. Monitoring land-use change effects on soil carbon in New Zealand: quantifying baseline soil carbon stocks. Environmental Pollution 116 (Supplement 1), S167–S186.
Shepherd, K.D., Walsh, M.G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal 66 (3), 988–998.
Simbahan, G.C., Dobermann, A., Goovaerts, P., Ping, J., Haddix, M.L., 2006. Fine-resolution mapping of soil organic carbon based on multivariate secondary data. Geoderma 132 (3–4), 471–489.
Smith, P., Smith, J.U., Powlson, D.S., McGill, W.B., Arah, J.R.M., Chertov, O.G., Coleman, K., Franko, U., Frolking, S., Jenkinson, D.S., Jensen, L.S., Kelly, R.H., Klein-Gunnewiek, H., Komarov, A.S., Li, C., Molina, J.A.E., Mueller, T., Parton, W.J., Thornley, J.H.M., Whitmore, A.P., 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. Geoderma 81 (1–2), 153–225.
Stalmans, M., Gertenbach, W.P.D., Carvalho-Serfontein, F., 2004. Plant communities and landscapes of the Parque nacional do Limpopo, Mocambique. Koedoe 47 (2), 61–81.
Tan, Z., Lal, R., Smeck, N.E., Calhoun, F.G., Slater, B.K., Parkinson, B., Gehring, R.M., 2004. Taxonomic and geographic distribution of soil organic carbon pools in Ohio. Soil Science Society of America Journal 68 (6), 1896–1904.
Tan, Z., Liu, S., Tieszen, L.L., Tachie-Obeng, E., 2009. Simulated dynamics of carbon stocks driven by changes in land use, management and climate in a tropical moist ecosystem of Ghana. Agriculture, Ecosystems & Environment 130 (3–4), 171–176.
Thompson, J.A., Kolka, R.K., 2005. Soil carbon storage estimation in a forested watershed using quantitative soil-landscape modeling. Soil Science Society of America Journal 69 (4), 1086–1093.
Thompson, J.A., Bell, J.C., Butler, C.A., 2001. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling. Geoderma 100 (1–2), 67–89.
Tiessen, H., Santos, M., 1989. Variability of C, N and P content of a tropical semiarid soil as affected by soil genesis, erosion and land clearing. Plant and Soil 119 (2), 337–341.
Ungaro, F., Staffilani, F., Tarocco, P., 2010. Assessing and mapping topsoil organic carbon stock at regional scale: a scorpan kriging approach conditional on soil map delineations and land use. Land Degradation & Development 21 (6), 565–581.
Vågen, T.-G., Lal, R., Singh, B.R., 2005. Soil carbon sequestration in sub-Saharan Africa: a review. Land Degradation & Development 16 (1), 53–71.
van Groenigen, J.W., Siderius, W., Stein, A., 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87 (3–4), 239–259.
van Reeuwijk, L.P., 2002. Procedures for soil analysis. ISRIC technical paper 9.
Wang, L., Okin, G.S., Caylor, K.K., Macko, S.A., 2009. Spatial heterogeneity and sources of soil carbon in southern African savannas. Geoderma 149 (3–4), 402–408.
Webster, R., Welham, S.J., Potts, J.M., Oliver, M.A., 2006. Estimating the spatial scales of regionalized variables by nested sampling, hierarchical analysis of variance and residual maximum likelihood. Computers & Geosciences 32 (9), 1320–1333.
Williams, M., Ryan, C.M., Rees, R.M., Sambane, E., Fernando, J., Grace, J., 2008. Carbon sequestration and biodiversity of re-growing miombo woodlands in Mozambique. Forest Ecology and Management 254 (2), 145–155.
Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. Journal of Hydrology 251 (3–4), 123–150.
Ziadat, F.M., 2005. Analyzing digital terrain attributes to predict soil attributes for a relatively large area. Soil Science Society of America Journal 69 (5), 1590–1599.