Catena 166 (2018) 298–309

Contents lists available at ScienceDirect

Catena

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

Soil organic carbon stocks and their determining factors in the Dano catchment (Southwest Burkina Faso)

Ozias K.L. Hounkpatin[a,*], Felix Op de Hipt[b], Aymar Yaovi Bossa[c], Gerhard Welp[a], Wulf Amelung[a]

[a] University of Bonn, Institute of Crop Science and Resource Conservation (INRES), Soil Science and Soil Ecology, Nussallee 13, D-53115 Bonn, Germany
[b] University of Bonn, Department of Geography, Meckenheimer Allee 166, 53115 Bonn, Germany
[c] West African Science Service Center on Climate Change and Adapted Land Use (WASCAL), Ouagadougou 06, P.O. Box 9507, Burkina Faso

ARTICLE INFO

Keywords:
Soil organic carbon
Carbon stock
Random forest
Multiple linear regression land use
Reference soil group

ABSTRACT

Although the evaluation of soil organic carbon (SOC) stocks across different types of land use and major reference soil groups is essential for mitigating climate change, there remains, to date, limited comprehensive understanding of whole tropical soil profiles. Therefore, this study aimed to explain the amount of SOC stocks in different land-use systems and across various soil groups, as well as its spatial pattern in the topsoil (0–30 cm) and subsoil (30–100 cm) within the savannah zone of Burkina Faso. Roughly 70 soil profiles were considered along with additional auger sampling to account for spatial variation in both cropland (CR) and savannah (SA). The machine learning technique random forest regression (RFR) and multiple linear regression (MLR) were used for modeling the surface and subsurface SOC stocks. For model calibration, covariates including land use, topographic, texture, and climatic data were considered as surrogate for soil forming factors. The prediction maps produced by the calibrated models were validated by an independent dataset. The results indicated that about 53% of the SOC stock over 1 m depth was held in the upper 30 cm. Only a marginal difference was recorded between the topsoil SOC stock in SA (41.4 t C ha^{-1}) and CR (39.1 t C ha^{-1}) soils. For the subsoil, a significant difference ($p < 0.05$) was observed between the SOC stock of CR soils recording about 40.2 t C ha^{-1} and SA soils with 26.3 t C ha^{-1}. Among the reference soil groups, the Gleysols located at lower elevation positions revealed the highest SOC stocks over 0–30 cm (44 t C ha^{-1}) and 100 cm depth (86.6 t C ha^{-1}). The Stagnosols (45.2 t C ha^{-1}) followed by the Gleysols (42.7 t C ha^{-1}) recorded the highest SOC stocks over 30–100 cm. The variability of SOC stock in the topsoil was primarily related to site-specific elements, such as particle-size fraction and wetness index, while its distribution in the subsoil was mainly associated with the topographical orientation represented by the slope aspect. Compared to the MLR, RFR estimated mean top- and subsoil SOC stocks of the catchment fairly well, along with lower statistical error metrics, though extreme values were not covered. Nevertheless, the findings on SOC stocks reinforce the view that the semi-arid ecosystems of West Africa still offer a significant opportunity for carbon sequestration for both topsoil and subsoil, and these results represent a baseline for future modeling of SOC dynamics in the region.

1. Introduction

Globally, soils contain the largest terrestrial carbon pool on earth. Though subject to regular change, the global amount of carbon in soils is estimated at 2500 Gt, including 1550 Gt of soil organic carbon (SOC) and 950 Gt of soil inorganic carbon (Batjes and Sombroek, 1997; Lal, 2008). As the SOC pool is 3.3 times the size of the atmospheric pool (760 Gt) and 4.5 times the size of the biotic pool (560 Gt) (Lal, 2004), slight changes in soil C cycling may significantly impact the global C cycle. Nevertheless, little is known on the role of tropical soils in these changes.

The ecosystems in West Africa are facing severe degradation due to changes in land use from perennial vegetation to cropping, increased cultivation in marginal lands, soil erosion, and nutrient mining (Bationo et al., 2007; UNEP, 2006), as well as climate change (Brevik, 2013). Models have predicted that climate change will lead to the conversion of soils from carbon sinks to carbon sources (Cox et al., 2000). However, prediction uncertainty remains significant (Cox et al., 2000; Smith, 2008), mainly due to the lack of adequate knowledge on SOC distribution across the landscape. Nowadays, different measures to conserve existing SOC stocks and trap the atmospheric carbon in the soil are being implemented in many areas in Africa, and comprise

* Corresponding author.
E-mail address: hozias@uni-bonn.de (O.K.L. Hounkpatin).

https://doi.org/10.1016/j.catena.2018.04.013
Received 9 April 2018; Accepted 10 April 2018
0341-8162/ © 2018 Elsevier B.V. All rights reserved.

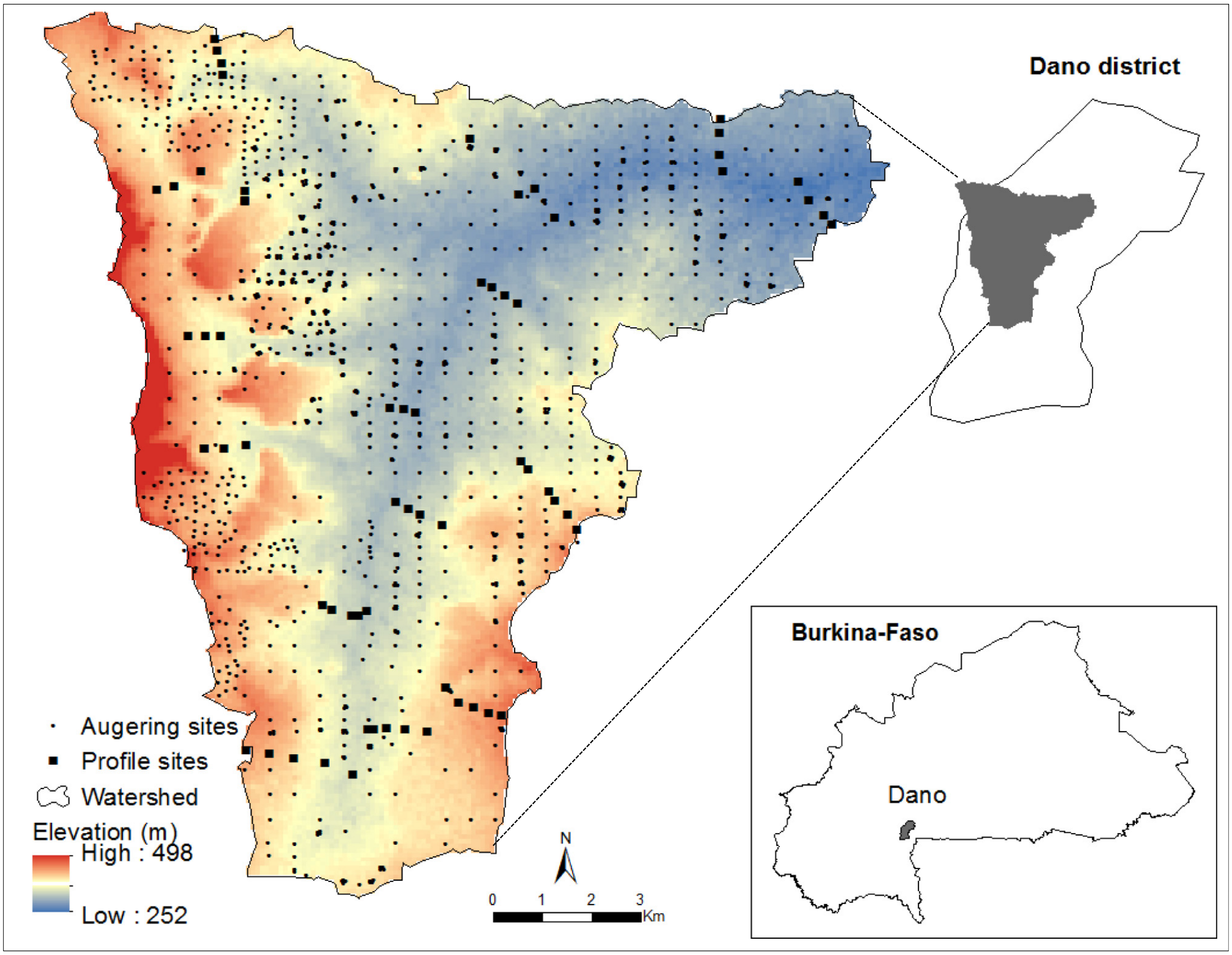

Fig. 1. The Dano catchment and sampling sites.

afforestation of degraded lands, agroforestry, and application of best agricultural practices and policies (Batjes, 2008). Recent works have addressed the quantification of soil carbon stocks in African countries; such studies include those by Akpa et al. (2016) for Nigeria and by Minasny et al. (2017) for South Africa, Tanzania, Kenya, and Nigeria. However, comprehensive data is still lacking on SOC stock for different agrosystems (Anikwe, 2010), especially for other countries in Western Africa. According to Batjes (2008), an estimation of the current carbon stock should be carried out before considering the carbon change dynamics under land use and climate change.

To accurately address land degradation, spatial information on soil properties is required for land evaluation. Spatial soil information as represented in soil maps is beneficial for farmers, scientists, and policy makers in identifying priority areas and for a sound and objective decision making. However, maps from traditional surveys are mostly qualitative, labor intensive, time consuming, and costly (Taghizadeh-Mehrjardi et al., 2015), and thus in most cases also obsolete (Kilasara, 2010). Recent advances in remote sensing and information systems have paved the way for digital soil mapping (DSM), which couples soil point data with statistically correlated auxiliary data (McBratney et al., 2003). This coupling of point and auxiliary data is carried out by using (geo-) statistical classification or regression models. Multiple linear regression (MLR) has been widely used in many studies for the prediction of SOC (Florinsky et al., 2002; Guo et al., 2015; Meersmans et al., 2008). However, soil–landscape relationships are often subject to nonlinear dynamics that might not be captured by MLR (Grimm et al., 2008). Many studies (Hengl et al., 2015; Rad et al., 2014; Wiesmeier et al., 2011) have reported on the potential of random forest regression (RFR), an ensemble machine learning approach, to overcome this limitation.

Potential factors that affect SOC stocks and are used as covariates for DSM include climatic and topographic elements (for example, mean annual precipitation and temperature, slope, etc.), land use, physical soil characteristics (texture, parent material, etc.), and microbial biomass (Albaladejo et al., 2013; Jobbágy and Jackson, 2000,; Ladd et al., 2013). Many of these factors have been investigated in various studies conducted across the globe (Albaladejo et al., 2013; Azlan et al., 2011; Bationo et al., 2007; Burke et al., 1989; Chaplot et al., 2010; Jobbágy and Jackson, 2000; Percival et al., 2000). However, these studies have mostly focused on surface soil horizons. Nevertheless, > 50% of SOC is usually allocated below 20 cm depth (Batjes, 1996). Fontaine et al. (2007) showed that this subsoil carbon is readily decomposable upon addition of a fresh C source, and Fierer et al. (2003) concluded that it is even more sensitive to changes in temperature or nutrient availability than topsoil carbon is. However, these latter studies were not performed with tropical soils, which may have specific SOC storage conditions—for instance, due to their special oxide assembly (Feller and Beare, 1997; Kögel-Knabner and Amelung, 2014).

The study detailed in this paper was performed in the Sudanian area of Burkina Faso, which is dominated by Plinthosols—that is, soils with high Fe oxide accrual—particularly in the subsoil. To the best of our knowledge, for such soils, neither (i) levels and distribution of SOC stocks nor (ii) the interactions between SOC stock and landscape properties have ever been investigated. Yet such quantitative data is crucial for estimating the local and regional carbon sequestration potential and the participation of developing countries in the Clean Development Mechanism (CDM) mentioned in the Kyoto Protocol, as well as the "4 per thousand" initiative launched during the twenty-first Conference of Parties (COP21) (Rhodes, 2016). Therefore, this study aims to estimate the surface and subsoil organic carbon stocks in different land-use systems and across various soil orders, as well as assessing the spatial variability of these carbon stocks and underlying factors using RFR and MLR models.

2. Study area

The study was carried out in the Dano catchment (Lat. 11°09′00″ N; Long. 3°04′00″ W), located in the Ioba province, southwest of Burkina Faso (Fig. 1). The Dano catchment covers an area of 155 km^2 and presents a relatively flat landscape with a mean average slope of 0.2%. The elevation ranges from 252 to 498 m above sea level, with a mean elevation of 295 m. The climate consists of one rainy season from May to October, with a mean annual rainfall ranging between 800 and 1200 mm $year^{-1}$ for the period 1951–2005, along with a mean annual temperature between 20.1 and 38.4 °C (Schmengler, 2010). The lithology is composed of partly volcanic formations from the middle precambrian period and is made up mainly of andesic rocks with massive texture, basalt, diabase, gabbro, and quartz-rich andesites.

The vegetation of the area belongs to the Sudanian domain with woody, arboraceous, or scrubby savannah, which is abundant in perennial grasses. However, a major part of the vegetation has been degraded and subject to cultivation and fallowing. The present study does not discriminate between fallow and degraded savannah and both are thus grouped under the term savannah. The farming is mostly rain fed, takes place on a small scale (Callo-Concha et al., 2012), and is subsistence-oriented with low investments regarding fertilizer and machinery. The soil orders in the catchments range from Gleysols, Cambisols, Lixisols, Stagnosols, Leptosols, and the Plinthosols, which are the predominant soil type (73.1% of the area) (Yira et al., 2016).

3. Materials and methods

3.1. Soil sampling

Representative soil units were chosen for sampling, based on existing maps—mainly soil (Bureau National des sols, 2000) and land use (Cord et al., 2010) maps—and a digital elevation model (DEM). The DEM was provided by the Shuttle Radar Topography Mission (SRTM) at a 90 m resolution. About 70 soil profiles were described along 16 transects from August to October 2012. These profiles were excavated up to 1 m where possible, and four soil cores (100 cm^3) were collected per horizon to determine bulk density (BD). These core samples were dried at 105 °C for 24 h before assessment of the total weight and of the stone content (SC). Additionally, composite soil samples were collected from different horizons. These latter samples were dried at 40 °C and sieved to ≤ 2 mm.

To account for the spatial variability of the SOC stock, intensive auger grid sampling was carried out in 2012 (August–October) and in 2013 (August–October 2013) over the entire study area. The sampling resulted in a total of 1257 samples for both soil and carbon mapping purposes. Each sampling location was geo referenced using a global positioning system device. At each auger point, composite samples, as well as core samples (four replicates), for bulk density were taken, but only from topsoil due to limitations in workload. Soil horizon description and soil classification were based on the World Reference Base for soil resources (IUSS et al., 2006). As Plinthosols were the dominant soil group across most of the study area (Yira et al., 2016)—that is, the other soils (Gleysols, Cambisols, Lixisols, Stagnosols, Leptosols) were hard to be found at a priori site selection. This resulted in an imbalanced dataset regarding the distribution of soil pedons among the reference soil groups.

3.2. Soil analysis and mid-infrared prediction

Many studies have proven the ability of mid-infrared spectroscopy (MIRS) to predict organic and inorganic soil properties (Baldock et al., 2013; Bornemann et al., 2010; Grinand et al., 2012; Reeves, 2010; Zimmermann et al., 2007). MIRS enables the identification of specific soil minerals and of organic matter functional groups, such as alkyl or carboxyl groups, carbohydrates, amides, amines, and aromatic functional groups (Janik et al., 2007). The estimation of soil properties is generated by calibrating spectral information against conventionally obtained data using multivariate statistical procedures such as partial least squares regression (Denef et al., 2009; Janik et al., 2007). As no fractionation or chemical reagent is required, MIRS offers a simple, rapid, and low-cost alternative, especially when dealing with large numbers of soil analyses.

For spectra measurement, about 20 mg of the profile samples were set into microplates and compacted with a plunger to get a level and plain surface in five replicates. The spectra were generated using a Bruker Tensor 27 having an automated high throughput device (Bruker HTS-XT). This extension is equipped with a liquid N_2-cooled mercury-cadmium telluride detector. The spectra recording was carried out using the OPUS/LAB software within the range of 8000–600 cm^{-1} (1250–16,700 nm) with a resolution of 4 cm^{-1} for each run. This software provides the most representative spectra upon applying principal component analysis, and about 50% of the corresponding profile samples were chosen for conventional laboratory analysis. These samples were analyzed in the laboratory as a ground truth database for SOC, nitrogen (N), pH, texture, and cation exchange capacity (CEC). These parameters were determined following the procedure described by Reeuwijk (2006).

A cross-validation approach was used to set up the models used for prediction, and a Test t validation was carried out to check the model's robustness. Different optimization schemes were run with the OPUS QUANT software, which provided different data processing procedures and combinations of spectral ranges. In the present study, data preprocessing consisted of the multiplicative scatter correction method (pH, CEC, silt fraction) and a combination of first-derivative and multiplicative scatter correction method (SOC, N, sand, and clay fraction). The quality of the different models was assessed based on their predictive ability with the ratio of performance to deviation (RPD) and the standard error of prediction (SEP) as presented in Table 1 (for more technical information, see the review carried out by Bellon-Maurel and McBratney (2011)). Here, only models exhibiting good predictive ability (RPD > 2), or close to that (RPD 1.7–2.0), were used to make predictions for the remaining samples (see Table 1 for parameter selection).

3.3. Determination of SOC stocks

Though the sampling was initially carried out on a horizon basis, we pooled all data as weighted averages of soil properties in order to provide a quantitative element for the 0–30 cm and 30–100 cm depth intervals. The carbon stock was therefore computed for these two depths. For each depth interval, the SOC stock (SOC stock, t C ha^{-1}) was determined by the product of the C content, the soil depth, and the bulk density of the fine earth along the soil profile. Bulk density was computed by dividing the weight of the oven-dry soil by the volume of the soil cores (Hartge and Horn, 1989). Each quantified bulk density was corrected for the coarse particle content, which was mainly made up of plinthites. $CaCO_3$ was absent in the collected soil samples; therefore, the SOC stocks for each soil profile were obtained based on the following equation:

Table 1
Soil parameter quantification by mid-infrared spectroscopy.

Parameters	Range	n	Cross validation						Test_t validation (V = 10%)				
			Rank	R^2 (%)	SEP	RMSECV	RPD	Slope	R^2 (%)	SEP	RMSEP	RPD	Slope
CEC (cmolc kg^{-1})	02.91–36.3	480	7	75.61	1.89	3.58	2.03	0.80	90.50	1.78	3.16	3.59	0.81
Sand (%)	02.20–51.7	480	7	70.45	2.60	6.76	1.84	0.76	80.93	2.39	5.73	2.52	0.77
Silt (%)	20.7–70.0	480	7	75.75	2.22	4.94	2.03	0.81	88.21	1.96	3.86	2.98	0.85
Clay (%)	10.7–62.9	480	8	77.58	2.48	6.16	2.11	0.82	80.55	2.34	5.49	2.41	0.82
SOC (g kg^{-1})	0.0–31.8	480	7	95.33	0.33	0.11	4.63	0.96	92.21	0.39	0.15	3.62	0.95
N (g kg^{-1})	0.0–2.14	480	7	85.45	0.12	0.02	2.62	0.88	85.73	0.14	0.02	2.95	0.82

n: number of spectra, SEP: standard error of prediction, RMSECV: root mean square error of cross validation, RMSEP: root mean square error of prediction, RPD: ratio of performance to deviation, V: validation set.

$$SOC_{stock} = SOC \times BD \times T \times \left(1 - \frac{CP}{100}\right) \quad (1)$$

where SOC is the organic carbon content (%) of the fine earth (< 2 mm), BD is the bulk density (g cm^{-3}), T is the soil thickness (cm), and CP is the coarse-particle content.

3.4. Selected variables to explain SOC stock variability

The variables (Table 2) considered as covariates consist of: terrain attributes, land use, temperature and precipitation, geomorphology, and lithology. These variables were maps covering the entire study area. All the sample points were overlaid on the different maps considered, and related values at each location were extracted using the ArcGIS 10.3.1 software (ESRI).

The terrain attributes were derived from the 90 m resolution DEM provided by the SRTM. Since Aspect is a circular variable, sin (Aspect) and cos (Aspect) were computed to account for the east–west and north–south topographical orientation, respectively (Lin et al., 2013; Rodríguez-Moreno and Bullock, 2014). The terrain parameters were clustered into local, regional, and combined terrain attributes, as defined by Grimm et al. (2008), without changing the spatial extent for each cell—that is, the 90 m resolution for all these covariates was maintained. The parent material (Geo) allocated to each sampling location was derived using a geological map (1/1000000) of Burkina Faso made by Hottin and Ouedraogo (1992). Since the former land-use map (Cord et al., 2010) used for collecting data was too coarse and obsolete, we consider the land use data from the map (5 m resolution) established by Forkuor (2014). Climatic data includes the mean annual temperature (Temp) and annual precipitation (Prep) at 1 km resolution from the WorldClim datasets. The climatic data were submitted to bicubic resampling prior to data extraction. Though most studies have mainly considered terrain attributes and land use data in predicting SOC, we additionally considered soil properties, as mentioned in Kumar and Lal (2011) and Were et al. (2015). We believed that using different categories of covariates, including soil properties, for digital soil mapping would improve predictive accuracy, as soil forming factors are better covered using more covariates. Therefore, soil texture fractions (sand, silt, clay) were considered in addition to the environmental variables. They were derived from interpolated maps using the ordinary kriging method with the ArcGIS 10.3.1 software (ESRI). Ordinary kriging has been used in many studies to predict soil properties at unsampled locations (Chaplot et al., 2010; Mishra et al., 2009; Were et al., 2015; Zhang and McGrath, 2004).

The predictors were reduced for the subsoil carbon stock model due to the smaller size of the dataset (n = 70). Feature selection was carried out using the RF recursive feature elimination algorithm of the R "caret" Package (Kuhn, 2015). The following variables were finally retained for the subsoil carbon stock prediction: elevation, distance to stream, sin (Aspect), cos (Aspect), ruggedness, curvature, catchment area, sand, silt, clay, precipitation, and temperature. Finally, the digital maps of the topsoil and subsoil carbon stock were predicted by the models at a resolution of 90 m for the whole study area.

Table 2
Selected variables to explain variability of SOC stocks.

Group	Parameters	Definition	Abbreviation	Units	Authors
Local	Slope	Inclination of the land surface from the horizontal	Slope.per	%	Allen et al. (2013)
	Slope length	Distance from origin of overland flow to deposition point	Slope.length	m	Wiesmeier et al. (2011)
	Curvature	Combination of horizontal and vertical curvature	A.curv	$^{\circ}$ m^{-1}	Grimm et al. (2008)
	Maximum curvature	Maximum curvature	S.max.cuv	$^{\circ}$ m^{-1}	Grimm et al. (2008)
	Minimum curvature	Minimum curvature	S.min.cuv	$^{\circ}$ m^{-1}	Grimm et al. (2008)
	Plan curvature	Horizontal (contour) curvature	S.Plan.cur	$^{\circ}$ m^{-1}	Wiesmeier et al. (2011)
	Profile curvature	Vertical rate of change of slope	S.Profile.cur	$^{\circ}$ m^{-1}	Davy and Koen (2013)
	Elevation	Vertical distance above sea level	Elevation	m	Davy and Koen (2013)
	cos (Aspect)	North-south topographical orientation	cosAsp		Wang et al. (2013)
	sin (Aspect)	East-west topographical orientation	sinAsp		Wang et al. (2013)
Regional	Catchment area	Discharge contributing upslope area	S.CA	m^2	Grimm et al. (2008)
	Distance to stream	Distance to stream network	Dist.stream	m	Grimm et al. (2008)
Combined	Topographic Wetness Index	Ratio of local catchment area to slope	A.TWI	–	Wiesmeier et al. (2011)
	Saga Wetness Index	Ratio of local catchment area to slope	S.Wet.Ind	–	Selige and Böhner (2006)
Climatic	Temperature	Temperature	Temp	°C	Page et al. (2013)
	Precipitation	Precipitation	Prep	mm	Page et al. (2013)
Soil properties	Sand	Sand	Sand	%	Were et al. (2015)
	Silt	Silt	Silt	%	Were et al. (2015)
	Clay	Clay	Clay	%	Were et al. (2015)
Others	Lithology	Lithology	Litho	–	Wiesmeier et al. (2011)
	Geormorphology	Geormorphology	Geo	–	
	Land use	Land use	LU	–	Wiesmeier et al. (2011)
	Reference soil group	Reference soil group	rsg	–	Allen et al. (2013)

Table 3
Basic soil characteristics under different land uses (mean values with SD).

		N	Sand	Silt	Clay	BD	pH
0–30 cm							
LU	CR	36	28.1^{a}(± 9.1)	43.2^{a}(± 7.1)	28.5^{a}(± 10.1)	1.4^{a}(± 0.1)	6.4^{a}(± 0.5)
	SA	34	29.9^{a}(± 12.3)	44.8^{a}(± 10.5)	25.9^{a}(± 9.5)	1.5^{a}(± 0.1)	6^{a}(± 0.4)
RSG	CM	8	25.5^{ac}(± 11.3)	42.2 ^{ac}(± 6.7)	32.2^{a}(± 13.6)	1.3^{ab}(± 0.1)	7^{a}(± 0.4)
	GL	12	19.1^{ba}(± 11.3)	50.3^{ba}(± 9.7)	31.7^{a}(± 11)	1.4^{ac}(± 0.1)	6.1^{bc}(± 0.3)
	LX	2	22.6^{ac}(± 3.5)	55.3^{ac}(± 8.6)	20.5^{a}(± 2.9)	1.4^{ac}(± 0.001)	6.2^{abc}(± 0.4)
	PT	44	32.8^{c}(± 8.9)	42^{c}(± 8.2)	25.2^{a}(± 8.4)	1.5^{c}(± 0.1)	6.1^{c}(± 0.4)
	ST	4	29^{ac}(± 10.7)	43.9^{ac}(± 8.9)	27.3^{a}(± 9.8)	1.4^{ac}(± 0.1)	6.5^{abc}(± 0.4)
30–100 cm							
LU	CR	36	21.6^{a}(± 6.9)	40.7^{a}(± 4.8)	37.2^{a}(± 7.9)	2^{a}(± 0.7)	6.3^{a}(± 0.5)
	SA	34	22.8^{a}(± 5.3)	41.8^{a}(± 6.2)	34.9^{a}(± 4.5)	2.1^{a}(± 0.7)	6.1^{a}(± 0.4)
RSG	CM	8	26.4^{a}(± 9.1)	39.5^{a}(± 2.7)	33.7^{a}(± 9.9)	1.7^{a}(± 0.6)	6.9^{a}(± 0.7)
	GL	12	19.7^{a}(± 7.5)	45.3^{a}(± 7.9)	34.5^{a}(± 5.8)	1.6^{a}(± 0.1)	6.1^{bc}(± 0.3)
	LX	2	17.9^{a}(± 6.1)	46^{a}(± 6.7)	34.4^{a}(± 2.3)	1.5^{a}(± 0.1)	6.1^{abc}(± 0.2)
	PT	44	22.2^{a}(± 4.6)	40.2^{a}(± 4.3)	37.1^{a}(± 6.3)	2.3^{a}(± 0.7)	6.1^{c}(± 0.3)
	ST	4	22.9^{a}(± 8.5)	41.3^{a}(± 8.3)	35.1^{a}(± 4.1)	1.8^{a}(± 0.8)	6.7^{abc}(± 0.7)

LU: land use, CR: cropland, SA: natural/semi natural vegetation, RSG: Reference soil groups, CM: Cambisols, GL: Gleysols, LX: Lixisols, PT: Plinthosols, ST: Stagnosols, n: number of samples, BD: bulk density. Means followed by the same letters are not significantly different ($p < 0.05$), topsoil and subsoil are to be considered separately.

3.5. Statistical analysis

Descriptive statistics (means and standard deviations [SDs] of the mean) were used to characterize the measured values of the variables. Normality of the carbon data was checked with the Shapiro–Wilk test. Student's *t*-test was used for comparison between the SOC stocks of the different land-use systems. Bartlett's test for homogeneity of variance was performed due to the unequal size of the data for the soil reference groups (Yu, 2011). The significance of the difference in the mean SOC stocks between the reference soil groups was examined using the Welch analysis of variance test, while for multiple means comparisons the Games–Howell test was performed, as carried out in Cornelissen et al. (2001). All statistical analyses were carried out using the R version 3.0.3 software (R Core Team, 2015).

3.6. Prediction models: multiple linear regression and random forest regression

In the present study, MLR and RFR were used as statistical models to predict the spatial distribution of the topsoil SOC stock. MLR is a commonly used statistical approach to predict the values of a dependent variable (here, the SOC stocks) based on a set of independent variables (here, the covariates in Table 2). In this study, MLR was implemented using the R "caret" package (Kuhn, 2015) using tenfold cross-validation with five repetitions.

RFR (Breiman, 2001) was used to get insight into the explanatory variables which affect the variability of the SOC stock. RFR functions are performed in a similar way as a regression tree, with the difference that many classification trees are generated and averaged to give a prediction for the response variable. Accuracy is reached with low bias and low variance as a large number of trees is averaged (Grimm et al., 2008). The reasons for using RFR in this study are related to its robustness towards nonlinearity, overfitting, and interaction, as well as to its capacity to deal with continuous and categorical variables (Breiman, 2001). Each tree is built from a unique bootstrap sample of the dataset, thereby providing a reliable error estimation based on the remaining set, called the out-of-bag data. The measure of the variable importance based on the mean decrease of the prediction accuracy before and after random shuffling of the predictors gives insight into the influence that each of the explanatory variables has on the response variable. The latter was considered for identifying the most important factors affecting the SOC stocks for both topsoil and subsoil. The RFR analysis was also carried out using the R "caret" package (Kuhn, 2015) using tenfold cross-validation with five repetitions.

3.7. Model training and mapping

The topsoil dataset was split, with 70% of the samples to train the model while 30% were used as an independent validation set. However, for the subsoil we conducted a split of 80% for model training and 20% for validation. The models derived from the MLR and RFR were used to make the respective prediction maps, which were corroborated by different validation sets. For the stability and robustness of the models, the different calibrations were carried out based on a five-time repeated 10-fold cross-validation using the "caret" R Package (Kuhn, 2015). The cross-validation was performed on the 70% and 80% of the data used for training the model for the topsoil and the subsoil, respectively. The root mean square error (RMSE) of cross-validation (RMSECV) and the RMSE from prediction based the validation set (RSMEPV), as well as the mean error (ME) were used to assess the model accuracies.

$$RMSE = \left(\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2\right)^{1/2} \quad (2)$$

$$ME = \frac{1}{n}\sum_{i=1}^{n}(P_i - O_i) \quad (3)$$

where "P" is the predicted value and "O" is the observed/measured value.

4. Results and discussion

4.1. Basic soil characteristics

The different land use and reference soil groups were dominated by silty soil texture with elevated clay contents in the subsoil (Table 3). Textural variations occurred among the different soil groups: the Gleysols were silty and less sandy than the Plinthosols which peaked in the opposite direction. Possibly, the latter was caused by pseudo-sand like oxide concretions which could not be destroyed completely during conventional texture analyses. The bulk density increased with depth, with larger values recorded in the subsoil for both land-use systems. Maximum bulk densities were found for the Plinthosol subsoils, which indicated the presence of petroplinthite in some of these profiles. The pH was slightly acidic and comparably similar among land use and reference soil groups at all soil depths. This trend is in line with values reported by Yoni et al. (2005) in Western Burkina Faso.

Table 4
SOC stocks in the different land use systems and reference soil groups.

LU	RSG	n	0–30 cm		30–100 cm		0–100 cm	
			Mean	Sd	Mean	Sd	Mean	Sd
CR & SA		70	39	± 16.7	33.9	± 23.8	73.5	± 30.7
CR (t C ha^{-1})		37	39.1^a	± 16.5	40.2^a	± 27.9	77.1^a	± 34.9
SA (t C ha^{-1})		33	41.4^a	± 17.4	26.3^b	± 15.9	67.7^a	± 27.3
CR (t C ha^{-1})	CM	6	40.2^a	± 12.6	48.7^a	30.7	88.9^a	± 40.5
SA (t C ha^{-1})		2	16.6^b	± 8.3	20.6^a	16.0	37.2^b	± 7.6
CR (t C ha^{-1})	GL	5	39.9^a	± 12.2	52.7^a	± 32	94.4^a	± 35.4
SA (t C ha^{-1})		7	46.6^a	± 18.9	35.6^a	± 15.1	82.5^a	± 31.2
CR (t C ha^{-1})	LX	1	27.6	.	26.0	.	53.6	.
SA (t C ha^{-1})		1	37.6	.	21.9	.	59.5	.
CR (t C ha^{-1})	PT	22	39.8^a	± 15	33.7^a	± 24.5	73.2^a	± 32.4
SA (t C ha^{-1})		22	42.4^a	± 16.9	24.6^a	± 16.3	67.0^b	± 25.9
CR (t C ha^{-1})	ST	3	9.0^a	± 5	54.6^a	± 42.7	63.6	± 46.8
SA (t C ha^{-1})		1	36.7^b	.	17.2^a	.	54.0	.
CR (t C ha^{-1})	CM	6	40.2^a	± 12.6	48.7^a	± 30.7	88.9^a	± 40.5
	GL	5	40.0^a	± 12.2	52.7^a	± 32	92.7^a	± 38.3
	PT	22	39.8^a	± 15	33.7^a	± 24.5	73.2^a	± 32.4
	ST	3	9.0^b	± 5	54.6^a	± 42.7	63.6^a	± 46.8
SA (t C ha^{-1})	CM	2	16.6^a	± 8.3	20.6^a	± 16.5	36.6^a	± 8.1
	GL	7	46.6^a	± 18.9	35.6^a	± 15.1	82.2^a	± 31.4
	PT	22	42.3^a	± 16.9	24.6^a	± 16.3	67.0^a	± 25.9

LU: land use, RSG: reference soil group, CR: cropland, SA: savannah, CM: Cambisols, GL: Gleysols, LX: Lixisols, PT: Plinthosols, ST: Stagnosols; n: number of samples, Means followed by the same letters are not significantly different ($p < 0.05$).

4.2. SOC stock in relation to land use and reference soil group

The distribution of the SOC stocks in the different land-use systems, as well as in the reference soil group of each specific land use, is presented in Table 4. About 73.5 t C ha^{-1} was recorded as the total average of SOC stock in the entire profile (0—100 cm) of the Dano catchment with 39 t C ha^{-1} found for the topsoil (0–30 cm) and 33.9 t C ha^{-1} for the subsoil (30–100 cm), amounting, respectively, to 53% and 47% of the total stock. These results coincide with the findings reported by other authors, with Batjes (1996) recording 39–70% of the SOC stock in the first 30 cm and Doetterl et al. (2015) reporting about 52% of SOC stock at the same depth. The total average of SOC stock over 100 cm recorded in the present study is higher than the range estimations of 42–45 t C ha^{-1} for West Africa and 64–67 t C ha^{-1} reported for Africa (Batjes, 2001); on the other hand, our average value is lower compared to the 82 t C ha^{-1} found by Hien et al. (2003) for southern Burkina Faso.

In the topsoil, the SOC stock was similar for both land-use systems. The average SOC stocks of the noncropped sites only slightly exceeded that of the croplands (2.3 t C ha^{-1}; not significant). The lack of significance was due to the Cambisols, which showed significantly larger SOC stocks in the surface soils of the croplands, likely due to former land degradation or just site preference of farmers for the better Cambisols. The larger SOC stocks in the surface soils for the other sites under natural vegetation are in line with other studies (Bruun et al., 2013; Singh et al., 2011).

A study in Ghana by Boakye-Danquah et al. (2014) reported 22.9 t C ha^{-1} for the topsoil of cultivated areas but 49.4 t C ha^{-1} for those under native vegetation. For Burkina Faso, Hien et al. (2006) recorded 16–25 t C ha^{-1} for cropland soil and 61 t C ha^{-1} for savannah soils. However, our data for the Dano catchment pointed to smaller differences among the two land-use systems (Table 4).

The small difference of SOC stocks between these two land-use systems in the Dano catchment suggest a high level of degradation of the sites under savannah, which is subject to overgrazing due to the absence of sufficient grazing areas and the nonexistence of straw and silage production (Callo-Concha et al., 2012). The pressure on these noncropped fields is worsened by the presence of migratory herding, which adds to the local livestock (Gonin and Tallet, 2012). Moreover, the production of a local beer (*dolo*) results in the use of about 6400 t of firewood per year from the native savannah sites; this also constitutes a major source for the degradation of natural resources (Blin and Sidibe, 2012). The sites under savannah may also include old fallow soils, which, because of current herding pressure, have failed to rebuild their carbon stock. Once degraded, it may take decades until SOC stocks in such savannah soils are restored (Preger et al., 2010).

One additional peculiarity was the presence of stone lines (field survey observation, Supplementary Information SI.1) in the croplands, which may have also reduced soil erosion, as observed by Schmengler (2010) in the same area. Zougmoré et al. (2004) reported a reduction of runoff by 45% with the use of stone lines as conservation practice. Therefore, the presence of these stone lines might have contributed to slowing down the SOC loss from the cropland.

Intriguingly, significantly different C stocks were found for the subsoils that contained more SOC in the cropland than in the savannah sites (Table 4). This SOC storage overcompensated SOC gains in the surface soils, so that significance disappeared on a whole soil profile basis. In part, the larger SOC stocks under cropland may be attributed to the presence of petroplinthite in the subsoil of the savannah soils, which are thus not used for cropping nowadays. In addition, intensive translocation processes in the croplands may have been induced at elevated precipitation events under a tropical climate, as formerly reported for the leaching of basic cations into the subsoil (Eze et al., 2014) along with clay and SOC, especially for low-acidity soils, which also prevailed in our study (Lorenz and Lal, 2005).

Considering the different reference soil groups in the topsoil, the Plinthosols (41.1 t C ha^{-1}) contained more or less as much SOC as the Gleysols (43.8 t C ha^{-1}). The latter also recorded the largest carbon stock over 100 cm depth (86.6 t C ha^{-1}) followed by the Cambisols (75.8 t C ha^{-1}) and the Plinthosols (70.1 t C ha^{-1}) (see Supplementary information, SI. 2). The prevalence of SOC in Gleysols might not be solely due to limited SOC decomposition under groundwater influence, but could mainly be related to the occurrence of local erosion processes, leading to the transport of SOC-rich sediments from upslope to the lower slope, and thus from other soils into the Gleysols under the combined effect of slope, elevation, and heavy tropical rain. Doetterl et al. (2013) reported a significant difference in SOC stocks between erosional and depositional areas due to soil relocation processes and local topographical features. However, with similar SOC stocks in the topsoil between Gleysols, Plinthosols, and Cambisols, depositional areas might not correspond only to Gleysols due to the variability of topographic features across the landscape. On the other hand, the periodic saturation by groundwater reduces oxidation processes in the subsoil.

The Stagnosols of the cropland exhibited the lowest SOC stocks (9 t C ha^{-1}, Table 4). As temporary saturation with water in the stagnosols should normally promote SOC storage, rather than distorting it, we attribute this finding first to their position at a relatively medium to high position in the landscape, favoring vulnerability to soil erosion, and second to stagnic conditions occurring at a relatively deeper depth regarding the high carbon stock in the subsoil (54.6 t C ha^{-1}). Moreover, exposition to a longer cultivation duration with very low input (Bationo and Buerkert, 2001) could also be responsible for the low carbon level of the topsoil, but investigation into the land use history is necessary before any sound conclusion can be reached. The Stagnosols, exhibiting larger SOC stocks in the subsoil of the croplands, could be taken as additional evidence that for mapping soil C storage, the consideration of whole soil profiles is needed.

4.3. Factors affecting the spatial variability of the topsoil SOC stocks

The analysis of variable importance characterizes the influences that different explanatory variables (see Table 2) have on the response variable (here, SOC stock). The analysis revealed different preeminent parameters controlling SOC stocks of topsoil and subsoil (Fig. 2) for

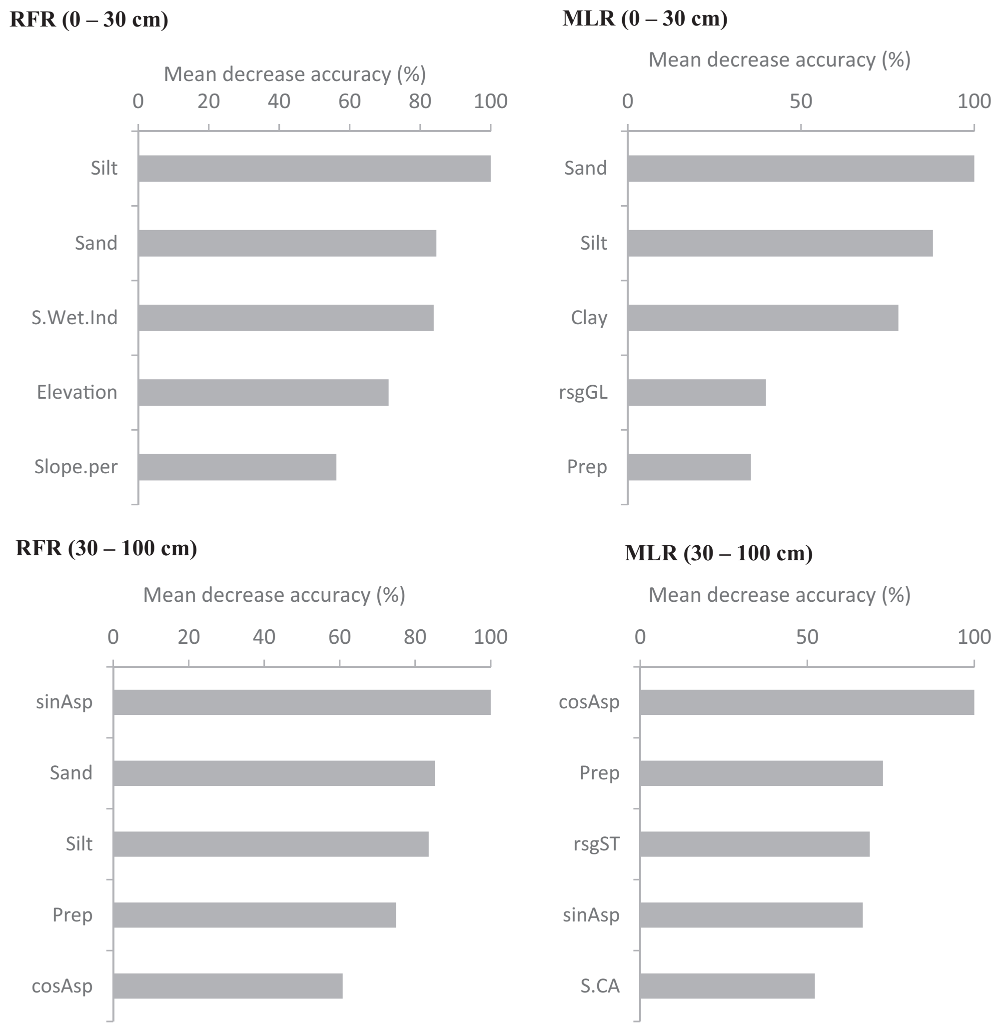

Fig. 2. Top five variables from the RFR and MLR models for the topsoil (0–30 cm) and subsoil (30–100 cm). S.Wet.Ind: wetness index, Slope.per: slope (%), rsgGL: Gleysols, Prep: Precipitation, S·CA: catchment area, rsgST: Stagnosols, RFR: random forest regression, MLR: multiple linear regression, cosAsp: cosine(Aspect), sinAsp: sine(Aspect).

RFR and MLR. Only the top five variables are considered in the figure (see Supplementary information, SI. 3–4 for scatter plots as a function of SOC stock).

The soil texture was revealed as the most prominent predictors for the topsoil SOC stock, with the silt and sand ranking first for the RFR while coming in inverse order with the MLR. The latter model also recorded clay next to the first two variables. Soil texture in general, and especially its fine particles (silt and clay), have been extensively discussed in the literature as important agents accounting for the variance of SOC through adsorption of organic matter (Bationo et al., 2007; Chaplot et al., 2010; Mao et al., 2015; Saiz et al., 2012; Zhang and Shao, 2014). The silt recorded a negative correlation with SOC stock (Supplementary information, SI. 3) contrary to some studies which reported a positive association (Sakin, 2012; Augustin and Cihacek, 2016). The presence of pseudosilt as reported for many rich oxide tropical soils might support this trend compared to previous studies (Embrechts and Sys, 1988; Jungerius et al., 1999; Gidigasu, 2012). However, it is obvious that not only linear patterns are at play within such soil-landscape system. As recorded in Table 2, the dominant content of silt in the topsoil makes it the most abundant soil particle involved in SOC storage by, for example, adsorption or aggregation. In addition, silt is known to affect the storage of plant-available water, which adds to the SOC storage via its control on primary plant production and carbon input. The correlation of SOC stock to wetness index (indicator of soil moisture) with the RFR may be taken as additional evidence for such interactions, as previously also reported by Zadorova et al. (2014). Beside its impact on plant growth, the wetness index also affects SOC dynamics as a hydrological factor at depositional and flat areas, where elevated humidity results in slower decomposition rates for soil organic matter (Doetterl et al., 2013). In logical consequence, elevation and slope are thus recorded as one of the next prominent variables that influence SOC storage (Fig. 2). The finding is in line with results of Hengl et al. (2015), who also reported elevation as a major factor affecting SOC stocks in Africa.

The subsequent parameters influencing surface SOC storage refer to Gleysols and precipitation (Fig. 2) for the MLR. Obviously, their higher level of humidity and position at lower elevation and slope areas explain Gleysols being positioned as the fourth-best predictor. Climate variables have been widely acknowledged as an influential variable for SOC stocks (Doetterl et al., 2013; Manning et al., 2015; Oueslati et al., 2013). Temperature and precipitation distribution affect the production of plant materials and soil fauna activity. Warmer temperatures and wetter conditions would most likely result in higher biomass production and microbial activity. Conversely, a lower heat transfer, coupled with lower humidity, could result in reduced C decomposition. The dry season of the study area is characterized by higher temperatures with very scarce rainfall, which suppresses biomass production, whereas the rainy season comes along with intense and heavy rainfall and subsequent vegetation growth. As a result, the final SOC stock at a given location is likely interactively affected by climatic elements, such as precipitation, which affect soil moisture (wetness index) distribution as a function of elevation, slope, and soil texture.

For the subsoil, sin (Aspect) (east–west topographical orientation) and cos (Aspect) (north–south topographical orientation) were revealed as the first-best predictor for the RFR and MLR, respectively. The slope aspect is reported to affect some microclimatic parameters, such as soil temperature, moisture, vegetation, and soil fauna, by regulating the incoming solar radiation and evapotranspiration (Dorji et al., 2014). Consequently, differences in aspect-induced microclimate resulting in higher biomass and moisture, as well as lower soil organism activities in shady areas (north-facing slopes) compared with sunny areas (south-facing slopes), would result in higher SOC stocks in the former compared to the latter. For the subsoil, about 77 t C ha^{-1} of SOC stock was recorded for the north (and northwestern) facing slopes, while 31 t C ha^{-1} and 19 t C ha^{-1} was recorded for the south- and west-facing slope areas, respectively (results not shown).

Other variables of lower rank but still of major importance, such as soil particle size (contents of sand and silt) and precipitation, were identified by the RFR. Apart from precipitation, the MLR also identified the Stagnosols and catchment area among the top-five variables. Earlier studies from other regions also identified precipitation, among others, as a major control variable for subsoil SOC stocks (Adhikari et al., 2014; Doetterl et al., 2013).

As indicated above, the amount of precipitation drives soil moisture and thus biomass production, contributes to waterlogging and thus SOC storage conditions in the subsoil, and affects the vertical and lateral transport of SOC-rich particles in the landscape. Consequently, Stagnosols, which are soils affected by temporary waterlogging, presented the highest SOC stock in the subsoil (Table 4).

4.4. Spatial distribution of SOC stocks

The spatial distribution pattern of SOC stocks in the topsoil (Fig. 3) based on the prediction of the RFR and MLR model presents an irregular pattern. There were innumerable patches of small and large SOC stocks across the study area, pointing to a pronounced variability of the SOC stock over small distances. On large scales, elevated SOC stocks in the topsoil were observed in the western and southeastern areas, especially with the RFR maps. However, the MLR map also showed much higher stock in the northern and northwestern part compared to the RFR map. The western and southeastern areas correspond to the high elevation part of the watershed (Fig. 1), with SOC stocks varying between 55 and 65 t C ha^{-1}. The remaining areas displayed low (28–40 t C ha^{-1}) to medium (40–55 t C ha^{-1}) SOC stocks. Though land use did not come up as key variable for SOC stocks in the topsoil, it had an indirect link with elevation, being one of the major influencing factors (Fig. 1). In our study area, the density of settlements and adjacent intensively cultivated fields was higher in the lower elevation areas due to the proximity of streams, which provide water for domestic purposes and for the irrigation of crops. Consequently, larger SOC stocks were found in the surface soils that belonged to areas in more remote and elevated parts of the watershed, which thus exhibited less cultivation intensity and larger areas covered by natural vegetation. This is further corroborated by the positive correlation observed between elevation and SOC stock in the topsoil (Supplementary information, SI. 3). The MLR map presents much lower (darker blue) and higher (green and yellow) values compared to the RFR map. This suggests that the former covered a larger range of the SOC stock and possibly allocated higher SOC stocks in the northern and northwestern area compared to the latter.

Fig. 3 presents a different pattern for spatial distribution of the SOC in the subsoil. The subsoil SOC stock map of the RFR displays the largest SOC stocks in the northern and northeastern part, as well as in the lower section of the southern part, though the latter covered only a small area. The largest subsoil SOC stocks are primarily located in the lowest elevation part, which correspond to the predominant area of the Gleysols, characterized by the smallest distance to the stream network. As noted above, this finding was likely additionally influenced by redepositional processes. Schmengler (2010) identified various colluvial deposits at footslope/toeslope level in the Dano catchment. The MLR subsoil map shows larger SOC stocks in the western area, while the remaining parts are mainly characterized by a much lower stock. The spatial distribution of SOC stock observed with the RFR model for the subsoil better reflected the high level of carbon stock observed in the Gleysols located at lower positions, while the MRL attributed the lowest amount to these soils. We suggest that the MLR failed to accurately capture the spatial distribution of the subsoil SOC stocks due to the small number of samples coupled with the existence of a nonlinear pattern which the RFR was able to capture.

4.5. Performance of the RFR and MLR models

In general, the accuracy of the RFR and MLR prediction models (Table 5) based on the R^2 was low (< 30%). The explained variances could not be improved even when some reference soil groups were removed from the dataset and modeling was carried out with the remainder (see Supplementary information, SI. 5). The results of this study regarding the low explained variances are consistent with some existing findings in the literature. Grimm et al. (2008) found only 6% as explained variance for topsoil and 8–25% for subsoil SOC content after using the RFR approach in a tropical island in Panama. Henderson et al. (2005) used a decision tree approach and reported an explained variance of 41% for topsoil SOC and 24% for the subsoil. Wiesmeier et al. (2014) analyzed the spatial distribution of SOC stocks and found 52% of explained variance for the carbon stock based on climate, land use, and environmental variables. Schulp and Verburg (2009) and Schulp et al. (2013) reported 21–43% explained variance for SOC contents and stocks in spite of the wide range of data from soil properties to terrain attributes considered. These authors pointed out that low explained variance for SOC prediction was recorded due to an intrinsic large spatial variability of SOC with the interplay of a large range of factors at local and regional level.

The low explained variance observed in the present study could be attributed to the existence of other environmental and soil parameters affecting SOC stock variability that have not have been investigated in this study. Such parameters may account for specific soil properties, such as soil structural stability, clay mineralogy, sesquioxide composition, as well as other factors beyond the scope of our design, such as socio-ecological impacts in soil resilience (e.g., Linstädter et al. (2016)). In addition, the RMSE obtained in this study is a reflection of errors related to field sampling, laboratory measurement, and statistics, as well as random errors. Since all of the soil properties used in the present study were interpolated by ordinary kriging, it is evident that related errors translated into the estimation of SOC stock. However, preliminary modeling without these soil properties revealed much lower variances (data not shown), proving them key variables to be taken into account. Auxiliary data coming from different sources and different scales infer variability in data quality, as also pointed out by Were et al. (2015). For example, the resampled lithology file was originally produced at a scale of one million; as result, its distribution in the study area might have been too coarse. Further model improvement would require additional explanatory variables at a finer scale, with the consideration of multi- or hyper-scale data in order to account for the possibility of SOC stock being subject to factors operating at different levels of scale (Behrens et al., 2010b; Behrens et al., 2010a).

Comparing the two models, the RFR performed marginally better than the MLR, with higher R^2 values of 14% and 25% for the topsoil and subsoil, respectively. In addition, the former recorded lower RMSEs for both cross-validation and independent validation for all depths compared to the latter (Table 5). This can be attributed to the nonlinear pattern in the SOC stock dataset, which could not be accounted for by

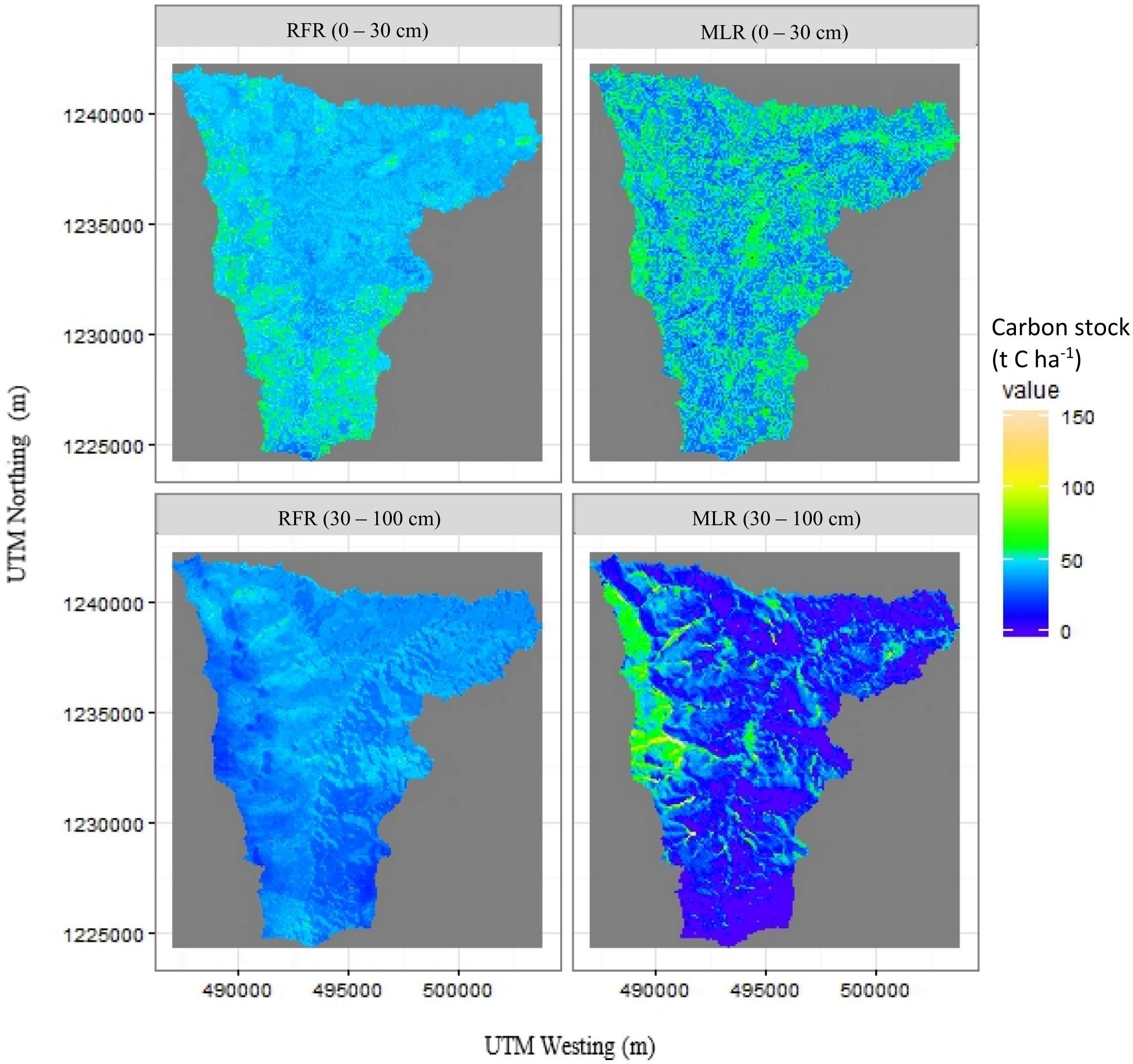

Fig. 3. Distribution of SOC stock in the topsoil (0–30 cm) and subsoil (30–100 cm) based on the RFR and MLR models. RFR: random forest regression, MLR: multiple linear regression.

Table 5
Performance statistics of the RFR and MLR models and general statistics for measured data and SOC stocks of the maps.

			R^2	RMSECV	RMSEPV	ME
Model and validation	Topsoil (0–30 cm)	RFR ($t\,C\,ha^{-1}$)	14.0	13.8	14.9	−0.8
		MLR ($t\,C\,ha^{-1}$)	10.8	14.2	18.8	−3.1
	Subsoil (30–100 cm)	RFR ($t\,C\,ha^{-1}$)	25.6	21.6	23.0	+0.2
		MLR ($t\,C\,ha^{-1}$)	18.23	25.07	34.07	−17.7

			Min	Max	Mean (± SD)	ME
Predicted map versus measured dataset	Topsoil (0–30 cm)	RFR predicted data ($t\,C\,ha^{-1}$)	25.4	58.4	45.2 (± 4.2)	−0.3
		MLR predicted data ($t\,C\,ha^{-1}$)	0	85.2	43.4 (± 9.9)	−2.1
		Measured data ($t\,C\,ha^{-1}$)	11.3	79.2	45.5 (± 14.9)	
	Subsoil (30–100 cm)	RFR predicted data ($t\,C\,ha^{-1}$)	17.3	54.4	36.3 (± 5.8)	+2.6
		MLR predicted data ($t\,C\,ha^{-1}$)	0.0	151.1	22.8 (± 21.1)	−10.9
		Measured data ($t\,C\,ha^{-1}$)	0.3	103.0	33.7 (± 23.9)	

RFR: random forest regression, MLR: multiple linear regression, RMSECV: root mean square error of cross validation, RMSEPV: root mean square error of prediction based on validation set, SD: standard deviation, ME: mean error, −: underestimation, +: overestimation.

the MLR. Other studies have also pointed out MLR's limitations with respect to handling nonlinear patterns in datasets - hence its lower performance compared to machine learning models such as RFR (Hengl et al., 2015; Zhang et al., 2017). Moreover, the error metrics (Table 5) showed that the RMSECV, as well as the RMSEPV from the RFR model, were all slightly lower or comparable to the SD of the measured values. This reveals that the predictions of the models from the RFR were as accurate as the training set, in spite of the low explained variance. A similar trend was also recorded by Were et al. (2015).

The general statistics for the predicted and measured SOC stocks (Table 5) revealed that the predicted minimum value for the RFR maps was larger than the measured one, while the predicted maximum values were lower (Table 5). The opposite was observed with the MLR, whose predictions were larger than the initial range of the measured data. This further explains why the maps of the MLR for both topsoil and subsoil presented much higher (green and yellow) and lower (darker blue) values compared to the RFR maps (Fig. 3). For the RFR, this may be attributed to the fact that the model considered the lowest and highest values of the training data as outliers, as also observed by Were et al. (2015). However, the mean SOC stocks measured for the topsoil (45.5 t C ha^{-1}) were very near to the mean SOC stocks predicted from the map (45.2 t C ha^{-1}). In general, a trend of underestimation was observed with the models when based on the mean error. However, this underestimation tended to be even larger with the MLR applied to the subsoil. In addition, the mean SOC stocks from the MLR map was about two times lower in the subsoil than the mean of the measured stocks, suggesting that pattern recognition with a machine learning method such as RFR might be a better option in the context of a limited dataset with an inherent nonlinear relationship.

5. Conclusion

This study provides insights into the quantitative status of topsoil (0–30 cm) and subsoil (30–100 cm) SOC stocks in the Dano catchment in different land-use systems and across different soil reference groups. Additionally, the driving factors and spatial distribution of the topsoil and subsoil carbon stock was investigated. RFR and MLR modeling were used as statistical methods for identifying these factors and for mapping the spatial distribution of SOC stocks for the topsoil and subsoil carbon stock.

The results indicate only a marginal difference between SOC stocks in the savannah and cropland, with most of the reference soil groups related to the former recording a slightly larger carbon stock in the topsoil. We attributed these findings to both site preferences by farmers for better sites selected for cropping, as well as advanced land-use degradation of the savannah land with increasing human grazing pressure. The findings also show a significantly higher SOC stock in the croplands compared to the savannah lands in the subsoil, suggesting a higher vulnerability of the former to intensive translocation processes induced by elevated precipitation events under a tropical climate.

The topsoil SOC stock variability was primarily affected by soil properties (e.g., silt content), followed by the soil moisture distribution with the wetness index. Sites at higher elevation exhibited elevated SOC stocks in the surface soil. This disentanglement was due to landscape controls on population density and cropping intensity, which were both concentrated in the lowlands. The subsoil SOC stock was mainly associated with the topographical orientation represented by the slope aspect. RFR performed slightly better than MLR in predicting the spatial distribution of the topsoil and subsoil SOC stock, as the latter could not account for the nonlinear pattern within the data.

Our findings reinforce the view that the semi-arid ecosystems of West Africa still offer a significant opportunity for carbon sequestration to offset ongoing C losses, with the spatial distribution of the topsoil SOC stock driven not only by soil and climate, but also by landscape-specific human pressure on ecosystems.

Acknowledgements

The authors would like to thank: (1) the German Federal Ministry of Education and Research (01LG1202E) for granting the financial support to carry out this research; and (2) the West African Science Service Center on Climate Change and Adapted Land use for making available all logistical resources for the field survey. We are also grateful to the field assistants and laboratory technicians for their support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.catena.2018.04.013.

References

Adhikari, K., Hartemink, A.E., Minasny, B., Kheir, R.B., Greve, M.B., Greve, M.H., 2014. Digital mapping of soil organic carbon contents and stocks in Denmark. PLoS One 9 (8), e105519.
Akpa, S.I., Odeh, I.O., Bishop, T.F., Hartemink, A.E., Amapu, I.Y., 2016. Total soil organic carbon and carbon sequestration potential in Nigeria. Geoderma 271 (Supplement C), 202–215. http://dx.doi.org/10.1016/j.geoderma.2016.02.021.
Albaladejo, J., Ortiz, R., Garcia-Franco, N., Navarro, A., Almagro, M., Pintado, J., Martínez-Mena, M., 2013. Land use and climate change impacts on soil organic carbon stocks in semi-arid Spain. J. Soils Sediments 13 (2), 265–277. http://dx.doi.org/10.1007/s11368-012-0617-7.
Allen, D.E., Pringle, M.J., Bray, S., Hall, T.J., Reagain, P.O., Phelps, D., Cobon, D.H., Bloesch, P.M., Dalal, R.C., 2013. What determines soil organic carbon stocks in the grazing lands of north-eastern Australia? Soil Res. 51, 695–706. http://dx.doi.org/10.1071/SR13041.
Anikwe, M., 2010. Carbon storage in soils of southeastern Nigeria under different management practices. Carb. Balance Manag. 5 (1), 5.
Augustin, C., Cihacek, L.J., 2016. Relationships between soil carbon and soil texture in the northern great plains. Soil Sci. 181 (8), 386–392.
Azlan, A., Aweng, E.R., Ibrahim, C.O., 2011. The correlation between total organic carbon (TOC), organic matter and water content in soil collected from different land use of Kota Bharu, Kelantan. J. Appl. Sci. Res. 7 (7), 915.
Bationo, A., Buerkert, A., 2001. Soil organic carbon management for sustainable land use in Sudano-Sahelian West Africa. Nutr. Cycl. Agroecosyst. 61 (1–2), 131–142.
Bationo, A., Kihara, J., Vanlauwe, B., Waswa, B., Kimetu, J., 2007. Soil organic carbon dynamics, functions and management in West African agro-ecosystems. Agric. Syst. 94 (1), 13–25.
Batjes, N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 47, 151–163.
Batjes, N.H., 2001. Options for increasing carbon sequestration in West African soils: an exploratory study with special focus on Senegal. Land Degrad. Dev. 12 (2), 131–142. http://dx.doi.org/10.1002/ldr.444.
Batjes, N.H., 2008. Mapping soil carbon stocks of Central Africa using SOTER. Geoderma 146 (1–2), 58–65. http://dx.doi.org/10.1016/j.geoderma.2008.05.006.
Batjes, N.H., Sombroek, W.G., 1997. Possibilities for carbon sequestration in tropical and sub-tropical soils. Glob. Chang. Biol. 3, 161–173.
Baldock, J.A., Hawke, B., Sanderman, J., Macdonald, L.M., 2013. Predicting contents of carbon and its component fractions in Australian soils from diffuse reflectance mid-infrared spectra. Soil Res. 51 (8), 577–595. http://dx.doi.org/10.1071/SR13077.
Behrens, T., Schmidt, K., Zhu, A.X., Scholten, T., 2010a. The ConMap approach for terrain-based digital soil mapping. Eur. J. Soil Sci. 61 (1), 133–143.
Behrens, T., Zhu, A.-X., Schmidt, K., Scholten, T., 2010b. Multi-scale digital terrain analysis and feature selection for digital soil mapping. Geoderma 155 (3), 175–185.
Blin, J., Sidibe, S., 2012. Caractérisation et amélioration d'un foyer de cuisson de "dolo" équipé d'un brûleur à huile végétale (jatropha). In: Proceding CIFEM2 2012. Ouagadougou, pp. 1–6 s.n.
Bellon-Maurel, V., McBratney, A., 2011. Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils–critical review and research perspectives. Soil Biol. Biochem. 43 (7), 1398–1410.
Boakye-Danquah, J., Antwi, E.K., Osamu, S., Abekoe, M.K., Takeuchi, K., 2014. Impact of farm management practices and agricultural land use on soil organic carbon storage potential in the savannah ecological zone of northern Ghana. J. Disaster Res. 9 (4), 484–500.
Bornemann, L., Welp, G., Amelung, W., 2010. Particulate organic matter at the field scale: rapid acquisition using mid-infrared spectroscopy. Soil Sci. Soc. Am. J. 74 (4), 1147–1156. http://dx.doi.org/10.2136/sssaj2009.0195.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. http://dx.doi.org/10.1023/A:1010933404324.
Brevik, E., 2013. The potential impact of climate change on soil properties and processes and corresponding influence on food security. Agriculture 2013 (3), 398–417.
Bruun, B., Elberling, B., de Neergaard, A., Magid, J., 2013. Organic carbon dynamics in different soil types after conversion of forest to agriculture. Land Degrad. Dev. 26 (3), 272–283. http://dx.doi.org/10.1002/ldr.2205.
Bureau National des sols, 2000. Etude morphologiaue des provinces de la Bougouriba et du Ioba, Echelle 1/100 000, Rapport technique n° 121, Burkina Faso.
Burke, I.C., Yonker, C.M., Parton, W.J., Cole, C.V., Schimel, D.S., Flach, K., 1989. Texture,

climate, and cultivation effects on soil organic matter content in U.S. grassland soils. Soil Sci. Soc. Am. J. 53 (3), 800–805. http://dx.doi.org/10.2136/sssaj1989.03615995005300030029x.
Chaplot, V., Bouahom, B., Valentin, C., 2010. Soil organic carbon stocks in Laos: spatial variations and controlling factors. Glob. Chang. Biol. 16 (4), 1380–1393. http://dx.doi.org/10.1111/j.1365-2486.2009.02013.x.
Callo-Concha, D., Gaiser, T., Ewert, F., 2012. Farming and cropping systems in the West African Sudanian Savanna. In: WASCAL Research Area: Northern Ghana, Southwest Burkina Faso and Northern Benin. ZEF Working Paper Series, No. 100, Bonn, 2012.
Cord, A., Conrad, C., Schmidt, M., Dech, S., 2010. Standardized FAO-LCCS land cover mapping in heterogeneous tree savannas of West Africa. J. Arid Environ. 74 (9), 1083–1091.
Cornelissen, J., Aerts, R., Cerabolini, B., Werger, M., Van Der Heijden, M., 2001. Carbon cycling traits of plant species are linked with mycorrhizal strategy. Oecologia 129 (4), 611–619.
Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A., Totterdell, I.J., 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. Nature 408 (6809), 184–187.
Davy, M.C., Koen, T.B., 2013. Variations in soil organic carbon for two soil types and six land uses in the Murray Catchment, New South Wales, Australia. Soil Res. 51, 631–644. http://dx.doi.org/10.1071/SR12353.
Denef, K., Plante, A.F., Six, J., 2009. Characterization of soil organic matter. In: Kutsch, W.L., Bahn, M., Heinemeyer, A. (Eds.), Soil Carbon Dynamics: An Integrated Methodology. Cambridge University Press, Cambridge.
Doetterl, S., Stevens, A., van Oost, K., Quine, T.A., van Wesemael, B., 2013. Spatially-explicit regional-scale prediction of soil organic carbon stocks in cropland using environmental variables and mixed model approaches. Geoderma 204, 31–42.
Doetterl, S., Kearsley, E., Bauters, M., Hufkens, K., Lisingo, J., Baert, G., Verbeeck, H., Boeckx, P., 2015. Aboveground vs. belowground carbon stocks in African tropical lowland rainforest: drivers and implications. PLoS One 10 (11), e0143209.
Dorji, T., Odeh, I.O.A., Field, D.J., 2014. Vertical distribution of soil organic carbon density in relation to land use/cover, altitude and slope aspect in the eastern Himalayas. Landscape 3 (4), 1232–1250.
Embrechts, J., Sys, C., 1988. Genesis of subsurface horizons of a soil catena in a humid tropical climate (Yaounde, Cameroon). Catena 15 (1), 53–63.
Eze, P.N., Udeigwe, T.K., Meadows, M.E., 2014. Plinthite and its associated evolutionary forms in soils and landscapes: a review. Pedosphere 24 (2), 153–166.
Fierer, N., Allen, A.S., Schimel, J.P., Holden, P.A., 2003. Controls on microbial CO_2 production: a comparison of surface and subsurface soil horizons. Glob. Chang. Biol. 9 (9), 1322–1332. http://dx.doi.org/10.1046/j.1365-2486.2003.00663.x.
Florinsky, I.V., Eilers, R.G., Manning, G.R., Fuller, L.G., 2002. Prediction of soil properties by digital terrain modelling. Environ. Model. Softw. 17 (3), 295–311.
Fontaine, S., Barot, S., Barre, P., Bdioui, N., Mary, B., Rumpel, C., 2007. Stability of organic carbon in deep soil layers controlled by fresh carbon supply. Nature 450 (7167), 277–280. http://dx.doi.org/10.1038/nature06275.
Forkuor, G., 2014. Agricultural Land Use Mapping in West Africa Using Multi-sensor Satellite Imagery. University of Wuerzburg, Wuerzburg, Germany, pp. 191.
Feller, C., Beare, M.H., 1997. Physical control of soil organic matter dynamics in the tropics. Geoderma 79 (1–4), 69–116. http://www.sciencedirect.com/science/article/pii/S0016706197000396.
Gidigasu, M., 2012. Laterite Soil Engineering: Pedogenesis and Engineering Principles. Elsevier Scientific Publishing, Amsterdam, The Netherlands.
Gonin, A., Tallet, B., 2012. Changements spatiaux et pratiques pastorales: les nouvelles voies de la transhumance dans l'Ouest du Burkina Faso. Cah. Agric. 21 (6), 448–454.
Grinand, C., Barthes, B.G., Brunet, D., Kouakoua, E., Arrouays, D., Jolivet, C., Caria, G., Bernoux, M., 2012. Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS). Eur. J. Soil Sci. 63 (2), 141–151. http://dx.doi.org/10.1111/j.1365-2389.2012.01429.x.
Grimm, R., Behrens, T., Marker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - digital soil mapping using Random Forests analysis. Geoderma 146 (1–2), 102–113. http://dx.doi.org/10.1016/j.geoderma.2008.05.008.
Guo, P.-T., Li, M.-F., Luo, W., Tang, Q.-F., Liu, Z.-W., Lin, Z.-M., 2015. Digital mapping of soil organic matter for rubber plantation at regional scale: an application of random forest plus residuals kriging approach. Geoderma 237, 49–59.
Hartge, K., Horn, R., 1989. Die physikalische Untersuchung von Böden. Enke Verlag, Stuttgart.
Hengl, T., Heuvelink, G.B., Kempen, B., Leenaars, J.G., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., Mendes de Jesus, J., Tamene, L., Tondoh, J.E., 2015. Mapping soil properties of Africa at 250 m resolution: Random Forests significantly improve current predictions. PLoS One 10 (6), e0125814. http://dx.doi.org/10.1371/journal.pone.0125814.
Henderson, B.L., Bui, E.N., Moran, C.J., Simon, D.A., 2005. Australia-wide predictions of soil properties using decision trees. Geoderma 124 (3), 383–398.
Hien, E., Ganry, F., Hien, V., Oliver, R., 2003. Carbon dynamics in a soil of southwestern Burkina savannah under the effect of the cultivation and farming practices. In: African Savannahs: The Changing Spaces, Actors Face New Challenges. May 2002, Garoua, Cameroon. Prasac, N'Djamena, Chad - CIRAD, Montpellier, France.
Hien, E., Ganry, F., Oliver, R., 2006. Carbon sequestration in a savannah soil in southwestern Burkina as affected by cropping and cultural practices. Arid Land Res. Manag. 20 (2), 133–146.
Hottin, G., Ouedraogo, O.F., 1992. Carte Géologique du Burkina Faso, 2 édition. ((B.M.G.B.). Échelle: 1: 1.000.000).
IUSS, ISRIC, FAO, 2006. World reference base for soil resources-a framework for international classification, correlation and communication. In: World Soil Resources, Report 103. FAO, Rome, Italy.
Janik, L.J., Skjemstad, J.O., Shepherd, K.D., Spouncer, L.R., 2007. The prediction of soil carbon fractions using mid-infrared-partial least square analysis. Soil Res. 45 (2), 73–81. http://dx.doi.org/10.1071/SR06083.
Jobbágy, E.G., Jackson, R., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. 10 (2), 423–436.
Jungerius, P.D., van den Ancker, J.A., Mücher, H.J., 1999. The contribution of termites to the microgranular structure of soils on the Uasin Gishu Plateau, Kenya. Catena 34 (3), 349–363.
Kilasara, M., 2010. Selection and use of soil characteristics in digital soil mapping in Tanzania. In: 19th World Congress of Soil Science. August 2. Brisbane, Australia.
Kögel-Knabner, I., Amelung, W., 2014. Dynamics, chemistry, and preservation of organic matter in soils. In: Reference Module in Earth Systems and Environmental Sciences, from Treatise on Geochemistry. 12. pp. 157–215I.
Kuhn, M., 2015. Caret: classification and regression training. In: Astrophysics Source Code Library. 1. pp. 5003.
Kumar, S., Lal, R., 2011. Mapping the organic carbon stocks of surface soils using local spatial interpolator. J. Environ. Monit. 13 (11), 3128–3135.
Lal, R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science 304 (5677), 1623–1627. http://dx.doi.org/10.1126/science.1097396.
Lal, R., 2008. Carbon sequestration. Philos. Trans. R. Soc. B Biol. Sci. 363 (1492), 815–830. http://dx.doi.org/10.1098/rstb.2007.2185.
Lin, G., Stralberg, D., Gong, G., Huang, Z., Ye, W., Wu, L., 2013. Separating the effects of environment and space on tree species distribution: from population to community. PLoS One 8 (2), e56171.
Ladd, B., Laffan, S.W., Amelung, W., Peri, P.L., Silva, Lucas C.R., Gervassi, P., Bonser, S.P., Navall, M., Sheil, D., 2013. Estimates of soil carbon concentration in tropical and temperate forest and woodland from available GIS data on three continents. Glob. Ecol. Biogeogr. 22, 461–469.
Lorenz, K., Lal, R., 2005. The depth distribution of soil organic carbon in relation to land use and management and the potential of carbon sequestration in subsoil horizons. Adv. Agron. 88, 35–66.
Linstädter, A., Kuhn, A., Naumann, C., Rasch, S., Sandhage-Hofmann, A., Amelung, W., Jordaan, J., Du Preez, C., Bollig, M., 2016. Assessing the resilience of a real-world social-ecological system: lessons from a multidisciplinary evaluation of a South African pastoral system. Ecol. Soc. 21 (3).
Manning, P., Vries, F.T., Tallowin, J.R.B., Smith, R., Mortimer, S.R., Pilgrim, E.S., Harrison, K.A., Wright, D.G., Quirk, H., Benson, J., 2015. Simple measures of climate, soil properties and plant traits predict national-scale grassland soil carbon stocks. J. Appl. Ecol. 52 (5), 1188–1196.
Mao, D.H., Wang, Z.M., Li, L., Miao, Z.H., Ma, W.H., Song, C.C., Ren, C.Y., Jia, M.M., 2015. Soil organic carbon in the Sanjiang Plain of China: storage, distribution and controlling factors. Biogeosciences 12 (6), 1635–1645.
Minasny, B., Malone, B.P., McBratney, A.B., Angers, D.A., Arrouays, D., Chambers, A., Chaplot, V., Chen, Z.-S., Cheng, K., Das, B.S., Field, D.J., Gimona, A., Hedley, C.B., Hong, S.Y., Mandal, B., Marchant, B.P., Martin, M., McConkey, B.G., Mulder, V.L., O'Rourke, S., Richer-de-Forges, A.C., Odeh, I., Padarian, J., Paustian, K., Pan, G., Poggio, L., Savin, I., Stolbovoy, V., Stockmann, U., Sulaeman, Y., Tsui, C.-C., Vågen, T.-G., van Wesemael, B., Winowiecki, L., 2017. Soil carbon 4 per mille. Geoderma 292 (Supplement C), 59–86. http://dx.doi.org/10.1016/j.geoderma.2017.01.002.
McBratney, A.B., Santos, M.M., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1), 3–52.
Meersmans, J., de Ridder, F., Canters, F., de Baets, S., van Molle, M., 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). Geoderma 143 (1), 1–13.
Mishra, U., Lal, R., Slater, B., Calhoun, F., Liu, D., van Meirvenne, M., 2009. Predicting soil organic carbon stock using profile depth distribution functions and ordinary kriging. Soil Sci. Soc. Am. J. 73 (2), 614–621.
Oueslati, I., Allamano, P., Bonifacio, E., Claps, P., 2013. Vegetation and topographic control on spatial variability of soil organic carbon. Pedosphere 23 (1), 48–58.
Page, K.L., Dalal, R.C., Pringle, M.J., Bell, M., Dang, Y.P., Radford, B., Bailey, K., 2013. Organic carbon stocks in cropping soils of Queensland, Australia, as affected by tillage management, climate, and soil characteristics. Soil Res. 51, 596–607. http://dx.doi.org/10.1071/SR12225.
Percival, H.J., Parfitt, R.L., Scott, N.A., 2000. Factors controlling soil carbon levels in New Zealand grasslands: is clay content important? Soil Sci. Soc. Am. J. 64 (5), 1623–1630. http://dx.doi.org/10.2136/sssaj2000.6451623x.
Preger, A.C., Kösters, R., Du Preez, C.C., Brodowski, S., Amelung, W., 2010. Carbon sequestration in secondary pasture soils: a chronosequence study in the South African Highveld. Eur. J. Soil Sci. 61 (4), 551–562. http://dx.doi.org/10.1111/j.1365-2389.2010.01248.x.
R Core Team, 2015. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria. Available. http://www.r-project.org/.
Rad, M.R.P., Toomanian, N., Khormali, F., Brungard, C.W., Komaki, C.B., Bogaert, P., 2014. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. Geoderma 232 (234), 97–106.
Reeuwijk, V.L., 2006. Procedures for soil analysis. In: Technical Report 9. Wageningen, Netherlands, ISRIC – World Soil Information, 7th edition. .
Reeves, J.B., 2010. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: where are we and what needs to be done? Geoderma 158 (1), 3–14. http://dx.doi.org/10.1016/j.geoderma.2009.04.005.
Rhodes, C.J., 2016. The 2015 Paris climate change conference: COP21. Sci. Prog. 99 (1), 97–104.
Rodríguez-Moreno, V.M., Bullock, S.H., 2014. Vegetation response to hydrologic and geomorphic factors in an arid region of the Baja California Peninsula. Environ. Monit.

Assess. 186 (2), 1009–1021.
Saiz, G., Bird, M.I., Domingues, T., Schrodt, F., Schwarz, M., Feldpausch, T.R., Veenendaal, E., Djagbletey, G., Hien, F., Compaore, H., Diallo, A., Lloyd, J., 2012. Variation in soil carbon stocks and their determinants across a precipitation gradient in West Africa. Glob. Chang. Biol. 18 (5), 1670–1683. http://dx.doi.org/10.1111/j.1365-2486.2012.02657.x.
Sakin, E., 2012. Relationships between of carbon, nitrogen stocks and texture of the Harran Plain soils in southeastern Turkey. Bulg. J. Agric. Sci. 18 (4), 626–634.
Schmengler, A.C., 2010. Modeling Soil Erosion and Reservoir Sedimentation at Hillslope and Catchment Scale in Semi-arid Burkina Faso, Bonn. (150 pp.).
Schulp, E., Verburg, P.H., 2009. Effect of land use history and site factors on spatial variation of soil organic carbon across a physiographic region. Agric. Ecosyst. Environ. 133 (1), 86–97.
Schulp, C.J.E., Verburg, P.H., Kuikman, P.J., Nabuurs, G.-J., Olivier, J.G.J., de Vries, W., Veldkamp, T., 2013. Improving national-scale carbon stock inventories using knowledge on land use history. Environ. Manag. 51 (3), 709–723.
Selige, T., Böhner, J., Schmidhalter, U., 2006. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. Geoderma 136, 235–244.
Singh, S.K., Pandey, C.B., Sidhu, G.S., Sarkar, D., Sagar, R., 2011. Concentration and stock of carbon in the soils affected by land uses and climates in the western Himalaya, India. Catena 87 (1), 78–89.
Smith, P., 2008. Land use change and soil organic carbon dynamics. Nutr. Cycl. Agroecosyst. 81 (2), 169–178. http://dx.doi.org/10.1007/s10705-007-9138-y.
Taghizadeh-Mehrjardi, R., Nabiollahi, K., Minasny, B., Triantafilis, J., 2015. Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. Geoderma 253, 67–77.
UNEP, 2006. Global Environment Outlook. 3. http://www.grida.no/geo/geo3/english/149.htm.
Wang, S., Wang, X., Guo, H., Fan, W., Lv, H., Duan, R., 2013. Distinguishing the importance between habitat specialization and dispersal limitation on species turnover. Ecol. Evol. 3, 3545–3553.
Were, K., Bui, D.T., Dick, Ø.B., Singh, B.R., 2015. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecol. Indic. 52, 394–403.
Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I., 2011. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 340 (1–2), 7–24. http://dx.doi.org/10.1007/s11104-010-0425-z.
Wiesmeier, M., Barthold, F., Spörlein, P., Geuß, U., Hangen, E., Reischl, A., Schilling, B., Angst, G., von Lützow, M., Kögel-Knabner, I., 2014. Estimation of total organic carbon storage and its driving factors in soils of Bavaria (southeast Germany). Geoderma Reg. 1, 67–78.
Yira, Y., Diekkrüger, B., Steup, G., Bossa, A.Y., 2016. Modeling land use change impacts on water resources in a tropical West African catchment (Dano, Burkina Faso). J. Hydrol. 537, 187–199.
Yoni, M., Hien, V., Abbadie, L., Serpantié, G., 2005. Dynamique de la matière organique du sol dans les savanes soudaniennes du Burkina Faso. Cah. Agric. 14 (6), 525–532.
Yu, C., 2011. Parametric tests. Available from. http://www.creative-wisdom.com/teaching/WBI/parametric_test.shtml.
Zadorova, T., Žížala, D., Penížek, V., Čejková, Š., 2014. Relating extent of colluvial soils to topographic derivatives and soil variables in a Luvisol sub-catchment, Central Bohemia, Czech Republic. Soil Water Res. 9 (2), 47–57.
Zhang, C., McGrath, D., 2004. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. Geoderma 119 (3), 261–275.
Zhang, P., Shao, M., 2014. Spatial variability and stocks of soil organic carbon in the Gobi desert of northwestern China. PLoS One 9 (4), e93584. http://dx.doi.org/10.1371/journal.pone.0093584.
Zhang, H., Wu, P., Yin, A., Yang, X., Zhang, M., Gao, C., 2017. Prediction of soil organic carbon in an intensively managed reclamation zone of eastern China: a comparison of multiple linear regressions and the random forest model. Sci. Total Environ. 592, 704–713.
Zimmermann, M., Leifeld, J., Schmidt, M.W.I., Smith, P., Fuhrer, J., 2007. Measured soil organic matter fractions can be related to pools in the RothC model. Eur. J. Soil Sci. 58 (3), 658–667. http://dx.doi.org/10.1111/j.1365-2389.2006.00855.x.
Zougmoré, R., Ouattara, K., Mando, A., Ouattara, B., 2004. Rôle des nutriments dans le succès des techniques de conservation des eaux et des sols (cordons pierreux, bandes enherbées, zaï et demi-lunes) au Burkina Faso. Sci. Change. Planét. 15 (1), 41–48.