Geoderma 263 (2016) 216–225

Contents lists available at ScienceDirect

Geoderma

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

Mapping of soil properties and land degradation risk in Africa using MODIS reflectance

Tor-G. Vågen [a,*], Leigh A. Winowiecki [b], Jerome E. Tondoh [a], Lulseged T. Desta [b], Thomas Gumbricht [c]

[a] World Agroforestry Centre (ICRAF), P.O. Box 30677-00100 GPO, Nairobi, Kenya
[b] International Center for Tropical Agriculture (CIAT), P.O. Box 823-00621, Nairobi, Kenya
[c] Karttur AB, Birger Jarlsgatan 102 B 3TR, 114 20 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form 3 June 2015
Accepted 30 June 2015
Available online 6 July 2015

Keywords:
Digital soil mapping
SOC
pH
Acidity
Land degradation
Africa

ABSTRACT

There is a need for up-to-date assessments and maps of soil properties and land health at scales relevant for decision-making and management, including for properties that are dynamic and hence change in response to management. Also, there is a need for approaches to soil mapping that capture the ever increasing effects that humans are having on the environment in general and specifically on soil properties worldwide. In this paper, we develop models for digital soil mapping based on remote sensing data from the Moderate Resolution Imaging Spectroradiometer (MODIS) platform for Africa. The article presents maps of soil organic carbon (SOC), pH, sand and sum of exchangeable bases, as well as prevalence of root-depth restrictions in the upper 50 cm of the soil profile. Prediction models were developed based on spatially balanced field survey data, representing all major climate zones on the continent. The prediction models for soil property mapping performed well, with overall RMSEP values of 10.6, 0.34, 9.1, and 6.5 for SOC, pH, sand, and sum of bases, respectively. The accuracy of the prediction model for root-depth restrictions was 77%, with an AUC of 0.85 and Cohen's kappa value of 0.52 when averaged across predictions run on independent test data. The methods and maps developed can provide much improved identification of soil and land health constraints, and spatial targeting of land management interventions at various scales, informing both policy and practice.

1. Introduction

It is widely recognized that soil degradation is becoming increasingly prevalent in many parts of sub-Saharan Africa (SSA), but its spatial extents and severity from local to continental scales are in general poorly understood (Bassett and Zuéli, 2000; Warren, 2002). Similarly, although soil and ecosystem health are recognized as important components of land and agricultural productivity, stakeholders such as farmers and rural communities, policy makers and governments often lack scientifically based information that is spatially and contextually explicit enough to effectively target soil and land management strategies that enhance and maintain critical ecosystem services (Chan et al., 2006). Consistent biophysical data and models are needed to answer questions related to the dynamics of key indicators of ecosystem health at multiple spatial and temporal scales, including feedback loops within nested hierarchies of socioeconomic and biophysical environments. This is particularly critical at present since humans are increasingly shaping the environment and altering the soil system, which means that models and maps are needed that capture the effects of anthropogenic impacts on soils (Grunwald et al., 2011). Patterns of various soil properties in landscapes are complex (Scull et al., 2003), and efforts to achieve reliable estimates (i.e. reduce uncertainty) of statistical prediction models for mapping of soil properties across a wide range of soil conditions require large sample sizes. Also, consistent sampling designs are needed, as well as harmonized soil analytical methods across a wide range of datasets. Indeed, some of the most important constraints to soil mapping to date, particularly on the African continent, include (i) a general lack of or sparse soil data, (ii) dated information, and (iii) a lack of consistency in available data, both in terms of field and laboratory methods.

The main reasons for increases in land degradation are often quoted as a complex nexus of overpopulation, poverty, overgrazing and/or poor agricultural practices and lack of appropriate policies (Drechsel et al., 2001; Duraiappah, 1998), but little is known about the interactions between these drivers and the biophysical environment, including biophysical stability domains (Gunderson, 2000). Land degradation also has important implications for climate change as it may lead to increases in CO_2 emissions through for instance deforestation, reductions in above- and below-ground storage of carbon (C), and through its influences on the ability of ecosystems to regulate soil-vegetation-atmosphere transfer (SVAT) processes (Cao et al., 2001) and to deliver vital ecosystem services. In many cases, ecological changes are strongly accelerated due to positive feedback mechanisms at various levels of

* Corresponding author.
E-mail address: t.vagen@cgiar.org (T.-G. Vågen).

http://dx.doi.org/10.1016/j.geoderma.2015.06.023

scale and between SVAT processes and the socioeconomic system that destabilize the system and lead to its collapse (Alcock, 2003).

Most studies involving the prediction of soil properties based on remote sensing have been conducted based on data collected at the plot scale (<1 km^2) (Mulder et al., 2011). Other efforts at mapping soil properties are based on dated soil survey reports that often lack accurate geospatial locations and where a number of different analytical techniques have been used, confounding analysis (e.g. Hengl et al., 2014). In contrast, this study draws on a coordinated collection of field data on a range of land health metrics, including soil condition and land degradation status, using a network of land degradation surveillance sites in SSA. The sites were sampled and characterized using consistent sampling protocols referred to as the Land Degradation Surveillance framework (LDSF) implemented as part of multiple projects during the period from January 2010 to July 2014, including the Bill and Melinda Gates funded Africa Soil Information Service (AfSIS) project (Vågen et al., 2010). These methods and projects have resulted in the availability of data at nested spatial scales, from plots (1000 m^2), to sampling clusters (1 km^2), and sites (100 km^2) covering the major climate zones of SSA. In the current paper we present an approach that combines the use of laboratory soil reference measurements, soil spectroscopy, remote sensing, and statistical modeling to predict and map both soil functional properties and root-depth restrictions. The aim of the study was to map soil functional properties and land degradation risk for SSA at a spatial scale of 500 m by applying data from the moderate resolution imaging spectroradiometer (MODIS) sensor.

The approach used in this study differs from traditional geostatistical models often used for soil mapping (e.g. Hengl et al., 2004; Heuvelink and Webster, 2001) which assume that points closer together are more related than points farther apart. As this assumption presents some challenges (Curriero, 2007) we used an approach where no assumptions were made in terms of the proximity of sampling points, but rather relied on the spectral properties of individual MODIS pixels by training statistical models based on a network of LDSF sites covering different (agro-)ecosystems in SSA. Prediction models were developed for the mapping of soil organic carbon (SOC), pH, sum of exchangeable bases (SB), and sand, which are important soil properties for management of soil health.

We also developed models to determine the likelihood of root-depth restrictions in each MODIS pixel. Effective rooting depth for agricultural crops such as wheat and maize range from 60 to 180 cm when grown in unrestricted conditions (Kirkegaard and Lilley, 2007), and hence root-depth restrictions in the upper 50 cm of the soil profile present challenges for agricultural management by potentially limiting plant growth (Unger and Kaspar, 1994) and soil infiltration capacity both in agricultural and forest systems (Lull, 1959). The developed maps set the stage for the development of contextual indicators of soil and ecosystem health that are sensitive to land use and hence applicable for spatially explicit monitoring of ecosystem health. The digital soil property maps presented are sensitive to the effects of land management and hence better reflect on-the-ground realities facing farmers and land managers.

2. Materials and methods

2.1. Field data collection

A network of 114 LDSF sites (Fig. 1), each covering a 10 by 10 km (100 km^2) area, were surveyed between February 2010 and June 2013, 62 of the sites representing a stratified random sample of

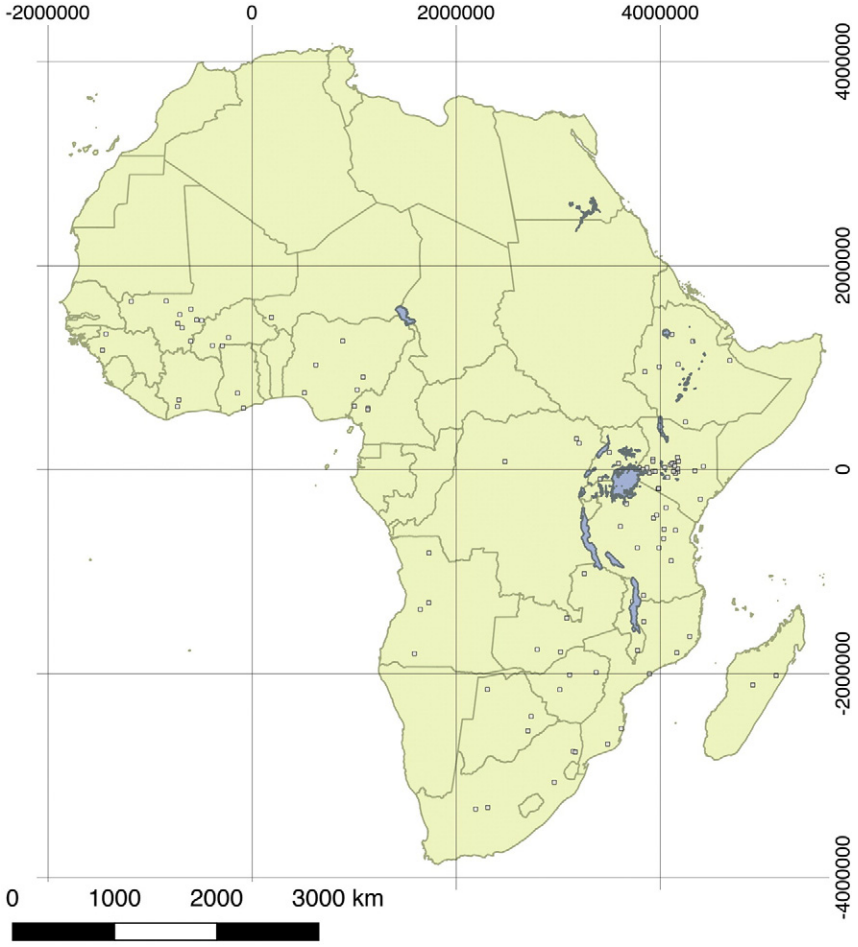

Fig. 1. Map of Africa showing the location of the LDSF sites (N = 114) included in the current study (each square on the map represents a 100 km^2 site, with 160 sampling plots). Projection is sinusoidal.

African landscapes sampled as part of the Africa Soil Information Service (AfSIS) project (Vågen et al., 2010). Major climate zones (Köppen, 1918; Kottek et al., 2006; Rubel and Kottek, 2010) were used to stratify the sampling of the AfSIS sites, while 52 additional sites were sampled as part of other projects applying the LDSF. These projects are usually applying the LDSF in the context of for example assessments of rangeland health, forest transition zones or land degradation and hence the sites are purposefully located to cover parts of project areas, either for baseline assessment or monitoring purposes. Within each LDSF site, cluster (N = 16) and plot (N = 160) locations were generated using a stratified random sampling approach (Vågen et al., 2010, 2013), making a total of 160 sampling plots in each site. The clusters represent 1 km^2 areas, each cluster consisting of 10 (1000 m^2) plots, which again have four sub-plots, each 100 m^2 in size.

A total of 10,473 topsoil samples were available for this study, collected at 0–20 cm depth from 71 of the 114 sites, while data on root-depth restrictions in the upper 50 cm of the soil profile (RDR50) was available for all 114 sites. The measurements of RDR50 were done using a soil auger in each LDSF sub-plot and recording the depth at which restrictions made it impossible to auger further within the upper 50 cm of the soil profile. If 50 cm depth was reached, this was recorded as "no restriction".

2.2. Laboratory analysis

Soil analyses were conducted using traditional wet chemistry methods for the determination of SOC (dry combustion), pH (1:1

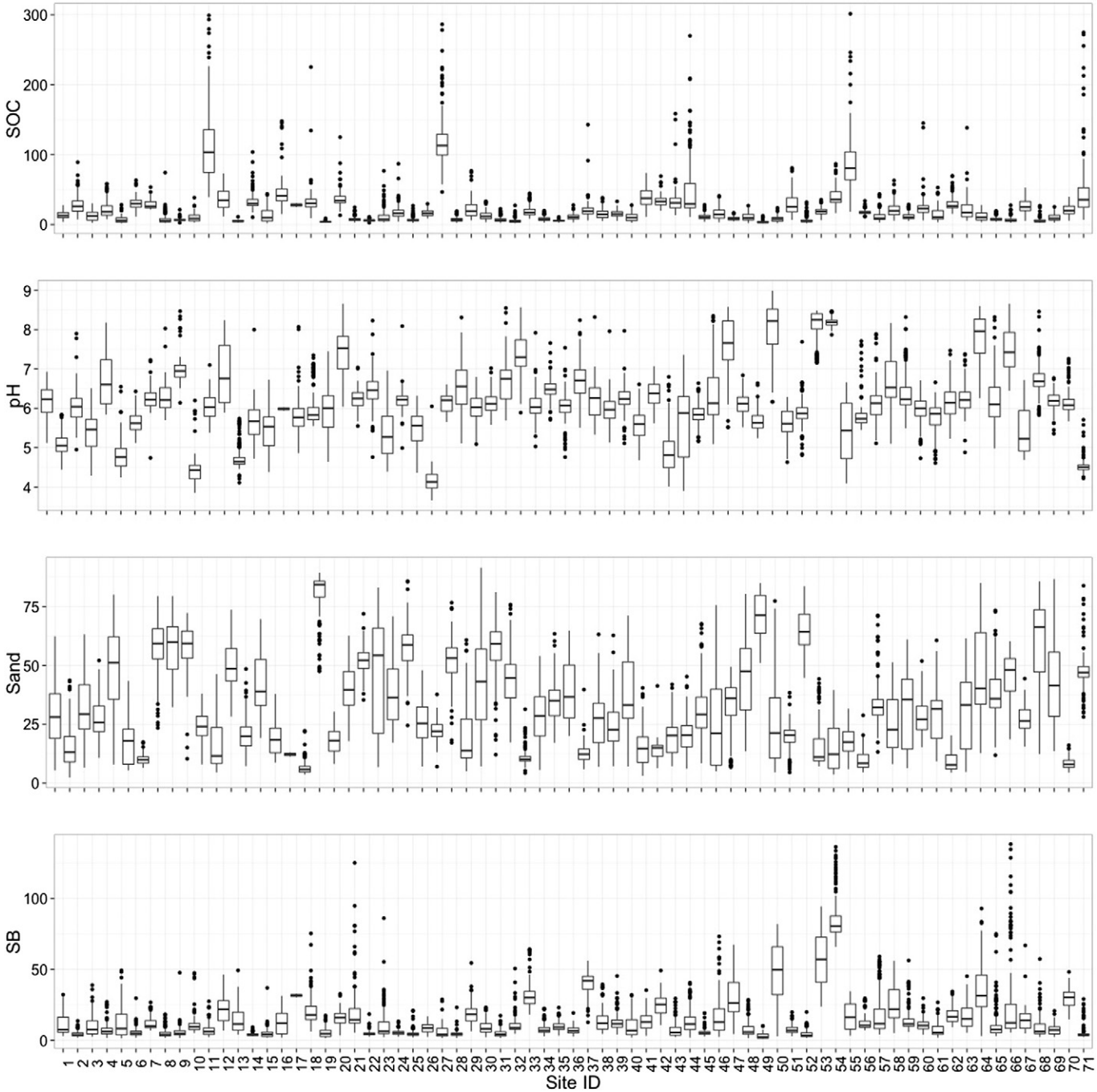

Fig. 2. Boxplots showing a summary of lab measured topsoil properties (0–20 cm depth) across the 71 sites with soil data.

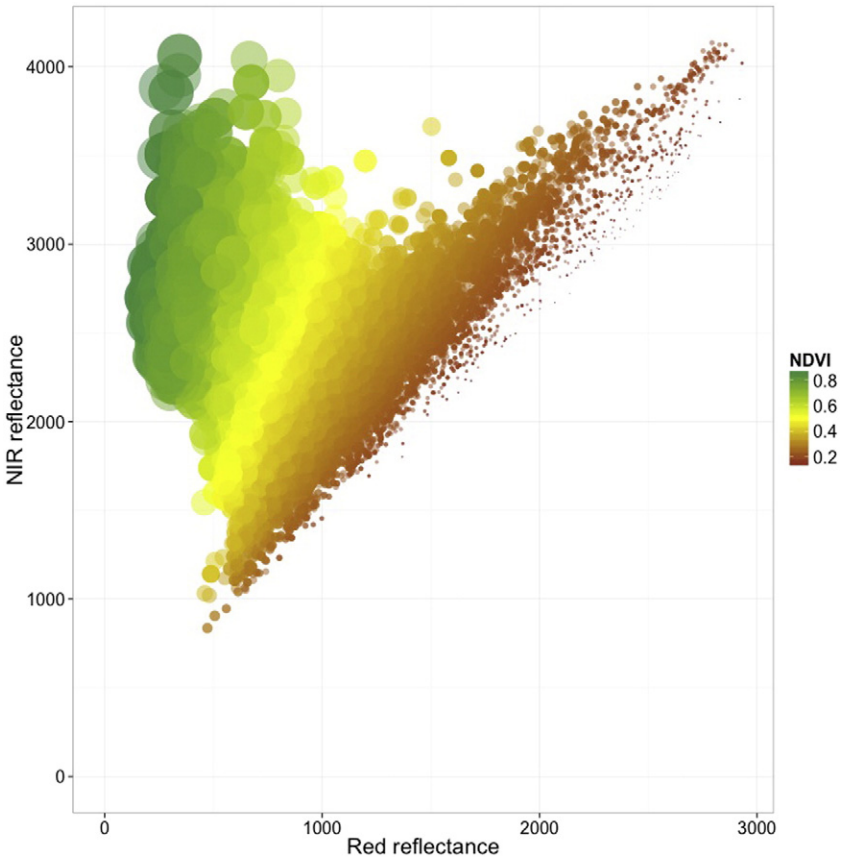

Fig. 3. Red vs NIR reflectance for the MODIS spectral library used in the study. Size of points and color reflect NDVI.

solution in water), base cations (Melich-3 extraction), and texture on 10% of the samples collected. Mid-infrared Spectroscopy (MIRS), which has become a well-established methodology for predicting important soil properties such as soil organic carbon (SOC), pH, base cations (SB), and texture (Madari et al., 2006; Reeves et al., 2006; Terhoeven-Urselmans et al., 2010; Vågen et al., 2006, 2013), was used to predict these properties on the remaining full set of soil samples. Soil MIRS is a non-destructive, rapid, and cost-effective methodology, enabling landscape-level assessments of soil health.

2.3. Remote sensing data processing

For this study we used Bidirectional Reflection Distribution Function (BRDF) corrected MODIS (MCD43A4) data for the calendar year 2012. This MODIS product is generally more stable and consistent since view angle effects have been removed from the directional reflectance values. To get consistent annual compo site reflectance values we opted to identify the date with the highest fractional vegetation cover based on the soil adjusted total vegetation (SATVI) index (Qi et al., 1994, 2002) and extracted the reflectance of each pixel for these dates in order to build a MODIS image spectral library for all of the sampled LDSF plots. The MODIS reflectance values (bands 1–7) were then used in the development of predictions models for soil properties and RDR50, respectively.

2.4. Prediction model development

We used Random Forest (RF) ensemble models (Breiman, 2001) for the prediction of soil properties and occurrence of RDR50. The main principle behind ensemble modeling techniques in is that a group of "weak learners" can be combined or bagged to form a "strong learner". A RF model uses a "decision tree" approach, which is a common machine learning technique, but rather than building one decision tree it builds several trees by creating subsets of the input data, each subset representing about two thirds of the total dataset. The individual decision trees were grown according to the CART algorithm (Breiman et al., 1984) and soil property values were predicted by averaging individual regression model predictions. In the case of RDR50, the dependent variable is binary (Yes or No) and predictions were made by counting the Yes votes in the ensemble and simply calculating the percent Yes votes received, which is the predicted probability of having RDR50 in each image pixel.

The application of RF models in ecology (Lawrence et al., 2006; Zhu, 2011) and soil science (Grimm et al., 2008; Kim et al., 2012; Vågen et al., 2013) is a relatively recent phenomenon, but has the potential to be a powerful approach for digital soil mapping, particularly when modeled

Table 1
Ranges in soil property values and summary of accuracy statistics (R^2 and RMSEP) for the RF model prediction of soil properties when applied to the three test datasets (averaged by sampling cluster) and compared to wet chemistry analysis results.

Variable	Soil property	Test			Wet chemistry
		Set 1	Set 2	Set 3	(N = 826)
Range	SOC (g kg^{-1})	3.5–114.0	3.6–117.1	3.3–139.6	1.3–161.6
	pH	4.3–8.6	4.4–8.7	4.3–8.8	3.4–10.5
	Sand (%)	3.1–83.4	5.7–83.2	3.9–84.5	1.0–98.7
	SB ($cmol_c$ kg^{-1})	3.4–99.7	3.2–111.6	2.9–91.7	0.7–179.2
R^2	SOC (g kg^{-1})	0.69	0.75	0.77	0.75–0.82
	pH	0.87	0.87	0.89	0.81–0.82
	Sand (%)	0.81	0.82	0.81	0.75–0.79
	SB ($cmol_c$ kg^{-1})	0.81	0.82	0.85	0.76–0.83
RMSEP	SOC (g kg^{-1})	11.46	10.67	9.60	9.30–11.05
	pH	0.35	0.36	0.32	0.47–0.48
	Sand (%)	9.24	8.99	9.17	12.81–13.67
	SB ($cmol_c$ kg^{-1})	6.76	6.74	6.12	9.00–10.58

relationships have high levels of dimensionality. Random Forest models can be applied both for regression and classification and are not prone to overfitting as the resampling does not use weighting (Gislason et al., 2006). The approach used here is similar to that applied in Vågen et al. (2013) for Landsat data.

The internal out-of-bag (OOB) prediction in the RF model, which is generated through bootstrapping, provides an estimate of model fit or accuracy across the decision trees in the model and was used for initial assessment of model performance. However, for the final prediction models, we used 2:1 cross-validation by randomly drawing two thirds of the data for calibration (N = 6982) and using the remaining one third of the samples (N = 3491) for validation (testing), without replacement. In other words, the three test datasets were independent of each other. This approach, while not perfect, gives reasonable

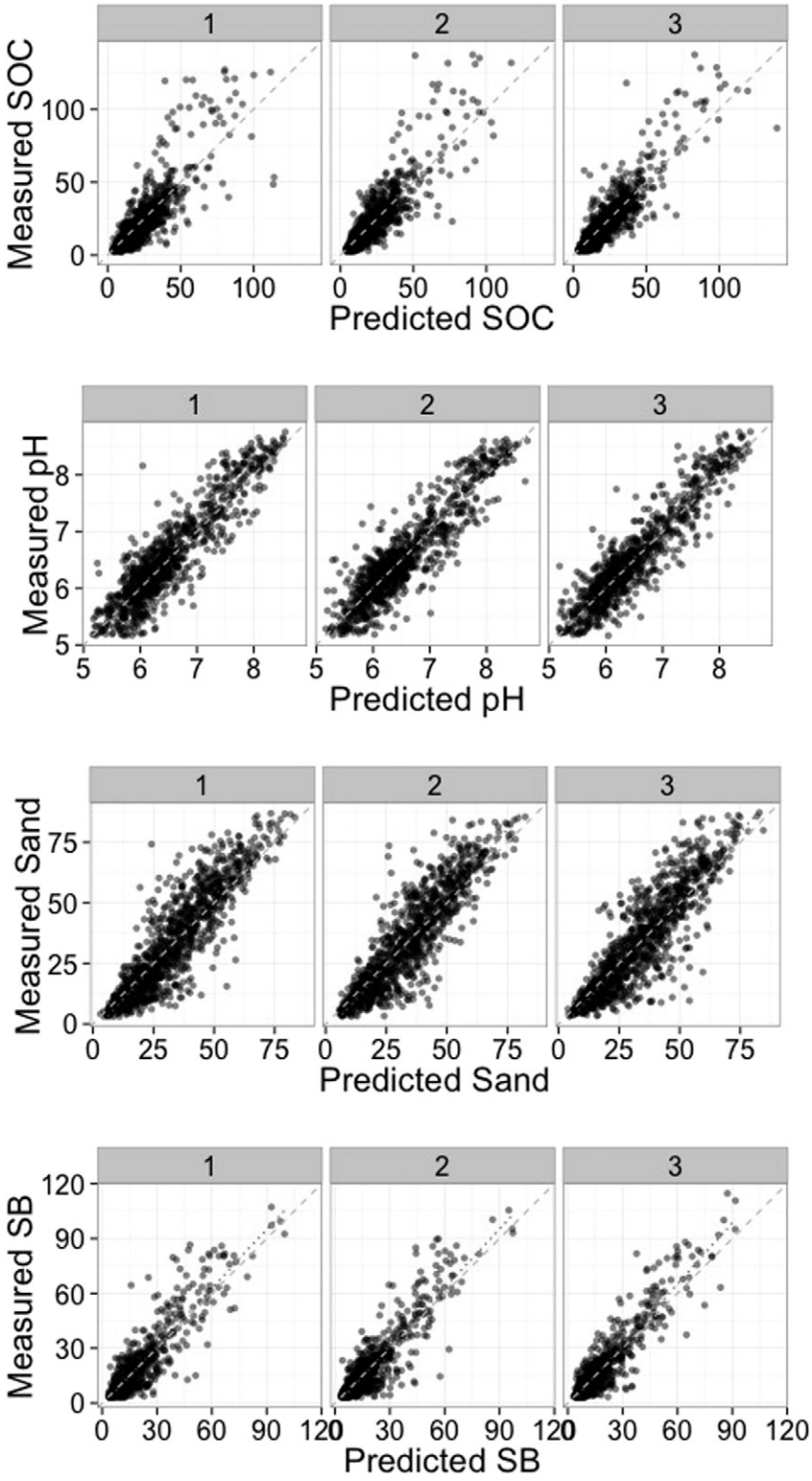

Fig. 4. Validation results for the predictions of SOC (g kg^{-1}), pH, sand (%), and exchangeable bases ($cmol_c$ kg^{-1}) based on MODIS. Prediction results are averaged to represent the mean for each MODIS pixel.

estimates of uncertainty in the developed models overall, including their stability across different sampling domains and leads to better model performance overall, while guarding against catastrophically poor model performance (Schaffer, 1993). The above RF regression models were applied to the MODIS reflectance image library for horizontal tiles 16 to 23 and vertical tiles 5 through 13 in order to create predictive maps of SOC, pH, sand, and SB for the African continent. Finally, the results from each calibration model were tested against wet chemistry analysis results for plots with complete reference analysis of SOC, pH, SB, and sand (N = 826). This additional validation step is helpful in testing for potential errors propagated through the MIR prediction of soil properties.

The probability of RDR50 in each pixel was mapped using the developed RF classification model (see also Vågen et al., 2013) based on field observations from 17,514 LDSF plots (from 114 sites). Accuracy statistics were calculated for each predicted outcome, using independent 2:1 cross-validation following the same procedure as that used for soil property predictions to assess model performance. We also assessed overall model performance by calculating precision (positive predicted value) and sensitivity (recall) for each validation run through Receiver Operating Characteristic (ROC) analysis. Overall model accuracy was assessed by calculating the area under the ROC (AUC) curve for each validation model run. The AUC is a measure of discrimination, or the ability of the model to correctly classify RDR50. We also calculated Cohen's kappa index (Cohen, 1960) for each validation run, which measures agreement corrected for chance on a scale from zero (no agreement) to one (total agreement).

3. Results and discussion

The individual sampling plots included in this study were representative of a wide range of conditions in terms of topsoil properties (Fig. 2), ranging in SOC concentrations from 1.5 to 300 g kg^{-1} (mean = 22.1 g kg^{-1}), pH from 3.7 to 9 (mean = 6.1), sand content from 2 to

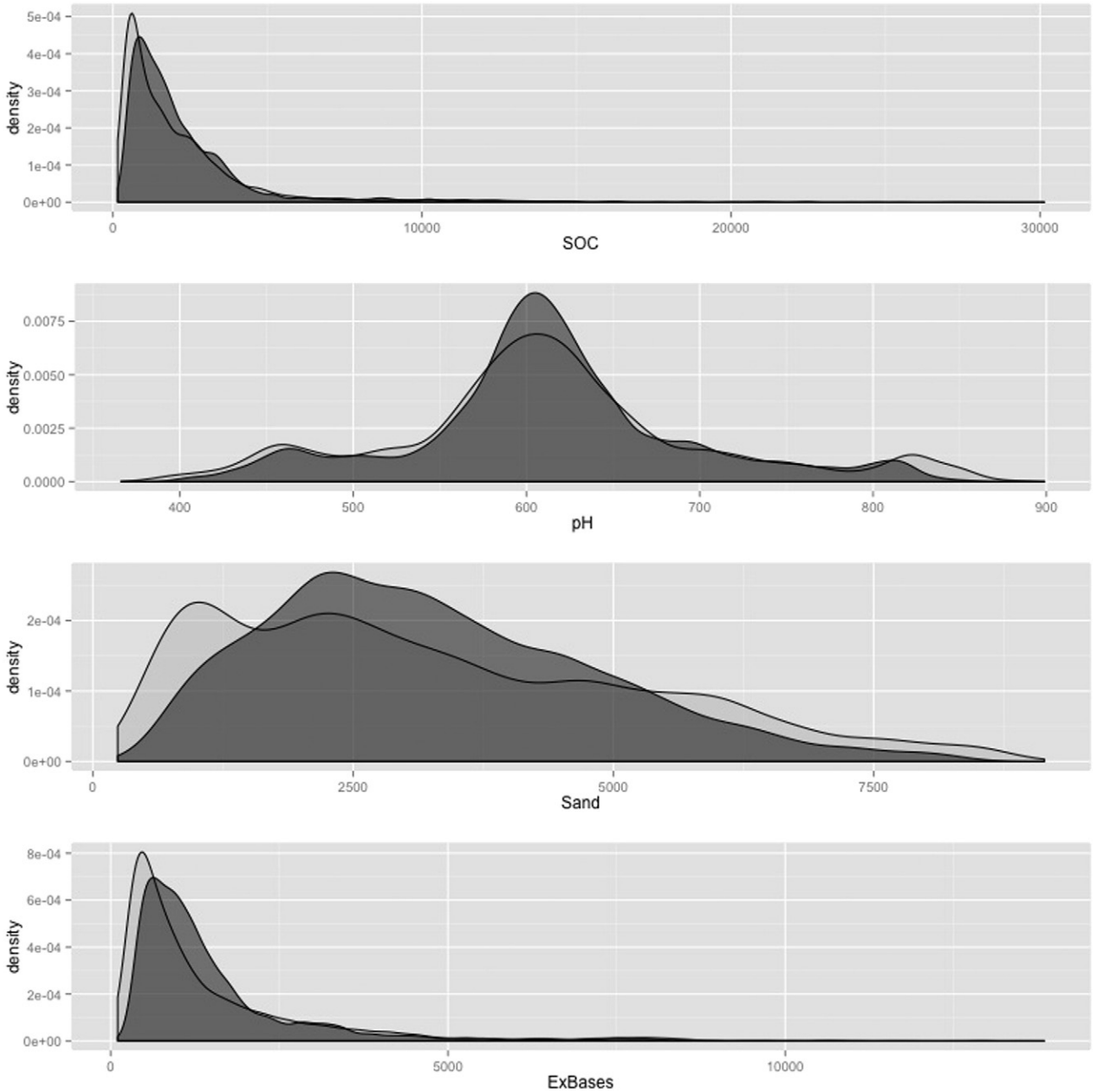

Fig. 5. Distributions of measured soil properties (light gray) and MODIS predicted soil properties (dark gray).

91% (mean = 33%), and SB from 1 to 138 $cmol_c\ kg^{-1}$ (mean = 15 $cmol_c\ kg^{-1}$). Mean annual precipitation (MAP) in the sampled sites ranged from about 276 to 2900 mm (mean = 1100 mm), based on Tropical Rainfall Monitoring Mission (TRMM) data for the period 1998 to 2012. The study also builds on sites from a diverse range of different ecosystems, from humid tropical forests, dry forests such as Miombo woodlands, natural savannah systems, semi-arid rangelands and croplands and arid ecosystems, as reflected in the range of MODIS spectral properties and NDVI values shown in Fig. 3.

Model performances for the prediction of soil properties based on MODIS are satisfactory overall (Table 1 and Fig. 4) and similar to those achieved in other studies using remote sensing for the mapping of SOC (Vågen et al., 2013; Wiesmeier et al., 2010). Model performance, when applied to the test datasets, for SOC shows a r^2 value of 0.74, with RMSEP between 9.6 and 11.46, which is also comparable to that achieved based on a globally distributed laboratory mid-infrared spectral library (Terhoeven-Urselmans et al., 2010). The results for pH are excellent with $r^2 = 0.88$ and RMSEP between 0.32 and 0.36 (Table 1). This is comparable to studies using proximal sensing in precision agriculture (Tekin et al., 2013). The RMSEP values for sand (8.99–9.24) and SB (6.12–6.76) are also similar to those reported by Terhoeven-Urselmans et al. (2010). Prediction performance was also good for all soil properties when the individual calibration model results were tested against laboratory measured soil properties (Table 1). We investigated prediction performance further by calculating the distributions of the measured and predicted values, respectively (Fig. 5). The results confirm the good correspondence between modeled and measured soil properties.

We proceeded with fitting the above RF prediction models to the MODIS image library described in Section 2.3, producing a set of soil property maps for the African continent for 2012 (Fig. 6). In brief, the maps show high SB and neutral to alkaline pH in arid to semi-arid ecosystems, including parts of the great Rift Valley of East Africa, while predictions for humid tropical regions are towards lower pH values. Predictions of SOC are highest in tropical forest systems, including montane forests in East Africa, the Congo basin, and the eastern rainforests of Madagascar, while predicted sand content is highest in dry lands (e.g. deserts and semi-deserts), as well as on

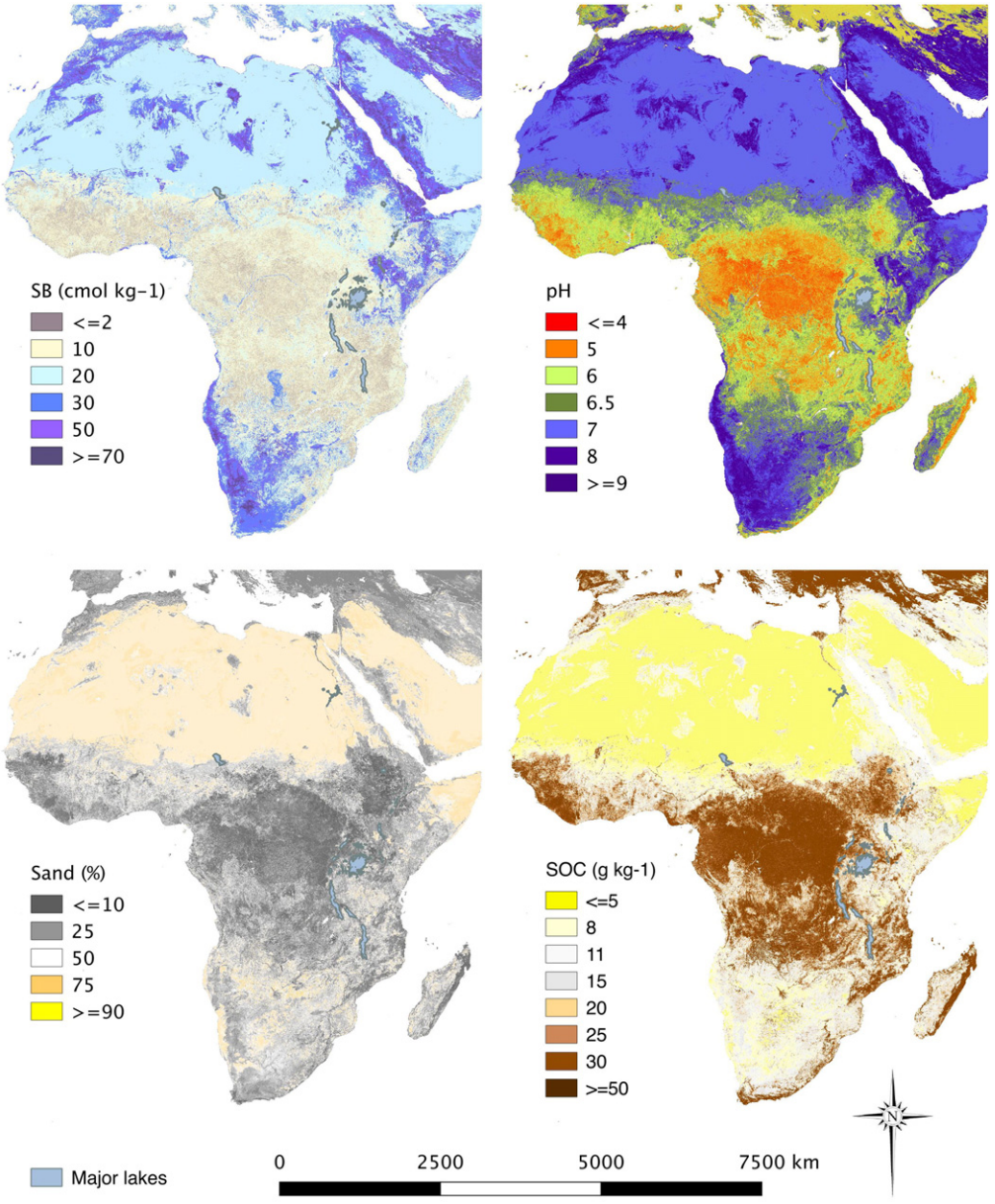

Fig. 6. Maps of sum of bases (SB), pH, SOC, and sand content using MODIS reflectance data for 2012.

granitic complexes such as in central Tanzania. The predicted SOC concentrations also reflect recent conversions following deforestation events.

In order to explore the relationships between predicted soil properties further, we randomly sampled the soil property maps using 5000 points (Fig. 7, top-left panel). The relationship between predicted sand and

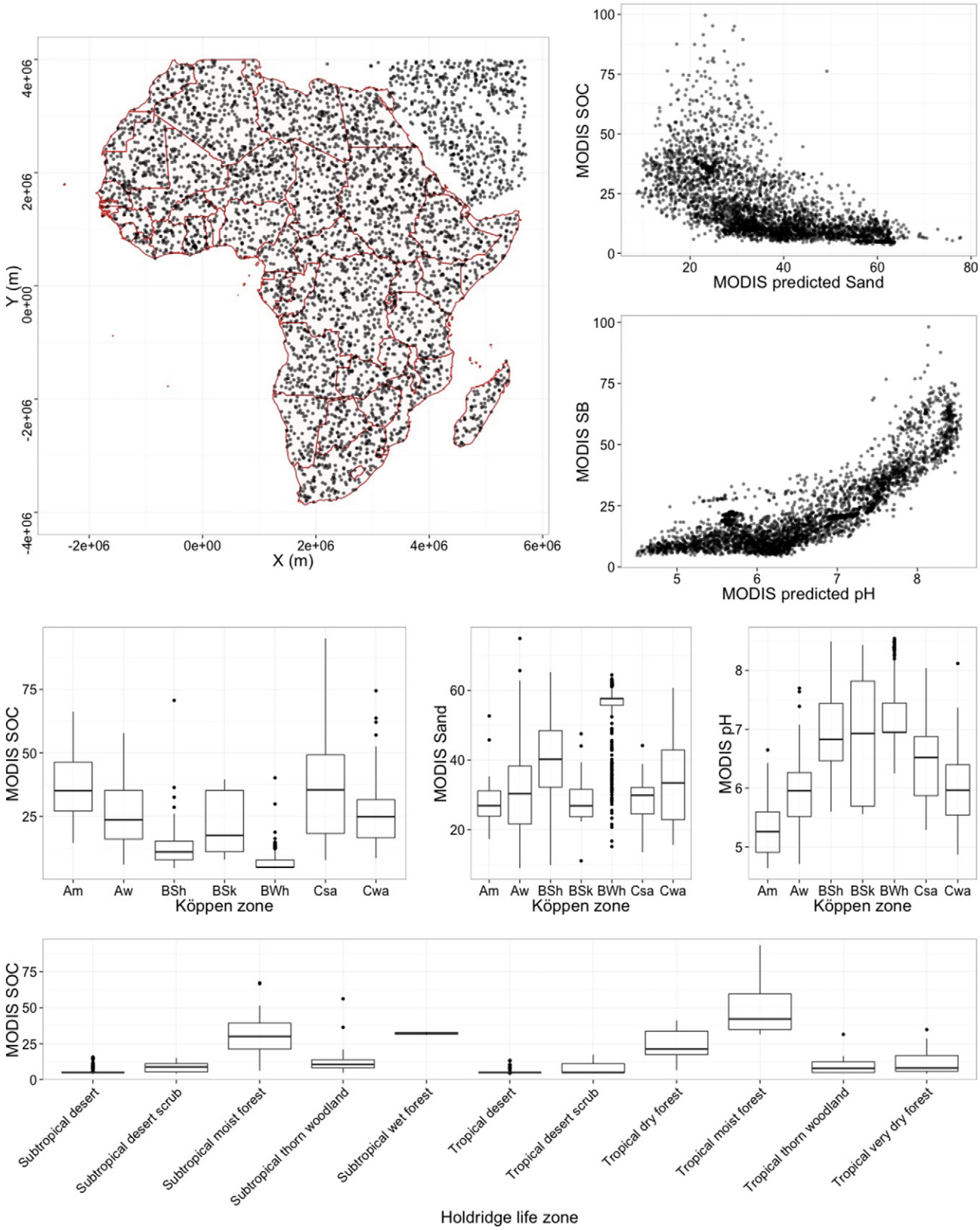

Fig. 7. Relationships between MODIS predicted soil properties based on a random sample of 5000 image pixels (points) from the predicted maps (top-left panel). The top-right panels show SOC (g kg^{-1}) vs sand (%) and pH vs SB ($cmol_c$ kg^{-1}), respectively. The lower panels show MODIS SOC, sand, and pH by Köppen climate zones and MODIS SOC for the most common Holdridge life zones (based on maps from Leemans, 1990) in Africa.

Table 2
Accuracy statistics for the prediction of root-depth restrictions (RDR50) based on the validation test sets.

	Test set 1	Test set 2	Test set 3
AUC	0.85	0.85	0.86
Precision (0/1)	0.75/0.76	0.79/0.74	0.76/0.74
Sensitivity (0/1)	0.83/0.67	0.81/0.72	0.81/0.69
Accuracy (overall)	0.77	0.77	0.75
Cohen's kappa	0.50	0.54	0.50

SOC (Fig. 7, top-right) in this study shows markedly lower SOC content as sand content increases beyond about 40 to 50%, and follows the same pattern as that found for laboratory measured SOC and sand. This relationship can be explored using the maps of SOC and sand in Fig. 6 to identify key constraints in terms of the potential for increasing SOC through management interventions. Generally, rates of increase or accumulation of SOC in soils following changes in management such as reduced cultivation intensity are inversely related to sand content in soils (Lugo et al., 1986). In other words, sand content forms an important constraint envelope in terms of SOC sequestration potential.

Given that the predicted soil properties are based on remote sensing data alone and do not incorporate climate information in the models explicitly, we explored the predictions for the most dominant Köppen–Geiger climate zones and Holdridge life zones on the continent (Fig. 7). The middle-left panel in Fig. 7 shows the distribution of predicted SOC from the 5000 plots sampled earlier using MODIS relative to the dominant Köppen–Geiger climate zones on the continent, while the middle-center panel shows sand by Köppen–Geiger climate zone. The predicted concentrations of SOC are highest in equatorial (Am, Aw) and warm temperate (Csa, Cwa) climates, lower in arid steppe (BSh) and lowest in arid desert climates (BWh). Arid desert climates also have the highest sand contents, while cold arid climates (BSk) have similar ranges of predicted sand as equatorial and warm temperate climates and intermediate SOC contents (Fig. 7).

As shown in the lower panel in Fig. 7, subtropical and tropical dry and moist forests have the highest predicted SOC based on Holdridge life zones (Leemans, 1990; Prentice et al., 1992). These results are consistent with a number of other studies showing increasing SOC with increasing precipitation and decreasing temperature for different levels of precipitation (Post et al., 1982), as well as studies of SOC in various biomes globally (Jobbágy and Jackson, 2000).

The predicted maps for SB and pH show higher SB in areas with alkaline soils (i.e. pH > 7), which is a well-established relationship (Fig. 7, middle-right) and in drier climate zones (Fig. 7). The maps of pH and SB in Fig. 6 can be used to identify areas with soil constraints such as acidity (i.e. pH < 5.5) or alkalinity (i.e. pH > = 8.0), as well as soil fertility constraints (i.e. areas with SB < 8). Predicted pH values are highly variable within Köppen–Geiger climate zones, but lowest in equatorial monsoonal climates (Am), as expected given the more highly weathered nature of soils in these climate zones.

Overall accuracy for the prediction of RDR50 was about 77% (kappa = 0.52), as shown in the summary of the modeled results in Table 2. Importantly, the RF model produced consistent results across the independent test data sets, indicating that it is stable across a wide range of conditions. Also, measured and predicted frequencies of RDR50 occurrence at the site level are consistent.

The resulting maps of predicted occurrence of RDR50 show root-depth restriction hotspots in for example southern Africa, as well as eastern Kenya (e.g. parts of the Tana River basin) (Fig. 8). Predictions show low prevalence of RDR50 in parts of the Sahel and in the Kalahari sands and wetlands of western Zambia, as well as for example the Okavango delta (Fig. 8). From a management perspective, identifying areas that have high prevalence of RDR50 will be important in order to target interventions that reduce soil compaction in agricultural areas more

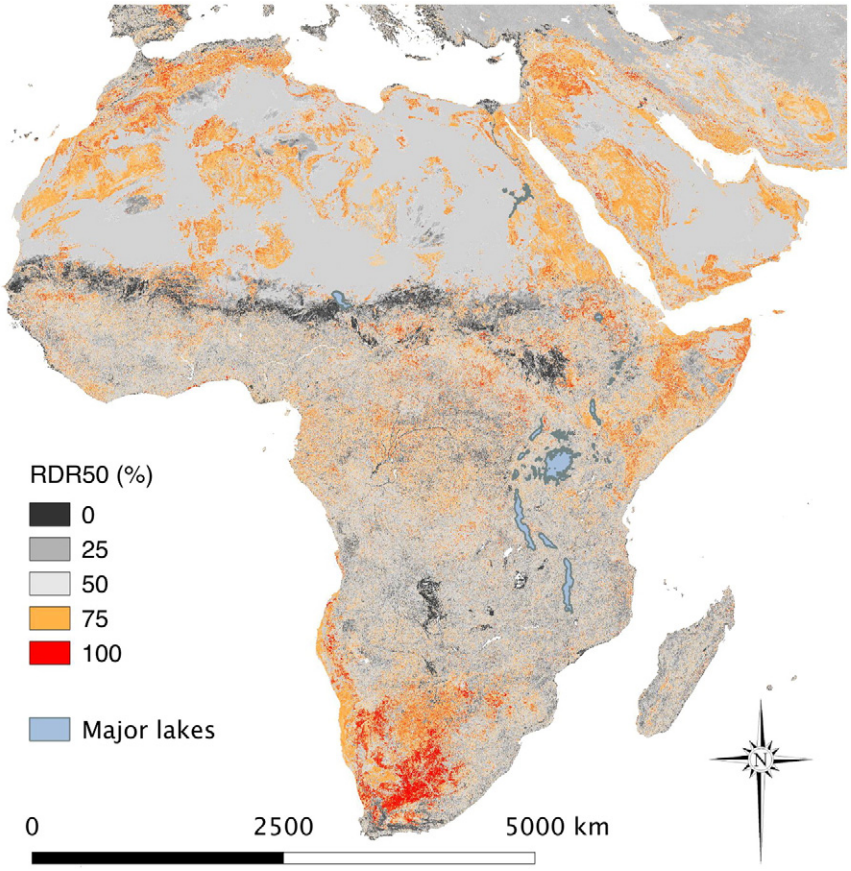

Fig. 8. Predicted maps of root-depth restrictions (RDR50) using MODIS reflectance data for 2012.

effectively, but also for identifying areas of high risk of developing such restrictions.

4. Conclusions

Combining laboratory IR spectroscopic soil analysis, MODIS remote sensing data and statistical methods, soil properties such as SOC, pH, sand content, and SB, as well as root-depth restrictions were mapped with unprecedented accuracy and spatial resolution for the African continent. The results indicate a strong relationship between predicted and measured values, with distributions of predicted soil property values similar to measured values. Also, prediction results for the mapping of restrictions to root-growth were satisfactory. The resulting maps can be used to identify areas with important soil fertility constraints, acidity or alkalinity, and for determining constraint envelopes for SOC sequestration by combining spatial predictions of SOC and sand. By combining predicted surfaces of soil properties, including cut-offs for various soil constraints, and the prevalence of root-depth restrictions, spatial patterns of soil degradation risk can be determined and management options better targeted both contextually and spatially.

The spatial resolution of the maps produced in this study have utility for assessments and monitoring of soil health at continental, regional, national and sub-national scales, and as prior estimates at more local scale. However, for applications in local assessments requiring finer resolution maps, approaches such as those shown in Vågen et al. (2013) or using high-resolution remote sensing data will be needed. A key challenge in future work will be to develop approaches for the prediction of soil health and land degradation risk that provide consistent estimates across spatial scales.

Acknowledgments

This work was funded with partial support from the Bill and Melinda Gates Foundation (BMGF) Grant Number 51353, the CGIAR Research Program on Forests Trees and Agroforestry, the CGIAR Climate Change, Agriculture and Food Security Research Program, and the CGIAR Challenge Program for Water and Food and Wajibu MS. The authors acknowledge the invaluable assistance of CIAT and ICRAF field and laboratory staff in Nairobi (Kenya), Lilongwe (Malawi), Arusha (Tanzania), and Bamako (Mali). We are also grateful to Professor Assefa Abegaz and Tesema Bekele Silewondim of Addis Ababa University, and Zelalem Hadush and Biadglign Demissie Mulawu of Mekelle University for coordinating the field surveys in Ethiopia.

References

Alcock, J., 2003. Positive feedback and system resilience from graphical and finite-difference models: the amazon ecosystem—an example. Earth Interact. 7, 1–23. http://dx.doi.org/10.1175/1087-3562(2003)007<0001:PFASRF>2.0.CO;2.
Bassett, T., Zuéli, K., 2000. Environmental discourses and the Ivorian savanna. Ann. Assoc. Am. 90, 67–95.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 35. http://dx.doi.org/10.1023/A:1010933404324.
Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and regression trees. Wadsworth Stat. Probab. Ser. 19 p. 368.
Cao, M., Zhang, Q., Shugart, H.H., 2001. Dynamic responses of African ecosystem carbon cycling to climate change. Clim. Res. 17, 183–193.
Chan, K.M.A., Shaw, M.R., Cameron, D.R., Underwood, E.C., Daily, G.C., 2006. Conservation planning for ecosystem services. PLoS Biol. 4, 15.
Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 37–46. http://dx.doi.org/10.1177/001316446002000104.
Curriero, F.C., 2007. On the use of non-euclidean distance measures in geostatistics. Math. Geol. 38, 907–926. http://dx.doi.org/10.1007/s11004-006-9055-7.
Drechsel, P., Gyiele, L., Kunze, D., Cofie, O., 2001. Population density, soil nutrient depletion, and economic growth in sub-Saharan Africa. Ecol. Econ. 38, 251–258.
Duraiappah, A., 1998. Poverty and environmental degradation: a review and analysis of the nexus. World Dev. 26, 2169–2179.
Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R., 2006. Random forests for land cover classification. Pattern Recogn. Lett. 27, 294–300. http://dx.doi.org/10.1016/j.patrec.2005.08.011.
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, 102–113. http://dx.doi.org/10.1016/j.geoderma.2008.05.008.
Grunwald, S., Thompson, J.a., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. Soil Sci. Soc. Am. J. 75, 1201. http://dx.doi.org/10.2136/sssaj2011.0025.
Gunderson, L.H., 2000. Ecological resilience — in theory and application. Annu. Rev. Ecol. Syst. 31, 425–439. http://dx.doi.org/10.1146/annurev.ecolsys.31.1.425.
Hengl, T., Heuvelink, G.B.M.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120, 75–93. http://dx.doi.org/10.1016/j.geoderma.2003.08.018.
Hengl, T., de Jesus, J., MacMillan, R., 2014. SoilGrids1km — Global Soil Information Based on Automated Mapping. PLoS One 9. http://dx.doi.org/10.1371/journal.pone.0114788.
Heuvelink, G., Webster, R., 2001. Modelling soil variation: past, present and future. Geoderma 100, 269-3.
Jobbágy, E., Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecol. Appl. 10, 423–436.
Kim, J., Grunwald, S., Rivero, R.G., Robbins, R., 2012. Multi-scale modeling of soil series using remote sensing in a wetland ecosystem. Soil Sci. Soc. Am. J. 76, 2327. http://dx.doi.org/10.2136/sssaj2012.0043.
Kirkegaard, J., Lilley, J., 2007. Root penetration rate–a benchmark to identify soil and plant limitations to rooting depth in wheat. Aust. J. Exp. Agric. 47, 590–602.
Köppen, W., 1918. Klassifikation der Klimate nach Temperatur. Niederschlag und Jahreslauf.
Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen–Geiger climate classification updated. Meteorol. Z. 15, 259–263. http://dx.doi.org/10.1127/0941-2948/2006/0130.
Lawrence, R.L., Wood, S.D., Sheley, R.L., 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). Remote Sens. Environ. 100, 356–362. http://dx.doi.org/10.1016/j.rse.2005.10.014.
Leemans, R., 1990. Global Data Sets Collected and Compiled by the Biosphere Project. Laxenburg, Austria.
Lugo, A.E., Sanchez, M.J., Brown, S., 1986. Land use and organic carbon content of some subtropical soils. Plant Soil 96, 185–196. http://dx.doi.org/10.1007/BF02374763.
Lull, H., 1959. Soil Compaction on Forest and Range Lands.
Madari, B.E., Reeves, J.B., Machado, P.L.O.A., Guimarães, C.M., Torres, E., Mccarty, G.W., 2006. Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. Geoderma 136, 245–259. http://dx.doi.org/10.1016/j.geoderma.2006.03.026.
Mulder, V.L., Bruin, S. De, Schaepman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping — a review. Geoderma 162, 1–19. http://dx.doi.org/10.1016/j.geoderma.2010.12.018.
Post, W.M., Emanuel, W.R., Zinke, P.J., Stangenberger, A.G., 1982. Soil carbon pools and world life zones. Nature 298, 156–159. http://dx.doi.org/10.1038/298156a0.
Prentice, I., Cramer, W., Harrison, S., 1992. Special paper: a global biome model based on plant physiology and dominance, soil properties and climate. J. Biogeogr. 19, 117–134.
Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126. http://dx.doi.org/10.1016/0034-4257(94)90134-1.
Qi, J., Marsett, R., Heilman, P., Bieden-bender, S., Moran, S., Goodrich, D., Weltz, M., 2002. RANGES improves satellite-based information and land cover assessments in southwest United States. EOS Trans. Am. Geophys. Union http://dx.doi.org/10.1029/2002EO000411.
Reeves III, J., Follett, R., McCarty, G., Kimble, J., 2006. Can near or mid-infrared diffuse reflectance spectroscopy be used to determine soil carbon pools? Commun. Soil Sci. Plant Anal. 37, 2307–2325. http://dx.doi.org/10.1080/00103620600819461.
Rubel, F., Kottek, M., 2010. Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen–Geiger climate classification. Meteorol. Z. 19, 135–141. http://dx.doi.org/10.1127/0941-2948/2010/0430.
Schaffer, C., 1993. Selecting a classification method by cross-validation. Mach. Learn. 13, 135–143. http://dx.doi.org/10.1007/BF00993106.
Scull, P., Franklin, J., Chadwick, O., McArthur, D., 2003. Predictive soil mapping: a review. Prog. Phys. Geogr. 27, 171–197.
Tekin, Y., Kuang, B., Mouazen, A.M., 2013. Potential of on-line visible and near infrared spectroscopy for measurement of pH for deriving variable rate lime recommendations. Sensors (Basel). 13, 10177–10190. http://dx.doi.org/10.3390/s130810177.
Terhoeven-Urselmans, T., Vagen, T.-G., Spaargaren, O., Shepherd, K.D., 2010. Prediction of soil fertility properties from a globally distributed soil mid-infrared spectral library. Soil Sci. Soc. Am. J. 74, 1–8. http://dx.doi.org/10.2136/sssaj2009.0218.
Unger, P., Kaspar, T., 1994. Soil compaction and root growth: a review. Agron. J. 86, 759–766.
Vågen, T.-G., Shepherd, K.D., Walsh, M.G., Va, T., 2006. Sensing landscape level change in soil fertility following deforestation and conversion in the highlands of Madagascar using Vis-NIR spectroscopy. Geoderma 133, 281–294. http://dx.doi.org/10.1016/j.geoderma.2005.07.014.
Vågen, T.-G., Shepherd, K.D., Walsh, M.G., Winowiecki, L.A., Tamene Desta, L., Tondoh, J.E., 2010. AfSIS Technical Specifications — Soil Health Surveillance, Africa. CIAT (the AfSIS project), Nairobi, Kenya.
Vågen, T.-G., Winowiecki, L.A., Abegaz, A., Hadgu, K.M., 2013. Landsat-based approaches for mapping of land degradation prevalence and soil functional properties in Ethiopia. Remote Sens. Environ. 134, 266–275.
Warren, A., 2002. Land degradation is contextual. Land Degrad. Dev. 13, 449–459.
Wiesmeier, M., Barthold, F., Blank, B., Kögel-Knabner, I., 2010. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant Soil 340, 7–24. http://dx.doi.org/10.1007/s11104-010-0425-z.
Zhu, X., 2011. Wetland mapping using remote sensing imagery and modelmap. Proceedings of the Surveying & Spatial Sciences Biennial Conference 2011. Scion, Wellington, New Zealand, pp. 353–364.