# Dynamics, drivers and impacts of land cover changes in the Lake Nakuru Drainage Basin and Eastern Mau Forest Reserve, Kenya

Forløp, årsaksforhold og virkninger av endringer i fordelingen av arealtyper i nedbørfeltet til Lake Nakuru og i Eastern Mau Forest Reserve, Kenya

Philosophiae Doctor (PhD) Thesis

**Kennedy Okello Were**

Department of Mathematical Sciences and Technology
Faculty of Environmental Science and Technology
Norwegian University of Life Sciences

Ås 2014

Thesis number 2014:53
ISSN 1894-6402
ISBN 978-82-575-1214-9

A doctoral thesis at the Norwegian University of Life Sciences consists of either a monograph or a compendium of papers in published or manuscript form. In the latter case, the introductory part, from an overall perspective, summarises and collates the research problems, results, and conclusions presented in the separate papers, and documents the coherence of the thesis.


**Evaluation committee:**
Prof. Rattan Lal
Distinguished university professor of soil science
School of Environment & Natural Resources
The Ohio State University, USA

Dr. Arnt-Børre Salberg
Senior research scientist
Department of Statistical Analysis, Pattern Recognition & Image Analysis
Norwegian Computing Center, Norway

Dr. Ingunn Burud (committee coordinator)
Department of Mathematical Sciences & Technology
Faculty of Environmental Science & Technology
Norwegian University of Life Sciences, Norway

**Supervisors:**
Prof. Øystein Bjarne Dick (main supervisor)
Geomatics section, Department of Mathematical Sciences & Technology
Faculty of Environmental Science & Technology
Norwegian University of Life Sciences, Norway

Prof. Bal Ram Singh (co-supervisor)
Department of Environmental Sciences
Faculty of Environmental Science & Technology
Norwegian University of Life Sciences, Norway

## ACKNOWLEDGEMENTS

"*When eating bamboo sprouts, remember the man who planted them.*" In the spirit of this saying, I wish to express my gratitude for the contributions I received from various sources towards my advancement and, which made the completion of this PhD thesis possible. First and foremost, special thanks go to the Almighty God for the gift of life and good health.

Profuse thanks are also due to the Research Council of Norway for funding the PhD studies through the Norwegian University of Life Sciences, and the Kenya Agricultural Research Institute for granting me leave to pursue the studies.

My supervisors, Prof. Øystein Bjarne Dick and Prof. Bal Ram Singh, deserve great credit for offering invaluable insights, guidance, encouragement, support, and selfless time all through. Together with Vegard Ophaug, a PhD colleague at the department, they also translated the English abstract of this thesis into Norwegian; *tusen takk for det*.

I highly appreciate the departmental staff, particularly Frode Hilmarsen, Anita Haugen Habbestad, Mona V. Kristiansen, and Tone Rasmussen, for providing vital information technology (IT), financial, and administrative services.

Additional thanks go to the PhD and research colleagues at the department, and all friends in Ås, for encouraging my progress and energy, and for their social support and warm companionship.

I am also grateful to Peter Owenga, Evanson Thairu, and Stephen Njuguna for their dedication, diligence, and team spirit during fieldwork.

Lastly, I unreservedly thank my parents and siblings for always being there, and for their steadfast love, affection, prayers, and support throughout my (academic) life; *asanteni sana*. I dedicate this thesis to them.

"*To speak gratitude is courteous and pleasant, to enact gratitude is generous and noble, and to live gratitude is to touch Heaven.*" (Johannes A. Gaertner)

Ås, May 20, 2014 Kennedy Okello Were

TABLE OF CONTENTS

**Acknowledgements** ........ i
**Summary** ........ v
**Sammendrag** ........ vii
**List of papers** ........ ix
**List of abbreviations** ........ xi
**1.0 Background and status of knowledge** ........ 1
1.1 Land cover and land cover changes ........ 1
1.1.1 Impacts of land cover changes ........ 3
1.1.1.1 Impacts of land cover changes on soils ........ 3
1.1.2 Remote sensing of land cover and land cover changes ........ 4
1.1.3 Modelling of land cover changes and soil landscapes ........ 6
1.2 Land cover changes in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve ........ 7
**2.0 Research problem and rationale** ........ 9
2.1 Research goal and objectives ........ 11
2.2 Research questions ........ 11
**3.0 Materials and methods** ........ 11
3.1 Study area ........ 11
3.2 Data ........ 13
3.2.1 Field data ........ 13
3.2.2 Remotely-sensed data ........ 15
3.2.3 Cartographical and GIS data ........ 15
3.3 Methods ........ 15
3.3.1 Spatio-temporal mapping and analyses of land cover changes ........ 16
3.3.2 Analyses of the linkages between land cover change processes and the geophysical and socio-economic determinants ........ 16
3.3.3 Assessment of the effects of land cover changes on SOC and TN stocks ........ 17
3.3.4 Spatially-distributed modelling and mapping of SOC and TN stocks ........ 18
**4.0 Results and discussion** ........ 18
4.1 The spatial and temporal dynamics of land cover changes ........ 18
4.2 The geophysical and socio-economic determinants of land cover changes ........ 21
4.3 The effects of land cover changes on SOC and TN stocks ........ 23
4.4 The spatially-distributed estimates and patterns of SOC and TN stocks ........ 26
4.5 Limitations of the research ........ 31
**5.0 Conclusions, recommendations and outlook** ........ 32
**References** ........ 35
**Errata** ........ 43
**Paper I - IV** (Individual page numbering)

## SUMMARY

Land cover in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve has been changing at different spatial and temporal scales. The changes, especially forest conversion, have interfered with the provision of ecosystem services, e.g., carbon sequestration. To mitigate the adverse impacts of the land cover changes and determine the future landscape scenario for environmental sustainability, reliable, consistent, multi-temporal, and spatially-explicit information on the state of land cover and other biophysical variables (e.g., soil properties, biodiversity) is necessary. Unfortunately, the archives of such information for this specific area, and Kenya in general, are poor.

This research aimed to fill the existing biophysical information gap and, thereby, contribute knowledge to support policy formulation for sustainable land management. The specific objectives were to: (i) detect, quantify, and map the land cover changes that had occurred over time, (ii) analyse the linkages between the land cover change processes and the geophysical and socio-economic factors that determine them, (iii) assess the effects of forest-cropland conversion on soil organic carbon (SOC) and total nitrogen (TN) stocks in the Eastern Mau Forest Reserve, and (iv) model and map the spatial distribution of SOC and TN stocks in the Eastern Mau Forest Reserve.

The magnitudes, rates, nature, and spatial patterns of land cover changes from 1973 to 2011 were derived through integration of field, satellite remote sensing, and GIS methods. Results revealed the transformation from natural to human-dominated landscape that occurred within the 38-year period. Forests-shrublands dominated the landscape from 1973 to 2000, but by 2011, croplands had become dominant. The important land cover change processes were conversion of native systems (forests-shrublands and grasslands) and expansion of croplands and built-up lands. Forest-shrublands, grasslands, and croplands had higher magnitudes of change than the other land cover types. The *hotspots* of forest-shrubland conversion were spread in the mid-regions and northern side of Lake Nakuru between 1973 and 1985, on the western side between 1985 and 2000, and around the Lake Nakuru National Park and on the western and southern parts between 2000 and 2011. Built-up lands were the most dynamic given their high annual average rates of change; for example, between 1985 and 2000, their annual average rate of change was 17%.

Moreover, the linkages between the geophysical and socio-economic determinants of the important land cover change processes were explored using binary logistic regression models and auxiliary data in a spatially-explicit framework. Results indicated that the significance, magnitude, and direction of the determining factors varied with time, as well as the nature of land cover change process. For example, between 1985 and 2000, rainfall, soil pH, soil cation exchange capacity (CEC), topographic wetness index (TWI), aspect, curvature, distance to road, and distance to town partly explained the occurrence of forest-shrubland conversion. But between 2000 and 2011, the foregoing factors, in addition to slope and distance to river, and with the exception of TWI, were the significant determinants of the observed forest-cropland conversions.

To establish the response of soils to the changing landscape in the Eastern Mau Forest Reserve, variations of SOC and TN stocks under natural forests (NF), plantation forests (PF), bamboo forests (BF), and croplands established after forest conversion (i.e., NF2C, PF2C, and BF2C) were assessed using a combination of field, laboratory, spatial, and linear mixed methods. Results showed that converting forests to croplands had reduced SOC and TN concentrations and stocks. For example, both SOC and TN stocks decreased significantly by about 51% in the surface and about 42% in the subsurface soils after conversion of NF. In the surface soils, the highest SOC and TN concentrations were in NF and the lowest in NF2C, while in the subsurface soils, the highest concentrations were in NF and the lowest in PF2C. The SOC and TN concentrations and stocks also decreased significantly as the soil depth increased.

Furthermore, the spatially-distributed patterns of SOC and TN stocks were modelled and mapped using field data, auxiliary spatial data, and four spatial statistical methods; namely, geographically weighted regression, geographically weighted regression-kriging, multiple linear regression, and multiple linear regression-kriging. Elevation, silt content, TN concentration, and Landsat 8 Operational Land Imager band 11 (proxy for land surface temperature) explained 72% of the variability in SOC stocks, while the same factors (excluding silt content) explained 71% of the variability in TN stocks. Soil properties, particularly TN and SOC concentrations, were more important than the other environmental factors in controlling the dynamics of SOC and TN stocks, respectively. The highest estimates of SOC and TN stocks (*hotspots*) were on the western and north-western parts where forests dominated, while the lowest estimates (*coldspots*) were on the eastern side where croplands had been established. Forests stored the highest amounts of SOC and TN (3.78 Tg C and 0.38 Tg N) followed by croplands (2.46 Tg C and 0.25 Tg N), and grasslands (0.57 Tg C and 0.06 Tg N). Overall, the Eastern Mau Forest Reserve stored about 6.81 Tg C and 0.69 Tg N.

The findings and outputs of this research enhance our understanding of human actions and their consequences in the study area. They provide a good base of spatially-explicit biophysical information to monitor land resources and formulate spatially-targeted policies for sustainable land management. The elaborate field sampling, satellite remote sensing, GIS, and (spatial) statistical approaches applied in the research can also be replicated in other data-poor environments in Eastern Africa to cost-effectively derive multi-purpose biophysical information. In a broader context, the resultant land cover and soil databases can support the activities of other programs, such as REDD+, FAO land use-land cover, and ISRIC-World Soil Information programs, to mention but a few.

In conclusion, spatially-targeted and time-specific policies that will restore and conserve the natural ecosystems, as well as enhance agricultural productivity for environmental sustainability and socio-economic well-being are recommended. Additionally, adoption of best management practices (BMPs), especially agro-forestry practices where fast-growing, highly productive, deep-rooted, and N-fixing trees are planted, will be beneficial for mitigating C and N losses in the croplands. For the forest soils, long-term storage of C and N will require proper management and protection of the forests from further deforestation and degradation.

**Keywords:** Land cover change • land cover change modelling • soil organic carbon • total nitrogen • soil landscape modelling • remote sensing • GIS • Eastern Mau • Lake Nakuru

## SAMMENDRAG

Fordelingen av arealtyper (land cover) i nedbørfeltet til Lake Nakuru og i skogsreservatet Eastern Mau Forest Reserve har vært i stadig endring, i varierende omfang, tidsmessig så vel som romlig. Endringene, spesielt overgang fra en type skog til annen, har virket inn på tilgjengeligheten av ulike økosystemtjenester, som for eksempel karbonfangst. For å dempe den uheldige virkningen av endringer i arealtype og samtidig legge et grunnlag for en fremtidig miljømessig bærekraftig landskapsutvikling, vil det være avgjørende viktig å ha tilgang til pålitelig, konsistent, multitemporal og stedfestet informasjon om arealtypenes tilstand og tilknyttede biofysiske variable, som for eksempel jordegenskaper og biologisk mangfold. Uheldigvis er slik informasjon svært mangelfull, i særdeleshet hva dette området angår, og i Kenya i sin alminnelighet.

Formålet med dette forskningsarbeidet har vært å bøte på den eksisterende mangel på biofysisk relatert informasjon, for dermed å kunne bidra med kunnskap til støtte for politiske beslutninger knyttet til bærekraftig arealforvaltning. De mer detaljerte målsettingene var: (i) kartlegging, overvåking og tallfesting av arealtype-endringer som hadde funnet sted over tid; (ii) analyse av koblingene mellom arealtype-endringsprosessene og de geofysiske og sosioøkonomiske faktorene som bestemte disse; (iii) vurdering av hvilken virkning overgang fra skog til dyrket mark har hatt på organisk karbon i jord - (SOC) og totalnitrogenmengden (TN) i skogsreservatet Eastern Mau Forest Reserve; og (iv): modellering av den romlige fordeling og kartlegging av organisk karbon i jord - og totalnitrogenmengden.

Omfang, art, grad av raskhet og romlig fordelingsmønster til endringene i arealtype fra 1973 til 2011 ble funnet ved å benytte en kombinasjon av GIS relaterte metoder, basert på feltarbeid, satellittfjernmåling og andre typer stedfestede data. Resultatet viste en tydelig endring fra et naturlig til et landskap dominert av menneskelig aktivitet i løpet av det 38 år lange tidsrommet. Mer presist kan det sies at skog og krattskog dominerte landskapet fra 1973 til 2000, mens det fra 2011 var dyrket mark som var den dominerende arealtype. De viktigste formene for endringer var reduksjon av de opprinnelige arealtypene (skog, krattskog og grasdekke) og utvidelse av dyrket mark og bebygde områder. Graden av endring var større for skog, krattskog, grasdekke og dyrket mark enn for de andre arealdekketypene. Områder der reduksjon av skog og krattskog i særlig grad fant sted (hotspots) var å finne spredt rundt i mellomregionene samt nord for Nakurusjøen mellom 1973 og 1985, på vestsiden mellom 1985 og 2000, og omkring Lake Nakuru nasjonalpark og i de vestlige og sørlige delene av området mellom 2000 og 2011. Med sin høye årlige grad av endring, fremsto arealtypen bebygde områder som mest dynamisk, eksempelvis var den årlige graden av endring 17% mellom 1985 og 2000.

I tillegg ble sammenhengen mellom naturgeografiske og sosioøkonomiske faktorer og de viktigste arealtypeendringsprosessene undersøkt med basis i et klart definert romlig rammeverk ved bruk av logistisk regresjon og tilleggsdata. Resultatene viste at betydningen, størrelsen og virkningsretningen til de bestemmende faktorene varierte både over tid og i forhold til arten av arealtypeendringsprosess. Som eksempel kan nevnes at nedbør, jords-pH, kationbyttekapasitet (CEC), topografisk fuktighetsindeks (TWI), eksposisjon, helningsform og avstand til vei bare delvis forklarte reduksjonen av skog og krattskog mellom 1985 og 2000. Mellom 2000 og 2011 har imidlertid de foregående faktorene med unntak av TWI, i tillegg til helningsgrad og avstand til elv vært de mest signifikante forklaringsvariablene for den observerte overgangen fra skog og krattskog til dyrket mark.

De jordrelaterte virkningene av arealtypeendringer, ble nærmere undersøkt. SOC og TN i arealkategoriene naturlig skog (NF), plantasjeskog (PF), bambusskog (BF) og de variantene av dyrket mark som var etablert etter overgang fra skog (dvs. NF2C, PF2C og BF2C) i Eastern Mau Forest Reserve ble bestemt. Felt- og laboratorie-arbeid i kombinasjon med romlige statistiske

metoder ble benyttet til dette. Resultatene viste at endringen av skog til dyrket mark hadde redusert mengden og konsentrasjonen av SOC og TN. For eksempel avtok mengden SOC og TN signifikant med 51% for hver av variablene på overflaten og med rundt 42% for hver av variablene i undergrunnslaget etter endring av NF. På overflaten var de høyeste konsentrasjonene av SOC og TN i NF, og de laveste i NF2C, mens de i undergrunnslaget var høyest i NF og lavest i PF2C. Mengden og konsentrasjonen av SOC og TN avtok signifikant med økt jorddybde.

Videre ble det romlige fordelingsmønsteret til SOC og TN modellert og kartlagt ved bruk av feltdata, romlige tilleggsdata og følgende fire romlige statistiske metoder; geografisk vektet regresjon, geografisk vektet regresjonskriging, multippel lineær regresjon og multippel lineær regresjonskriging. Terrenghøyde, siltinnhold, totalnitrogenkonsentrasjon og Landsat 8 Operational Land Imager band 11 (båndet for avledning av overflatetemperatur) forklarte 72% av variasjonen i mengden av lagret SOC, mens de samme faktorer med unntak av siltinnhold forklarte 71% av variasjonen i mengden av lagret TN. Konsentrasjonen av SOC og TN var viktigere for å forklare den romlige variasjonen av henholdsvis lagret SOC og lagret TN, enn de andre miljøfaktorene. De høyeste estimatene for lagret SOC og TN var i de vestlige og nordvestlige delene av området som er dominert av skog, mens de laveste estimatene var i øst der områder med dyrket mark hadde blitt opprettet. Skog lagret den største mengden SOC og TN (3,70 Tg C og 0,38 Tg N) etterfulgt av dyrket mark (2,47 Tg C og 0,25 Tg N) og gresslandskap (0,57 Tg C og 0,06 Tg N). Sammenlagt lagret Eastern Mau Forest Reserve tilnærmelsesvis 6,84 Tg C og 0,69 Tg N.

Funnene og den tilhørende dokumentasjonen fra dette forskningsarbeidet er med på å øke vår forståelse av hvilke konsekvenser menneskelig inngripen vil ha i studieområdet. De gir et godt romlig eksplisitt biogeografisk informasjonsgrunnlag til støtte for overvåking av arealressurser og stedsrettede politiske beslutninger for bærekraftig arealforvaltning. Den vitenskapelige tilnærmingen som er anvendt i dette arbeidet kan også benyttes som et verktøy i andre områder med mangelfullt datagrunnlag for å skaffe til veie relevant informasjon, anvendbar i flere sammenhenger, på en kostnadseffektiv måte. I en større sammenheng vil de resulterende arealtype- og jorddatabasene kunne være til støtte for aktiviteter i andre programmer, som for eksempel programmene REDD+, FAO land use – land cover og ISRIC – World Soil Information.

Til slutt gis det en anbefaling om at det formuleres og iverksettes stedsrettede og tidsspesifikke politiske beslutninger for gjenoppretting og konservering av naturlige økosystemer i tillegg til at det tas sikte på å oppnå økt miljømessig bærekraftig landbruksproduktivitet og velfungerende sosioøkonomiske forhold. Bruk av best mulig forvaltningspraksis vil, særlig introduksjon av rasktvoksende, høyproduktive, nitrogensamlende trær med dype røtter (dvs. landbruk integrert med produksjon av skog – såkalt *agroforestry*), være fordelaktig for å minske karbon- og nitrogentapet i områdene med dyrket mark. For jord i skogsområdene, vil langtidslagring av karbon og nitrogen kreve riktig forvaltning og vern av skogen mot ytterligere avskoging og forfall.

**Nøkkelord:** Arealtypeendringer • modellering av arealtypeendringer • organisk karbon i jord • totalnitrogen • jord landskapsmodellering • fjernmåling • GIS • Eastern Mau • Lake Nakuru

## LIST OF PAPERS

The manuscripts written were published, or submitted for publication, in international peer-reviewed journals as follows:

Paper I: **Were, K.O.,** Dick, Ø.B., Singh, B.R. (2013). Remotely sensing the spatial and temporal land cover changes in the Eastern Mau Forest Reserve and Lake Nakuru drainage basin, Kenya. *Applied Geography* 41, 75-86.

Paper II: **Were, K.O.,** Dick, Ø.B., Singh, B.R. (2014). Exploring the geophysical and socio-economic determinants of land cover changes in the Eastern Mau Forest Reserve and Lake Nakuru drainage basin, Kenya. *GeoJournal*, doi: 10.1007/s10708-014-9525-2.

Paper III: **Were, K.O.,** Singh, B.R., Dick, Ø.B. (2014). Effects of land cover changes on soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya (Chapter 6). In: Lal, R., Singh, B.R., Mwaseba, D.L., Kraybill, D., Hansen, D.O., Eik, L.O. (eds.), *Sustainable intensification to advance food security and enhance climate resilience in Africa*. Springer International Publishing, Switzerland, doi: 10.1007/978-3-319-09360-4_6 (in press).

Paper IV: **Were, K.O.,** Singh, B.R., Dick, Ø.B. (2014). Spatially-distributed modelling and mapping of soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya. Submitted to *Catena* (under review).

## LIST OF ABBREVIATIONS

| | |
|---|---|
| BD | Bulk density |
| BF | Bamboo forest |
| BF2C | Bamboo forest to cropland |
| C | Carbon |
| CEC | Cation exchange capacity |
| CTI | Compound topographic index |
| DEM | Digital elevation model |
| ETM+ | (Landsat 7) Enhanced thematic mapper plus |
| FAO | Food and agricultural organisation of the United Nations |
| GIS | Geographic information systems |
| GPS | Global positioning systems |
| GWR | Geographically weighted regression |
| GWRK | Geographically weighted regression-kriging |
| LULCC | Land use, land cover changes |
| MLR | Multiple linear regression |
| MLRK | Multiple linear regression-kriging |
| MSS | (Landsat 1) Multi-spectral scanner system |
| NDVI | Normalized difference vegetation index |
| NF | Natural forest |
| NF2C | Natural forest to cropland |
| OLI | (Landsat 8) Operational land imager |
| PF | Plantation forest |
| PF2C | Plantation forest to cropland |
| Pg | Petagrams (1 Pg = $10^{15}$ g = 1 billion tons) |
| REDD | Reduction of emissions from deforestation and forest degradation |
| SOC | Soil organic carbon |
| Tg | Teragrams (1 Tg = $10^{12}$ g = 1 million tons) |
| TM | (Landsat 5) Thematic mapper |
| TN | Total nitrogen |
| TWI | Topographic wetness index |
| USGS | United States geological survey |
| UTM | Universal transverse Mercator |
| WGS | World geodetic system |

## 1.0 BACKGROUND AND STATUS OF KNOWLEDGE

### 1.1 Land cover and land cover changes

Land cover, defined as the biophysical attributes of the Earth surface and immediate subsurface (Lambin et al., 2003), is a key determinant of the state of physical and human environments. It is quite distinct from land use, which refers to the manipulation of these attributes by humans to meet different needs; for example, agriculture, ranching, and grazing. This implies that land use affects land cover, and the attendant land cover changes affect land use. It also suggests that land cover links the physical and human environments by providing various beneficial goods and services, so-called ecosystem services, to humans. The ecosystem services *per se* include: provisioning services (e.g., food, water, fuel, fibre), regulating services (e.g., climate regulation, disease control), cultural services (e.g., spiritual and aesthetic benefits), and supporting services (e.g., nutrient cycling) (Millennium Ecosystems Assessment (MEA), 2005). These goods and services are vital for human survival and general well-being (Fig. 1).

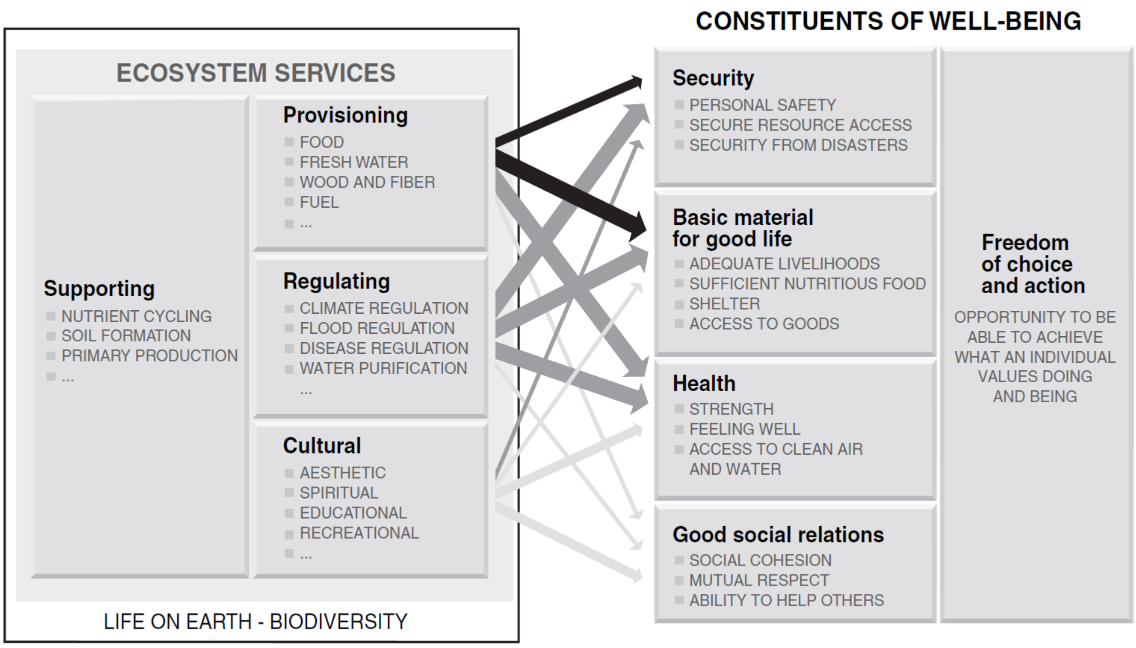



**Fig. 1:** Ecosystem services and their linkages to human well-being. (Source: MEA, 2005).

Land cover varies in space and time. Human manipulations (i.e., land uses) and natural processes (e.g., climatic variability) have produced shifts on the Earth's surface for centuries. However, the current rates and magnitudes of land cover changes are unprecedented and drive global environmental changes owing to the rapid demographic changes and technological advances. Agricultural activities have expanded into the native systems (forests, savannas, and steppes) following the increasing demands for food, fuel, and fibre by the rising human population. The land cover changes can either be complete replacements of one cover type by another (land cover conversions), or subtle changes that affect the character of the land cover without changing its overall classification (land cover modifications) (Lambin et al., 2003). Further, the changes can either be gradual and localized (progressive), or rapid and abrupt (episodic) due to the interactions between land use and climatic factors. The patterns of land cover change are often a culmination of complex interactions between the behavioural and structural factors in specific spatial and temporal contexts (Briassoulis, 2000; Lambin et al., 2003; Overmars and Verburg, 2005). Thus, an understanding of how these factors influence land use decision-making within these contexts is important to resolve land cover change issues. Of equal importance is the knowledge of the state of land cover, changes that occur, where and when they occur, and the rates at which they occur (Lambin, 1997). Human actions that originate from the intended land use and alter land cover are known as proximate drivers (e.g., agricultural expansion, deforestation), while the forces that underpin these actions are referred to as underlying drivers (e.g., population dynamics, policies, climate variability). The former mainly operate at the local level, while the latter operate diffusely at the regional and global levels. Environmental impacts of land cover changes and socio-ecological responses to these impacts may feedback to amplify, or suppress the driving factors leading to new changes (Geist and Lambin, 2001). Figure 2 illustrates this conceptualization of the links between the proximate drivers, underlying drivers, land-use, and land-cover change.

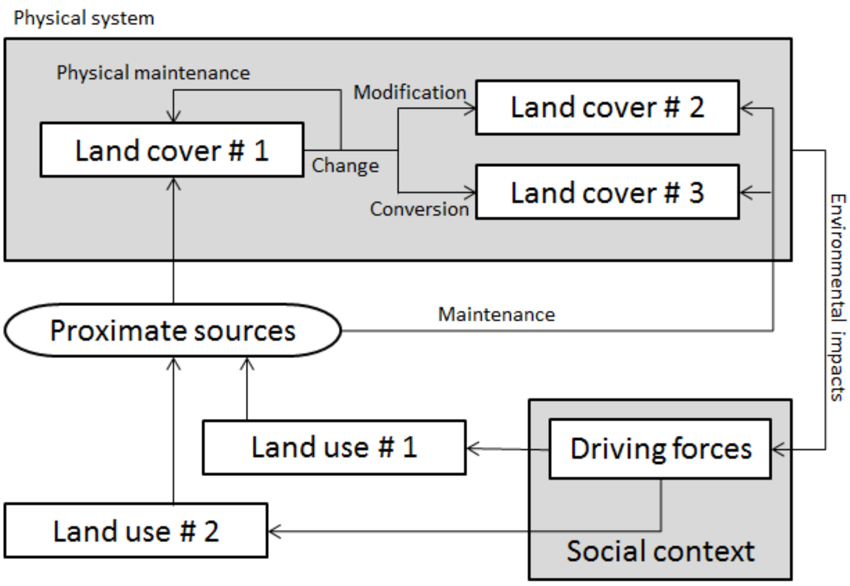



**Fig. 2:** Linkages between human activities, land use, and land cover. (Source: Geist and Lambin, 2001, after Turner et al., 1993).

### 1.1.1 Impacts of land cover changes

The impacts of land cover changes are severe, and have been felt globally because of the role of land cover on the Earth system's processes. Some of the impacts include: loss of biodiversity, desertification, soil degradation (leading to low productivity, food insecurity, and poverty), pollution through biomass burning and agro-chemicals (leading to health problems and acidification of precipitation), climatic changes through modification of surface albedo and carbon (C) sources, and hydrological changes through alteration of the evapotranspiration regime. Land use and land cover change (LULCC) has been an item for assessment and action on the agenda of several global environmental change forums over the last few decades because of these impacts. For example, calls for substantive studies on land use changes were made at the Conference on Human Environment held in Stockholm in 1972, as well as at the United Nations Conference on Environment and Development (UNCED) in 1992 (Fan et al., 2007). In 1993, the International Geosphere and Biosphere Programme (IGBP) and International Human Dimension Programme (IHDP) constituted a working group to define the research agenda and promote LULCC research. This working group suggested three core areas for LULCC research; namely, situation assessment, modelling and projecting, and conceptual scaling. The famous Kyoto Protocol and United Nation Framework Convention on Climate Change (UNFCCC) also embraced Land Use, Land Use Change, and Forestry (LULUCF) activities among the measures to mitigate climate change by the Parties.

#### 1.1.1.1 Impacts of land cover changes on soils

Soil is a valuable natural resource that sustains life on Earth by providing various ecosystem services. For example, it offers a physical matrix, chemical environment, and biological setting for producing food, fibre, fodder, renewable energy, and raw materials, as well as for regulating the exchange of material, energy, water, and gas within the lithosphere–hydrosphere–biosphere–atmosphere system (Osman, 2014). The rapid LULCC, especially conversion of native- to agro-ecosystems, due to the rising global population is straining the soils. Agricultural uses of soil alter its physical, chemical, and biological properties leading to soil degradation. This is manifested through erosion, acidification, salinization, nutrient and organic matter depletion, compaction, crusting, hardsetting, and decline in soil biodiversity, among others. Such alterations further interfere with the composition and functions of ecosystems. In the face of climate change, research spotlight has been on the impacts of LULCC on soil organic carbon (SOC) storage. The rationale for this is that the world's soils contain about 1500 Pg C (1 Pg=$10^{15}$g) to 1m depth, which is twice the amount of C in the atmospheric (750 Pg C) pool and almost three times the amount in the biotic pool (610 Pg C) (Lal, 2004; Smith, 2004, 2008). Thus, even slight changes in SOC pool can have repercussions on the global climate and biogeochemical cycles. Many studies have reported that converting native systems to croplands diminishes SOC pool (Brown and Lugo, 1990; Murty et al., 2002; Osher et al., 2003; Braimoh and Vlek, 2004; Evrendilek et al., 2004; Houghton and Goodale, 2004; Powers, 2004; Lemenih et al., 2005; Yimer et al., 2007; Eaton et al., 2008; Don et al., 2011; Girmay and Singh, 2012; Muñoz-Rojas et al., 2012; Wiesmeier et al., 2012; Demessie et al., 2013; Jafarian and Kavian, 2013). This is ascribed to disruption of the balance between the inputs of C through litterfall, dead roots, belowground biomass, and root exudates, and the outputs through leaching, decomposition, and erosion in the soil system (Detwiler, 1986; Eaton et al, 2008).

### 1.1.2 Remote sensing of land cover and land cover changes

The global concern about the impacts of land cover changes has seen the execution of numerous studies with a view to understanding the dynamics and trends of land cover changes, processes that drive them, impacts on Earth systems, as well as the future trajectories. This has largely comprised characterization of the biophysical cover of the Earth's surface over time to establish what changes occur, where and when they occur, and the rates at which they occur. Remote sensing, which is "... the science and art of obtaining information about an object, area, or phenomenon through the analysis of data acquired by a device that is not in contact with the object, area, or phenomena" (Lillesand et al., 2008, p.1), has been instrumental in this respect since the launch of the first Earth Resources Technology Satellite (ERTS-1 or Landsat 1) in 1972 for Earth observation. Thereafter, many other satellites (e.g., ENVISAT, Terra, SPOT, IRS, EOS) with different sensors (e.g., TM, ETM+, OLI, LISS, MODIS, MERIS, VGT, HRV, HRG, HRVIR, ASTER, AVHRR, MISR, Hyperion) have been launched (Rogan and Chen, 2003). These provide improved data in terms of spatial, spectral, and radiometric resolutions. Essentially, when electromagnetic radiation (EMR) from an energy source reaches the Earth's surface, it is reflected, transmitted, or absorbed depending on the properties of the surface features. Remote sensors aboard aerial and space-borne platforms detect and record the emitted or reflected EMR, which is processed to form remotely-sensed imagery. In the ideal, each feature on the Earth's surface has a unique spectral signature (spectral response over a range of wavelengths), which permits its clear discrimination on the remotely-sensed imagery.

There has been a growing application of remote sensing in LULCC research because it offers synoptic view and inexpensive, detailed, consistent, multi-date, quantitative, spatially-explicit, and repetitive Earth's surface data, which is compatible with geographic information systems (GIS). Remote sensing systems also allow processing of large quantities of multi-temporal, multi-resolution, and multi-spectral data, as well as multi-sensor data fusion (Lu et al., 2004). Over the last two decades, a number of global and regional land cover mapping projects have been successfully implemented using remote sensing and GIS technologies, including, *inter alia*, the GLOBCOVER 2005, Global Land Cover 2000 Project, IGBP Global Land Cover Mapping Project 1997, GeoCover LC, FAO Africover Land Cover Mapping Project 2004, FAO Asiacover Land Cover Mapping Project, and CORINE Land Cover project of the European Union countries.

The substantial advances in development of algorithms for discriminating land cover and detecting changes on remotely-sensed imagery have also given impetus to the growing popularity of remote sensing in LULCC research. Algorithm development is still an active area of research (Lu et al., 2014). The image classification algorithms that have been developed and applied in LULCC research so far include: parametric per-pixel classifiers (e.g., maximum likelihood), non-parametric per-pixel classifiers (e.g., artificial neural networks, support vector machine, expert system, decision trees), sub-pixel classifiers (e.g., spectral mixture analysis, fuzzy-set), object-oriented classifiers, GIS-based per-field classifiers, and contextual classifiers (Lu and Weng, 2007). Similarly, the change detection algorithms that have been developed and used in LULCC research vary from simple and straightforward ones based on spectral classification of the input remotely-sensed data (e.g., post-classification comparison, direct multi-date classification) to complex ones based on radiometric changes between the input remotely-sensed data (e.g., principal component analysis, change vector analysis, multi-dimensional feature space analysis, temporal trajectory analysis, image differencing, vegetation index differencing, image regression, image ratioing, background substitution, artificial neural networks, Gramm-Schimdt, chi-square and bi-temporal linear data transformations, Li-Strahler reflectance models, spectral mixture models, biophysical parameter estimation models). These algorithms find their best description in Singh (1989), Mas (1999), Coppin et al. (2004), Lu et al. (2004), Chen et al. (2012), and Hussain et al. (2013). Since the contexts of landscape changes are diverse and complex, selection of appropriate satellite imagery, imagery acquisition dates, image classification and change detection schemes and algorithms, as well as the analyst's skills and knowledge of the area are important for successful change analyses.

### 1.1.3 Modelling of land cover changes and soil landscapes

Mitigation of LULCC impacts demands, not only, the characterization of land cover changes, but also the appreciation of the interactions between the biogeophysical and socio-economic factors, which operate at different spatial and temporal scales resulting in the changes (Overmars and Verburg, 2005). However, the interactions are functionally and structurally complex; for example, they involve multiple factors that are highly interrelated and vary in space and time. Such complexities can limit the understanding of the interactions. There are various approaches to complexity. One approach views complexity of human-nature systems as the result of a small number of controlling processes, and not random association of a large number of interacting factors (Holling, 2001). That is, the system's nature can be captured and described by single key variables since most of the system's features tend to shift based on a small set of key state variables.

In LULCC research, (spatial) models are pivotal tools for unravelling the complexity of the driving factors, as well as for projecting the future evolution of the patterns of LULCC. The models use artificial representations of the interactions (i.e., simple and easily interpreted proxies) within the land use-land cover system to explore and understand its dynamics and possible futures (Verburg et al., 2006). The artificial representations of the interactions are constantly refined as deeper insights into the organization and functioning of land use-land cover systems is gained. Modelling reveals gaps in knowledge; for example, when important LULCC mechanisms in an area, which could not be observed in the field, are detected through analysis of the sensitivity of land cover patterns to variations in driving factors. The findings may also lead to new insights, guide further analysis of LULCC processes, or formalize knowledge (Verburg et al., 2006).

Over time, several modelling approaches have emerged in LULCC research thanks to the utility of models in understanding the human-nature systems. These approaches have been classified variously by scholars according to the underlying theories, processes being studied, techniques applied, and purpose of the models, among others. Based on the techniques applied, Briassoulis (2000), Lambin et al. (2000), and Heistermann et al. (2006) categorized LULCC models as: (i) empirical-statistical models, which derive quantitative relationships between the observed land cover changes and spatial variables using multivariate statistics (e.g., logistic regression, CLUE models), (ii) stochastic simulation models, which describe stochastically processes that move in a sequence of steps through a set of states (e.g., Markov chain models), (iii) optimisation (economic) models, which apply optimisation techniques (e.g., linear programming, general equilibrium model) for optimal allocation of land resources (e.g., GTAPE-L, GTAPEM), (iv) dynamic (process-based) simulation models, which simulate the biophysical and socio-economic processes that produce the patterns of LULCC by, systematically, reducing complex ecosystems into a small number of differential equations (e.g., SALU model), and (v) integrated (hybrid) models, which couple the best elements of the existing models in the most appropriate way to answer specific questions (e.g., land use choice module that links IFPSIM and EPIC models, IMAGE, Integrated Model to Predict European Land Use (IMPEL)).

Besides the human-nature systems, numerical tools have also been instrumental in analysing soil-landscape interactions and deriving spatially-exhaustive information on soil functional properties (e.g., C and N stocks) to assess, monitor, and manage ecosystems. This is the essence of digital soil modelling and mapping (DSMM), which has been an active research front since the late 1990s. In DSMM, the variability of a target soil property is explained by its relationships with soil-forming factors, such as topography, climate, land use, vegetation, and soil type. This is underpinned by Jenny's (1941) seminal work, which considered soil development as a function of climate (*c*), organisms (*o*), relief (*r*), parent material (*p*), and time (*a*). The function was later expanded by McBratney et al. (2003) to include soil properties (*s*) and space (*n*) under the well-known *scorpan* framework. The array of statistical, geostatistical, and machine learning tools that have been used in DSMM thus far include: multiple linear regression (Meersmans et al., 2008), partial least square regression (Amare et al., 2013), generalized linear models (Yang et al., 2008), linear mixed models (Doetterl et al., 2013; Karunaratne et al., 2014), geographically weighted regression (Mishra et al., 2010; Kumar et al., 2013), kriging (Cambule et al., 2014), regression-kriging[1] (Hengl et al., 2004, 2007; Kumar et al., 2012; Dorji et al., 2014; Martin et al., 2014), artificial neural networks (Malone et al., 2009; Li et al., 2013; Dai et al., 2014), boosted regression trees (Martin et al., 2011), random forests (Grimm et al., 2008), support vector regression (Ballabio, 2009), and rule-based models (Lacoste et al., 2014). The interested reader is referred to McKenzie and Ryan (1999), McBratney et al. (2003), Scull et al. (2003), and Grunwald (2009) for detailed discussions of DSMM.

## 1.2 Land cover changes in Lake Nakuru drainage basin and Eastern Mau Forest Reserve

Land cover has changed rapidly in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve, since the pre-colonial era to date. Before the British settlement, the area was barely

[1] This includes a combination of any regression-based technique with ordinary kriging.

populated and dominated by natural vegetation and wildlife (Krupnik, 2004; Odada, 2006). Indigenous trees species characterized the Eastern Mau forest on the eastern slopes of Mau escarpment, Eburru forest in the south, Menengai forest in the north, and Dundori (Bahati) forest in the north-east. The *Ogieks* who were hunter-gatherers and bee keepers inhabited parts of Eastern Mau forest, while shifting cultivators and pastoralists occupied the valley floors and grasslands. By 1900, the area was firmly under the British, who were interested in extracting timber from the forests and settling in the fertile lands. Exploitation of forest resources intensified as demand on European colonies to supply raw materials (e.g., wood) during World War I increased. In areas where indigenous trees were felled, fast-growing, exotic tree species were replanted to meet the rising industrial and domestic demand for wood. To sustain this, the *shamba* system was introduced. This system allowed the local farmers to grow food crops in small plots where trees had been felled and, simultaneously, plant and nurture tree seedlings for a specific period of time.

After independence in 1963, felling of trees was mainly undertaken by large timber enterprises (e.g., Timsales Ltd). During this time until the early 1980s, there was pressure on the government to allocate land to the landless people, which led to the establishment of several resettlement schemes (i.e., Keriri, Gichobo, Naishi, and Bagaria) on the large-scale farms formerly owned by the white settlers. The large-scale farms were divided into smaller units and allocated to individuals; hence, the proliferation of small-scale farms in the area. Demand for land increased as the population expanded in the 1980s and 1990s resulting in illegal encroachments and loggings in the adjacent indigenous forests for settlement, cultivation, and fuel. The UNEP (2009) reported that 47% of the Lake Nakuru drainage basin was under forest and natural vegetation in 1970, but between 1973 and 2003, 49% of these had been cleared. Baldyga et al. (2007) also found that between 1986 and 2003, one-fifth of the forests in the upper reaches of the Njoro River watershed had been lost.

The ill-advised political decision made in 1994 (and legitimized in 2001) to excise 353 $km^2$ of the Eastern Mau Forest Reserve (Odada, 2006) partly explains the loss of forests. In theory, the government exerted powers provided by the Forest Act 1942 (Cap.385) to excise part of the forest and resettle about 3,000 *Ogiek* families that dwelt in the indigenous forests, and victims of ethnic clashes that occurred in Molo, Likia, Mauche, and Njoro in the 1990s (Government of Kenya, 2009). But in practice, patronage politics ensured that most of the people who secured land in Mariashoni, Nessuiet, Teret, Likia, Baraget, and Sururu forests were others who came from Koibatek, Baringo, Bomet, Kericho, Bureti, and Transmara districts. Consequently, by 1997, almost 50% of the Eastern

Mau Forest Reserve had undergone deforestation and degradation (Fig. 3), and over 30,000 people had migrated to the area (Krupnik, 2004; Odada, 2006). The Lake Nakuru drainage basin (including the Eastern Mau Forest Reserve) has also experienced rapid urban growth and increase in human settlements because of population growth. Currently, it has over 1.5 million residents with over 300,000 living in Nakuru municipality. The rest live in small towns, market centres, and rural settlements.


**Fig. 3:** Human activities in the Eastern Mau Forest Reserve: (**a**, **b**, and **c**) illegal felling of trees; (**d**) charcoal burning; and, (**e**) agricultural expansion and human settlement. (Source: Author).

## 2.0 RESEARCH PROBLEM AND RATIONALE

The demographic and landscape changes, especially forest conversion, have had implications for ecosystem services in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve (Fig. 4). Some studies conducted in the area revealed that water quality (Shivoga et al., 2007), hydrological regime (Mwetu et al., 2009), temperature distribution (Hesslerová and Pokorný, 2010), soil properties (Enanga et al., 2011), and biodiversity (Kibichii et al., 2007; Raini, 2009) had been adversely affected. For example, farmers have had to grapple with erratic weather patterns, while wildlife, bird, and fish populations in the Lake Nakuru National Park have been threatened by the dwindling discharge of the Njoro River. Additionally, sediments and agro-chemicals (e.g., phosphorous) transported by rain water from the cultivated areas have severely affected the surface and ground water quality.

In view of this, there is need to make effective decisions to mitigate the adverse impacts of the ongoing land cover changes, and to determine the future landscape scenario for environmental protection and management in the area. This is essential to achieve sustainable development, millennium development goals, and *Vision 2030* in Kenya as a whole. Effective decision-making calls for availability of accurate, consistent, quantitative, multi-temporal, and spatially-explicit information on the state of land cover and other important biophysical variables, such as soil properties, biodiversity, etc. Such information can, for instance, guide the formulation of policies for delineating and managing priority areas and land resources. It can also be useful for assessing trends and impacts (e.g., decline in SOC storage), explaining processes, and predicting future patterns of land cover changes. Unfortunately, the archives of biophysical information are poor in Kenya. Thus, remote sensing offers a practical means of deriving such information for systematic mapping and monitoring of land cover at multiple spatial and temporal scales, as well as for modelling and mapping the spatial patterns of environmental variables (e.g., soil properties). Geographic information systems (GIS), on the other hand, offer a platform for analysing, manipulating, modelling, and visualizing the remotely-sensed and other spatial data.

This research aimed to bridge the existing biophysical information gap and, thereby, contribute knowledge towards rational policy making for sustainable land management in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve.

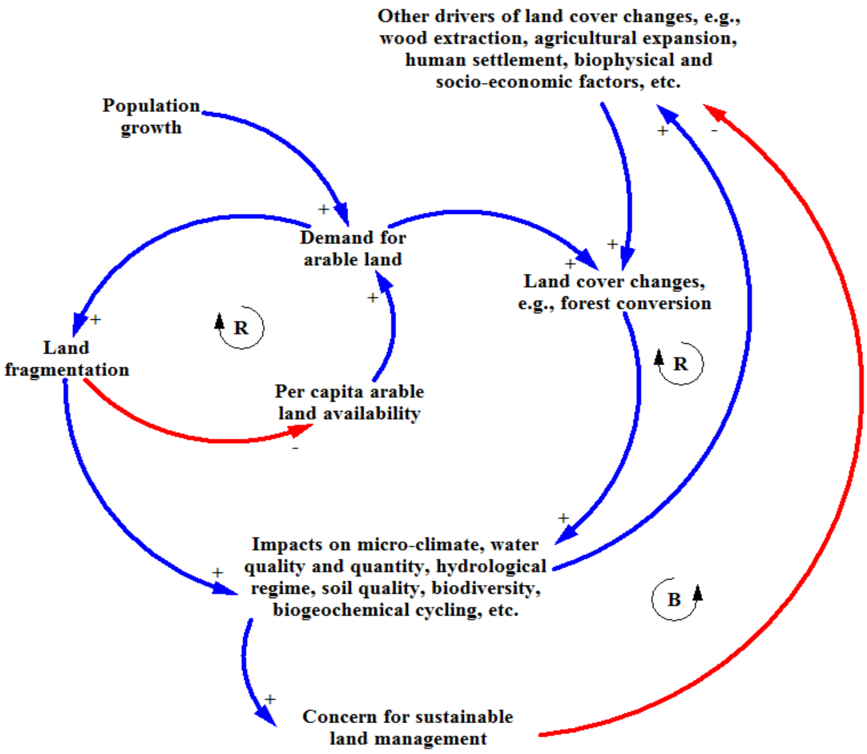



**Fig. 4:** Causal loop diagram of the research problem. Each arrow indicates a causal relationship, which can be large or small, immediate or delayed, an increasing (+) or decreasing (-) effect. The letters R and B denote reinforcing and balancing feedback loops, respectively.

## 2.1 Research goal and objectives

The overarching goal of the research was to contribute knowledge to support the formulation of policies for mitigating the adverse impacts of land cover changes and sustainable management of land resources in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve. To achieve this, the specific objectives were:

a) To detect, quantify, and map the land cover changes that had occurred in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve over time,

b) To analyse the linkages between the observed land cover change processes and the geophysical and socio-economic factors in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve,

c) To assess the effects of forest-cropland conversion on soil organic carbon (SOC) and total nitrogen (TN) stocks in the Eastern Mau Forest Reserve, and

d) To model and map the spatial distribution of SOC and TN stocks in the Eastern Mau Forest Reserve.

## 2.2 Research questions

a) What are the spatial patterns, rates, magnitudes, and nature (trajectories) of the land cover changes that have occurred in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve over time?

b) Can remote sensing and GIS techniques classify the past and present land cover in a spatially heterogeneous Kenyan landscape to acceptable levels of accuracy?

c) What are the significant geophysical and socio-economic determinants of the observed land cover change processes in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve?

d) How have SOC and TN stocks in the Eastern Mau Forest Reserve responded to the forest-cropland conversion? Are the stocks between forests and croplands equal, or significantly different?

e) What are the significant environmental factors that control, and can be used to estimate and map the spatial patterns of, SOC and TN stocks in the Eastern Mau Forest Reserve?

# 3.0 MATERIALS AND METHODS

## 3.1 Study area

The study area was the Lake Nakuru drainage basin and Eastern Mau Forest Reserve (Fig. 5) covering ~2000 km$^2$. It lies in the Kenyan Rift Valley system, bounded by the latitudes 0º

10´- 0º 45´S and longitudes 35º 40´- 36º 5´E, with the altitudes ranging from 1750 to 3090 m above sea level. The Makalia, Njoro, Naishi, Lamuriak, Enderit, and Ngosorr Rivers drain down the area into Lake Nakuru, while the Nessuiet flow into Lake Bogoria and the Rongai River into the Baringo. The landforms include mountains and major scarps, hills and minor scarps, plateaus, volcanic footridges, uplands, volcanic and lacustrine plains, and bottomlands. The soils are classified as *Andosols, Planosols, Vertisols, Nitisols, Regosols, Calcisols, Solonetz*, and *Phaeozems* (Jaetzold et al., 2010; Wanjogu et al., 2010), the parent materials of which originated from volcanic rocks (e.g., basalts, trachytes, phonolites, pumice tuffs, lavas) and associated sediments of tertiary-quaternary age (McCall, 1967).

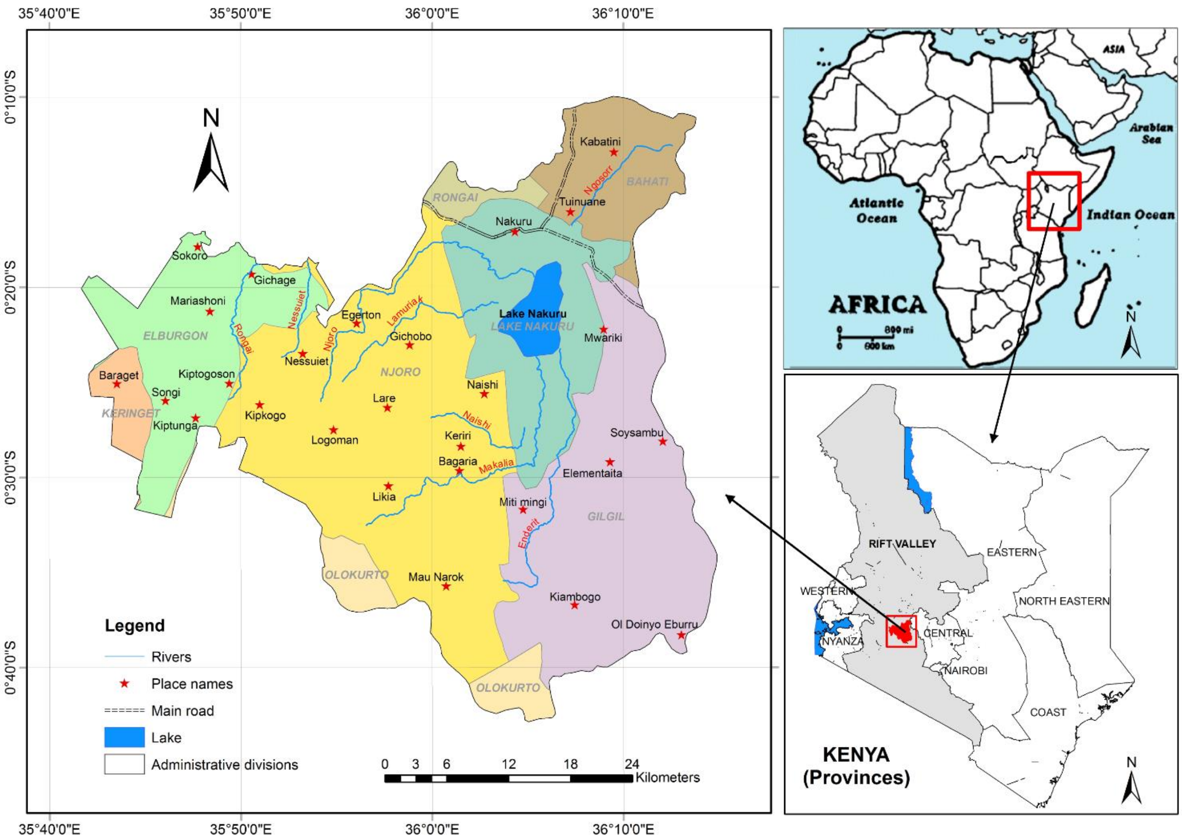



**Fig. 5:** Geographical location of the study area

The climate varies from cool and humid to hot and humid depending on the altitude and topography. Higher areas at Mau escarpment receive substantial rainfall (~1200mm), which decreases notably (~700mm) on the lower areas around Lake Nakuru. The rainfall pattern is bimodal with the long rains falling between March and May, and short rains between November and December because of the seasonal north-south movement of the Inter-Tropical Convergence Zone (ITCZ) (Odada et al., 2006). The vegetation comprises grasslands and scrublands in the lower parts, acacia trees along the lakeshore, riverine vegetation along the rivers, and forests in the higher areas. The major land use systems, which also contribute to human economy, are agriculture, ranching, pastoralism, forestry, urban and industrial centres, and tourism and wildlife conservation. Land ownership is varied with the government owning the national park and forest reserves, subsistence farmers owning the small-scale farms, and commercial farmers leasing the large-scale farms, or ranches. The area is also an important centre for a growing human population. Currently, it has over 1.5 million inhabitants with over 300,000 living in the rapidly expanding Nakuru Municipality (www.opendata.go.ke/). The rest live in small towns, market centres, and rural settlements.

## 3.2 Data

Figure 6 summarizes the data and methods used in the research.

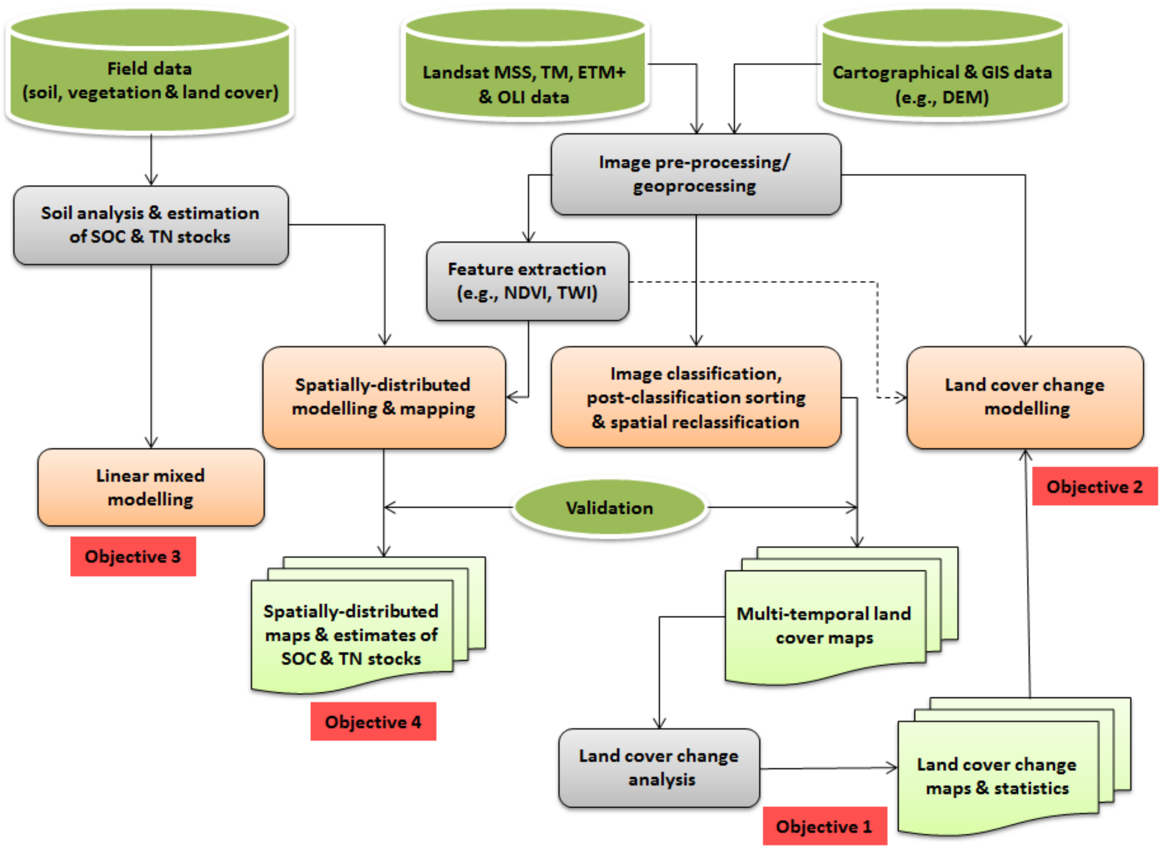



**Fig. 6:** An overview of the research data and methodological flow

### 3.2.1 Field data

Fieldwork was conducted between July and August 2012 to collect land cover, soil, and land use management data. Prior to the land cover campaign in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve, random sampling strategy was designed for objective selection of sampling units into the sample. Sampling units were randomly generated in a GIS, the ground size of which was 30×30m to coincide with the spatial resolution of the Landsat imagery. In the field, the biophysical land attributes (e.g., percent tree cover) were recorded and georeferenced using a hand-held GPS receiver at 450 sampling locations. Interviews were also conducted with key informants, particularly the local administrators, forest managers, farmers, and community group leaders. The data were used to determine the land cover types on Landsat imagery, validate the extracted land cover map for 2011, and understand the land use-land cover history (Paper I, II, and III).


**Fig. 7:** Different land cover types: (**a**) natural forest (mixed; NF); (**b**) plantation forest (pine and cypress; PF); (**c**) Bamboo forest (BF); and, (**d**) croplands converted from forests (NF2C, PF2C, or BF2C). (Source: Author).

Similarly, before the soil campaign in the Eastern Mau Forest Reserve, sampling points were generated in a completely randomized design using agro-ecological zone map as the base in a GIS. A map showing the distribution of the sampling points was produced to support their identification in the field. At each sampling point, a 30×30m plot was laid, and an auger used to collect samples from the centre and four corners of the plot, at 0-15cm and 15-30cm depths. The samples taken from corresponding depths were mixed thoroughly and bulked into one composite sample of about 500g. To determine bulk density (BD), a core sampler (5 cm in diameter and 5cm in height) was used to collect one undisturbed sample at the centre of each plot and each depth. The geographical coordinates, elevation, vegetation, and land management practices were also recorded at each plot. To fulfil the third objective, soil samples were taken in a similar way at 60 plots within sites with similar climate, soil type, and slope in order to minimize variations. Four to fifteen sampling plots were laid out in a completely randomized design within the natural forests (NF), plantation forests (PF), bamboo forests (BF), and croplands that had been established on natural forests (NF2C), plantation forests (PF2C), and bamboo forests (BF2C) at these sites (Fig. 7). In total, 440 samples (220 for each depth) were collected for chemical and physical analysis, and another 440 samples for BD determination at the National Agricultural Research Laboratories. The results from these data are presented in Paper III and IV.

### 3.2.2 Remotely-sensed data

Terrain-corrected Landsat 1 Multispectral Scanner System (MSS), Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper plus (ETM+), and Landsat 8 Operational Land Imager (OLI) imagery acquired in 1973, 1985, 2000, 2011, and 2013, respectively, were procured from the USGS archive (http://earthexplorer.usgs.gov/). Normalized Difference Vegetation Index (NDVI) was derived after conversion of the digital numbers of OLI band 4 (red) and 5 (near infra-red) to top-of-atmosphere reflectance, while principal components bands were obtained from principal component analysis of OLI bands 2, 3, 4, 5, 6, and 7. These data were used for the analyses in Papers I, II, III, and IV.

### 3.2.3 Cartographical and GIS data

Digital data in raster and vector formats were obtained from the existing spatial databases, pre-processed, and used for the various analyses (Papers I, II, III, and IV). The data included: topographical maps (Survey of Kenya), Google™ earth imagery, Africover land cover map (www.fao.org), administrative units, towns, villages, roads, forests, rivers and agro-ecological zones (www.ilri.org/gis), population (1989 and 2009) (www.opendata.go.ke/), digital elevation model (DEM) (http://srtm.csi.cgiar.org/), soils (Kenya Soil Survey), and rainfall and temperature (www.worldclim.org). Primary and secondary terrain attributes, including slope, aspect, curvature, and compound topographic index (CTI) (or topographic wetness index (TWI)) were extracted from the DEM.

## 3.3 Methods

Various procedures were followed in order to fulfil the stated research goal and objectives, as well as to answer the research questions (see also Fig. 6):

### 3.3.1 Spatio-temporal mapping and analyses of land cover changes

a) **Data pre-processing**: The downloaded Landsat files were unzipped and stacked, transformed to the Universal Transverse Mercator system (UTM WGS84 Zone 37S), geometrically co-registered, atmospherically and radiometrically corrected using image-based COST method (Chavez, 1996), and subsets extracted.

b) **Classification scheme design**: Land cover classes were defined based on the field land cover data, as well as on modification of the definitions used by Anderson et al. (1976) and the USGS' National Land Cover Database 2006 (http://www.mrlc.gov/nlcd06_leg.php).

c) **Image classification, post-classification processing and spatial reclassification**: Partitioning, hybrid classification, and spatial reclassification technique was applied to discriminate the land cover types on the image subsets. This produced land cover maps for 1973, 1985, 2000, and 2011.

d) **Accuracy assessment**: Visual inspections, ancillary data (topographical maps, Africover land cover map), field data, and temporally-invariant land cover data (Fortier et al., 2011) were used to qualitatively and quantitatively validate the land cover maps. Statistical measures of map accuracy (overall, producer's, and user's accuracy, Kappa statistic) were computed and presented on error matrices (Campbell, 2002; Foody, 2002).

e) **Land cover change detection**: The land cover maps for 1973 (resampled to 30 m), 1985, 2000, and 2011 were overlaid in post-classification comparison to detect the pixel by pixel land cover changes between 1973-1985, 1985-2000, 2000-2011, and 1973-2011. This generated cross-tabulation matrices and land cover change maps showing the pathways and spatial patterns of land cover change, respectively.

### 3.3.2 Analyses of the linkages between land cover change processes and the geophysical and socio-economic determinants

a) **Data preparation**: Binary maps of the response variables (presence or absence of land cover change) were extracted from the land cover change maps. The maps identified areas of: (i) forest conversion versus stable forests, (ii) grassland conversion versus stable grasslands, and (iii) conversion to croplands versus stable croplands. A suite of 13 candidate geophysical and socio-economic explanatory variables were then selected *a priori* based on existing land use theories, fieldwork experience, data availability, and literature review (Chomitz and Gray, 1996; Geist and Lambin, 2002; Lambin et al., 2003; Braimoh and Vlek, 2005; Aguiar et al., 2007). The variables were rainfall, temperature, soil cation exchange capacity (CEC), soil pH, elevation, slope, aspect, TWI, curvature, distance to road, distance to river, distance to town, and population density. These data were rasterized, transformed to UTM WGS84 Zone 37S, clipped according to the areas of interest, resampled to 100m, and integrated into the GIS database.

b) **Statistical modelling**: This consisted of sampling spatially-balanced random points from the binary maps of response variables in a GIS, overlaying the sample points with the maps of explanatory and response variables to extract attribute values to the points, cleaning the sample points data (e.g., deleting spurious values), exploring the data (e.g., correlation analysis), and modelling the probability of occurrence of each land cover change process given the set of explanatory variables using binary logistic regression method (Montgomery et al., 2006a; Agresti, 2007).

### 3.3.3 Assessment of the effects of land cover changes on SOC and TN stocks

a) **Physical and chemical soil analysis**: The soil samples were air-dried, ground, sieved, and analysed for different properties. SOC was determined using Walkley-Black wet oxidation method (Nelson and Sommers, 1982), TN using Kjeldahl digestion method (Bremner and Mulvaney, 1982), BD using core method (Blake, 1965), particle size distribution using hydrometer method (Day, 1965), potassium (K) using flame-photometer, calcium (Ca) and magnesium (Mg) using atomic absorption spectrophotometer, pH (1:2.5 soil-water) using pH meter (Okalebo et al., 2002), and phosphorous (P) using Mehlich method (Okalebo et al., 2002).

b) **Estimation of SOC and TN stocks:** SOC and TN stocks (mass C or TN per unit area) for each depth, in addition to the percentage changes in the stocks after forest conversions were calculated.

c) **Data preparation**: This involved transforming climate (mean annual temperature and rainfall), soil (soil type), agro-ecological zones, elevation, slope, and aspect data to UTM WGS84 Zone 36S, creating subsets from the thematic layers, rasterizing the vector layers, resampling the data to 100m, integrating the field and laboratory data into the GIS database as points, extracting attribute values from the raster datasets (slope, rainfall, soil type, etc.) to the points, and summarizing the point data by land cover types and soil depths for statistical analysis.

d) **Statistical analysis:** The summarized data were explored using descriptive and correlation statistics. Linear mixed models (Montgomery et al., 2006b) were then fitted to test the effects of land cover, soil depth, and sampling plot on SOC concentrations and stocks, TN concentrations and stocks, and BD for each category of forest-cropland conversion: NF to NF2C, PF to PF2C, and BF to BF2C in the Eastern Mau Forest Reserve.

### 3.3.4 Spatially-distributed modelling and mapping of SOC and TN stocks

a) **Physical and chemical soil analysis**: see section 2.3.3

b) **Estimation of SOC and TN stocks:** see section 2.3.3

c) **Data preparation**: Twenty candidate environmental predictors were selected based on the *scorpan* conceptual model (McBratney et al., 2003), including climate (mean annual temperature, mean annual rainfall), land cover, elevation, Landsat 8 OLI thermal bands, slope, curvature, aspect, CTI, NDVI, soil properties, and principal component band of the Landsat 8 OLI optical bands. The data were pre-processed by transforming them to UTM WGS84 Zone 36S, extracting subsets from each, resampling to 30m where necessary, integrating soil data from the laboratory analysis (sand content, silt content, clay content, TN, C, pH, Mg, Ca, P, and K) into the geodatabase both as feature points and as raster grids after interpolation, and extracting the attribute values of the other raster datasets (e.g., slope, rainfall, temperature) to the feature points for spatial modelling.

d) **Spatial modelling**: The pre-processed data were explored using descriptive and correlation statistics. Different models were then calibrated, validated, and applied to map the spatial patterns of SOC and TN stocks in the Eastern Mau Forest Reserve. Multiple linear regression (MLR) (Montgomery et al., 2006a), regression-kriging (MLRK) (Hengl et al., 2004, 2007), geographically weighted regression (GWR) (Fotheringham et al., 2002), and geographically weighted regression-kriging (GWRK) (Kumar and Lal, 2011; Kumar et al., 2012) techniques were used to calibrate the models.

## 4.0 RESULTS AND DISCUSSION

### 4.1 The spatial and temporal dynamics of land cover changes

Analysis of multi-temporal Landsat imagery discriminated the six main land cover types in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve between 1973 and 2011;

namely, forests-shrublands, grasslands, croplands, built-up lands, bare lands, and water bodies (Fig. 8). Forests-shrublands dominated the landscape from 1973 to 2000, but by 2011, croplands had become dominant. This highlights transition from a natural to human-dominated landscape. The overall accuracy of the extracted land cover maps for 1973, 1985, 2000, and 2011 ranged between 80 and 89%. These levels of accuracy suggest that the partitioning, hybrid classification, and spatial reclassification approach applied in this study is a promising alternative for successful mapping of heterogeneous landscapes in Kenya.

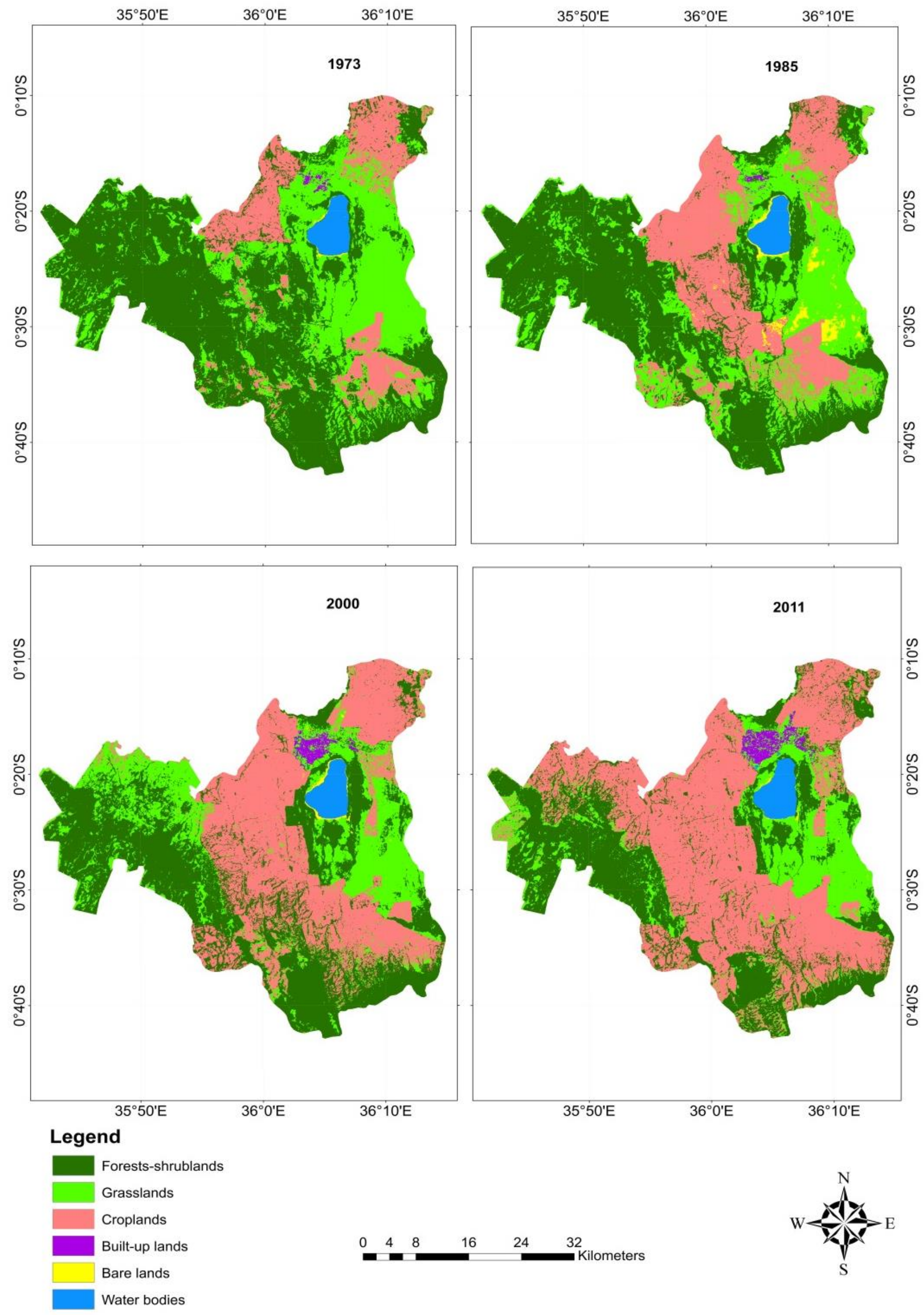



**Fig. 8:** Land cover maps for the Eastern Mau Forest Reserve and Lake Nakuru drainage basin

Moreover, the results revealed that over the 38-year period, the dominant land cover change processes were conversion of native ecosystems (forests-shrublands and grasslands), and expansion of croplands and built-up lands (Table 1). That is, the main proximate drivers of the land cover changes were agricultural expansion, built-up expansion, and wood extraction. Forest-shrublands, grasslands, and croplands had higher magnitudes of change than the other land cover types throughout the period. For example, croplands expanded from 293km$^2$ in 1973 to 953km$^2$ in 2011 at the expense of forest-shrublands and grasslands (Fig. 9; Table 1). This corresponds with various environmental reports of the region (Odada et al., 2006; Baldyga et al., 2007; UNEP, 2009). The ill-advised political decision to excise parts of the Eastern Mau Forest Reserve is one of the societal processes that underpinned the conversion of forests-shrublands. The *hotspots* of forest-shrubland conversion were distributed in the mid-regions and northern side of Lake Nakuru between 1973 and 1985, on the western side between 1985 and 2000, and around the Lake Nakuru National Park, as well as the western and southern parts between 2000 and 2011 (Fig. 9). Built-up lands were the most dynamic judging by their high annual average rates of change; for example, their annual average rate of change was 17% between 1985 and 2000 (Table 1).  This is ascribed to high population growth rates.

**Table 1:** Areal extent, magnitude and the annual average rates of land cover changes

| | Area (km$^2$) | | | | Magnitude of change (km$^2$) | | | | Average rate of change p.a (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Land cover class | 1973 | 1985 | 2000 | 2011 | 73 - 85 | 85 - 00 | 00 - 11 | 73 - 11 | 73 - 85 | 85 - 00 | 00 - 11 | 73 - 11 |
| Forests-shrublands | 1067 | 893 | 797 | 639 | -174 | -96 | -158 | -428 | -1 | -1 | -2 | -1 |
| Grasslands | 589 | 531 | 421 | 331 | -58 | -110 | -90 | -258 | -1 | -1 | -2 | -1 |
| Croplands | 293 | 521 | 714 | 953 | 228 | 193 | 239 | 660 | 6 | 2 | 3 | 6 |
| Built-up lands | 4 | 5 | 18 | 28 | 1 | 13 | 10 | 24 | 2 | 17 | 5 | 16 |
| Bare lands | 4 | 7 | 7 | 2 | 3 | 0 | -5 | -2 | 6 | 0 | -6 | -1 |
| Water bodies | 42 | 40 | 40 | 43 | -2 | 0 | 3 | 1 | 0 | 0 | 1 | 0 |

Note: see also the errata.

The above results depict a clearer picture of land cover changes in the area in terms of what, where, and when the changes occurred, as well as the magnitudes and rates of change. The outputs are a good base of spatially-explicit land cover information for: (i) formulating spatially-targeted policies to restore and conserve the natural ecosystems (e.g., forests); (ii) analysing the biophysical and socio-economic drivers of LULCC; and, (iii) predicting the future spatial patterns of LULCC. The resultant land cover database can also support other research programs, such as soil carbon sequestration, REDD+, and FAO land use-land cover programs. Finally, the elaborate classification approach employed in the study is a useful tool for successful mapping and monitoring of land cover in other complex landscapes in Eastern Africa using the freely available, or inexpensive remotely-sensed data.

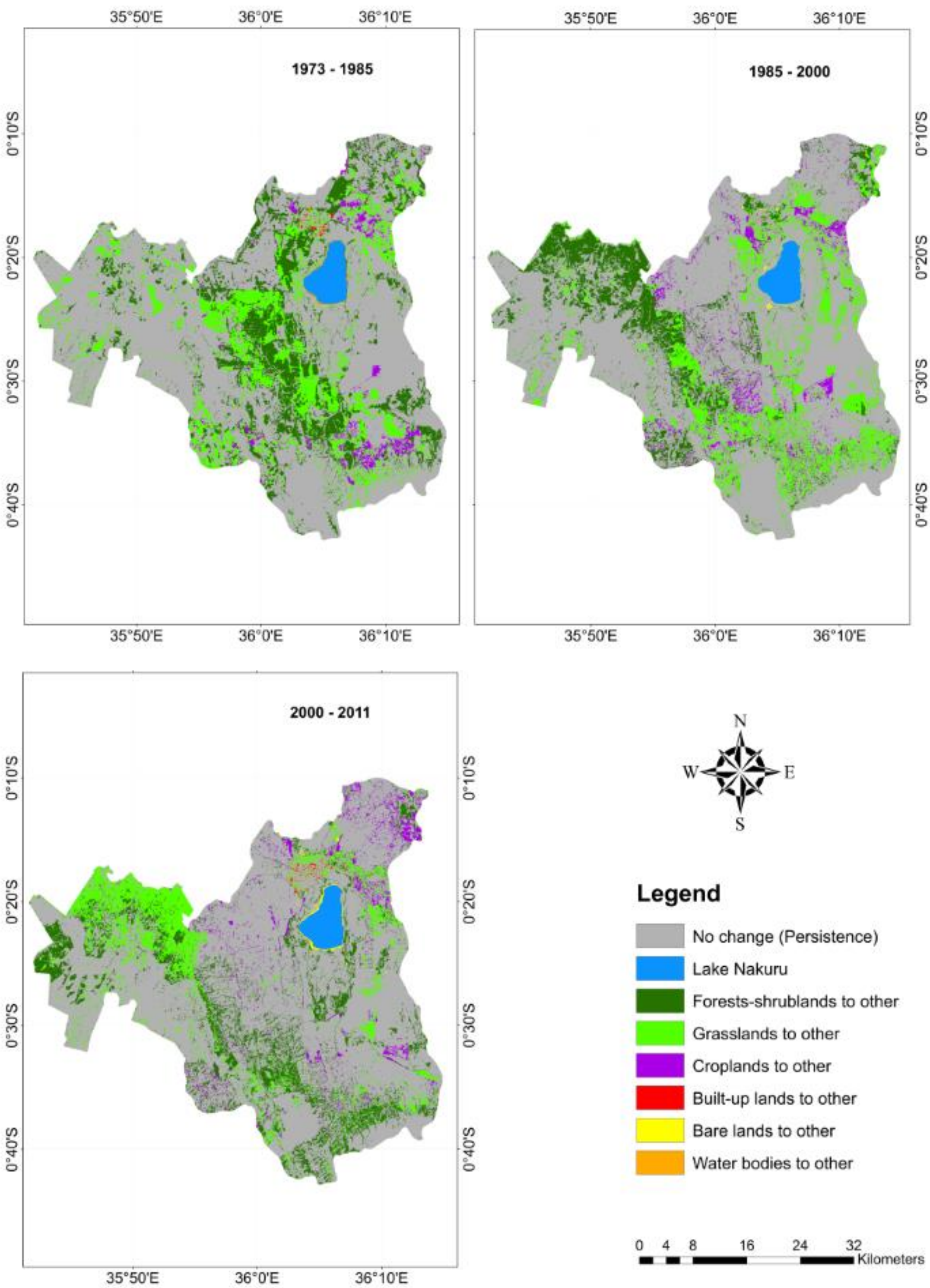



**Fig 9:** Maps showing the spatial distribution of land cover changes in the Eastern Mau Forest Reserve and Lake Nakuru drainage basin from 1973 to 2011. Quantitative information on the nature of these land cover changes is shown in Table 5, Paper I.

## 4.2 The geophysical and socio-economic determinants of land cover changes

Binary logistic regression analyses estimated that between 1985 and 2000, rainfall, soil pH, soil CEC, TWI, aspect, curvature, distance to road, and distance to town were important in determining the occurrence of forest-shrubland conversions (Table 2). As rainfall and soil CEC increased, the probability of forest-shrubland conversion increased, but as distance to road and town increased, the converse was true. This was expected because the good road network, especially on the western part, provided access to agricultural inputs and markets for forest products (charcoal, logs, fuel wood, and timber) in Elburgon, Njoro and Nakuru towns. This proximity to roads also lowered the cost of migration, land access, and land clearance for subsistence farming (Chomitz and Gray, 1996). This is consistent with previous findings; for instance, Müller and Mburu (2009) found that forest clearance in Kakamega forest, western Kenya, tended to occur near roads and market centres. In the same period, rainfall, slope, aspect, curvature, TWI, and distance to road, town and river best explained the occurrence of grassland conversion. Their directions suggested that water and land accessibility were more important in agricultural land use decision-making than market accessibility, which was in contradiction with the von Thünen's theory of agricultural land use. However, this is common in developing countries where subsistence is the overriding goal of agricultural production. In case of surpluses, the middlemen often collect and transport the produce from the farm-gates to the market. Further, in the same period, temperature, aspect, TWI, soil pH, soil CEC, population density, and distance to road, town and river determined the presence of cropland expansion.

**Table 2:** Summary statistics of the logistic regression models for 1985 - 2000

| Explanatory variable | Forest-shrubland conversions | | | Grassland conversions | | | Cropland expansions | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF |
| (Intercept) | -4.216*** | - | - | 6.362*** | - | - | 10.938*** | - | - |
| Rainfall | 0.002*** | 1.00 | 1.37 | -0.003*** | 1.00 | 1.61 | ni | - | - |
| Temperature | ni | - | - | ni | - | - | -0.851*** | 0.43 | 1.95 |
| Elevation | ni | - | - | ni | - | - | ni | - | - |
| Slope | ns | - | - | 0.021· | 1.02 | 2.65 | ns | - | - |
| Aspect | -0.002*** | 1.00 | 1.02 | -0.002*** | 1.00 | 1.06 | 0.002*** | 1.00 | 1.05 |
| Curvature | 0.069*** | 1.07 | 1.08 | -0.121** | 0.89 | 1.74 | ns | - | - |
| TWI | 0.355*** | 1.43 | 1.14 | -0.203*** | 0.82 | 1.95 | -0.075* | 0.93 | 1.28 |
| Soil pH | -0.287*** | 0.75 | 1.90 | ns | - | - | 0.497*** | 1.64 | 1.43 |
| Soil CEC | 0.061*** | 1.06 | 1.87 | ns | - | - | -0.034*** | 0.97 | 1.33 |
| DIST_road | -0.342*** | 0.71 | 1.09 | -0.161*** | 0.85 | 1.32 | 0.505*** | 1.66 | 1.27 |
| DIST_town | -0.183*** | 0.83 | 1.23 | 0.039 ** | 1.04 | 1.17 | 0.036* | 1.04 | 1.48 |
| DIST_river | ns | - | - | -0.317*** | 0.73 | 1.42 | 0.362*** | 1.44 | 1.36 |
| POP_89 | ns | - | - | ns | - | - | 0.001*** | 1.00 | 1.26 |
| *N* | 3870 | | | 1890 | | | 2902 | | |
| Pr >LR $\chi^2$ | 0.00 | | | 0.00 | | | 0.00 | | |
| Pr >Pearson $\chi^2$ | 0.52 | | | 0.41 | | | 0.71 | | |
| Nagelkerke $R^2$ | 0.38 | | | 0.20 | | | 0.36 | | |
| AUC | 0.81 | | | 0.72 | | | 0.80 | | |
| Moran's I. | 0.16 | | | 0.18 | | | 0.19 | | |

Notes: Significance level 0 = '***', 0.001 = '**', 0.01 = '*' and 0.05 = '.'; ni = not included due to high correlation with either elevation, temperature or rainfall; ns = not significant; VIF = variance inflation factor; *N* = number of observations; LR = likelihood ratio, and; AUC = area under curve.

The significance, magnitude, and direction of the explanatory factors varied depending on the process of land cover change and time. For example, between 2000 and

2011, the foregoing factors, in addition to slope and distance to river, and with the exception of TWI, were the significant determinants of forest-cropland conversion (Table 3). This indicates that generalizations about the underlying drivers of land cover changes must be time-specific. Thus, policies should be both spatially-targeted and time-specific.

These results enhance our knowledge of the processes involved in land cover change in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve where empirical evidence of the underlying causal factors is scarce. The knowledge gained can be applied, for instance, to spatially predict the possible future trends of land cover changes. Such predictions, coupled with the knowledge, are beneficial for environmental policy makers, planners, and managers since they can: (i) inform the selection of priority areas for targeted policies or detailed analyses in an effective and efficient manner, and (ii) be linked with biophysical data, e.g., species distribution or carbon storage data, to identify *hotspots* of biodiversity, or carbon losses after land cover conversion.

**Table 3:** Summary statistics of the logistic regression models for 2000 - 2011

| Explanatory variable | Forest-shrubland conversions | | | Grassland conversions | | | Cropland expansions | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF |
| (Intercept) | 4.714*** | - | - | 9.461*** | - | - | -7.558*** | - | - |
| Rainfall | -0.002*** | 1.00 | 1.54 | ni | - | - | 0.008*** | 1.01 | 1.30 |
| Temperature | ni | - | - | -0.471*** | 0.62 | 1.86 | ni | - | - |
| Elevation | ni | - | - | ni | - | - | ni | - | - |
| Slope | -0.020** | 0.98 | 1.55 | ns | - | - | ns | - | - |
| Aspect | -0.002*** | 1.00 | 1.05 | ns | - | - | 0.002*** | 1.00 | 1.05 |
| Curvature | 0.039* | 1.04 | 1.49 | ns | - | - | -0.097*** | 0.91 | 1.08 |
| TWI | ns | - | - | ns | - | - | -0.297*** | 0.74 | 1.35 |
| Soil pH | -0.391*** | 0.68 | 2.61 | -0.306* | 0.74 | 3.12 | 0.827*** | 2.29 | 2.05 |
| Soil CEC | 0.054*** | 1.06 | 2.71 | 0.061*** | 1.06 | 3.70 | -0.009*** | 0.99 | 1.80 |
| DIST_road | -0.278*** | 0.76 | 1.22 | -0.528*** | 0.59 | 1.40 | -0.119*** | 0.89 | 1.53 |
| DIST_town | -0.086*** | 0.92 | 1.34 | -0.041* | 0.96 | 1.44 | 0.060*** | 1.06 | 1.72 |
| DIST_river | 0.077** | 1.08 | 1.26 | -0.244*** | 0.78 | 1.48 | 0.217*** | 1.24 | 1.29 |
| POP_09 | ns | - | - | ns | - | - | -0.003*** | 1.00 | 1.13 |
| *N* | 3408 | | | 1618 | | | 3898 | | |
| Pr >LR $\chi^2$ | 0.00 | | | 0.00 | | | 0.00 | | |
| Pr >Pearson $\chi^2$ | 0.13 | | | 0.08 | | | 0.76 | | |
| Nagelkerke $R^2$ | 0.21 | | | 0.49 | | | 0.43 | | |
| AUC | 0.74 | | | 0.86 | | | 0.84 | | |
| Moran's I. | 0.19 | | | 0.13 | | | 0.14 | | |

Notes: Significance level 0 = '***', 0.001 = '**', 0.01 = '*' and 0.05 = '.'; ni = not included due to high correlation with either elevation, temperature or rainfall; ns = not significant; VIF = variance inflation factor; *N* = number of observations; LR = likelihood ratio, and; AUC = area under curve.

## 4.3 The effects of land cover changes on SOC and TN stocks

Linear mixed analyses revealed that soil properties were responsive to the changing landscape. Land cover had a highly significant effect on SOC and TN ($p$<0.0001) in the NF vs. NF2C category (Table 4). SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), and $TN_{st}$ ($p$<0.001) in NF differed from NF2C. In the surface soils (0-15cm), the highest SOC and TN stocks were in NF (71.6 and 7.1 Mg $ha^{-1}$, respectively) and the lowest in NF2C (35.4 and 3.5 Mg $ha^{-1}$). In the subsurface soils (15-30cm), the highest stocks were in NF (55.7 and

5.6 Mg ha$^{-1}$, respectively) and the lowest in PF2C (32.3 and 3.2 Mg ha$^{-1}$). Cultivation of NF reduced both SOC and TN stocks by 51% in the surface and 42% in the subsurface soils (Fig. 10 & 11). This indicates soil degradation upon cultivation of NF. Soil depth also had a highly significant effect on SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0002), TN ($p$<0.0001), and $TN_{st}$ ($p$=0.0002). SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), and $TN_{st}$ ($p$<0.001) between soil depths 0-15cm and 15-30cm were significantly different.

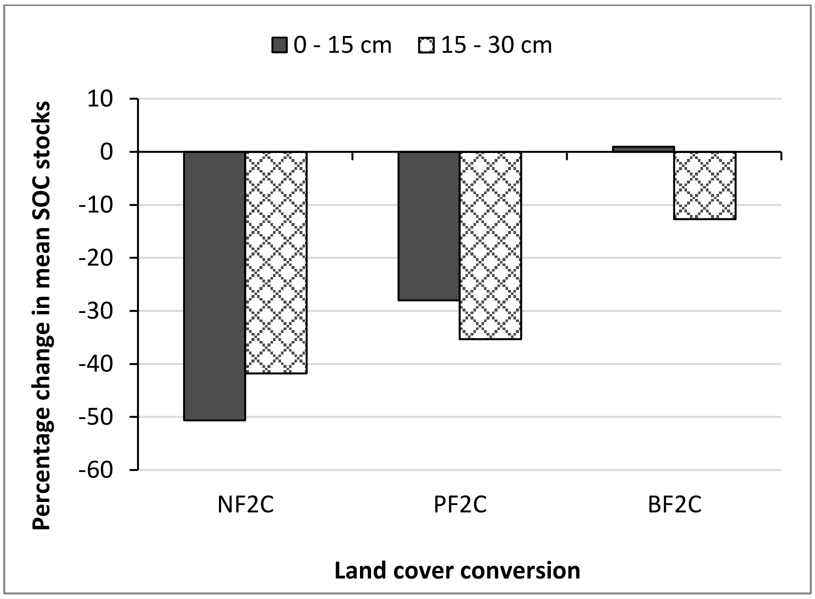



**Fig. 10:** Percentage change in SOC stocks following forest conversions. NF2C = natural forest converted to cropland; PF2C = plantation forest converted to cropland; and, BF2C = bamboo forest converted to cropland. The SOC concentrations and BD data used to estimate SOC stocks are shown in Table 1, Paper III.

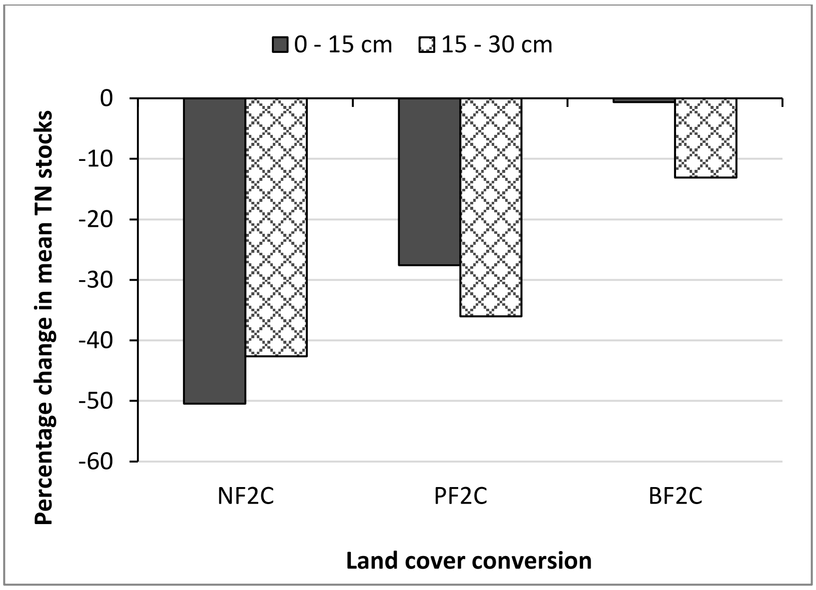



**Fig. 11:** Percentage change in TN stocks following forest conversions. NF2C = natural forest converted to cropland; PF2C = plantation forest converted to cropland; and, BF2C = bamboo forest converted to cropland. The TN concentrations and BD data used to estimate TN stocks are shown in Table 1, Paper III.

Similarly, there was a highly significant land-cover effect on SOC ($p$=0.0001), $SOC_{st}$ ($p$=0.0001), TN ($p$=0.0002), and $TN_{st}$ ($p$<0.0001) in the PF vs. PF2C category. SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), and $TN_{st}$ ($p$<0.001) between PF and PF2C were significantly different. Cultivation of PF reduced both SOC and TN stocks by 28% in the surface and 36% in the subsurface soils. Hence, there is the risk of soil degradation upon cultivation of PF. There was also a highly significant soil-depth effect on SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0038), TN ($p$<0.0001), and $TN_{st}$ ($p$=0.0026). SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0041), TN ($p$<0.0001), and $TN_{st}$ ($p$=0.0026) between soil depths 0-15cm and 15-30cm differed significantly.

**Table 4:** Statistical summary of land-cover and soil-depth effects on SOC, $SOC_{st}$, TN, $TN_{st}$, and BD

| | **SOC** | | **$SOC_{st}$** | | **TN** | | **$TN_{st}$** | | **BD** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **(a) NF vs. NF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 63.22 | <0.0001 | 56.46 | <0.0001 | 66.09 | <0.0001 | 57.44 | <0.0001 | 0.19 | 0.6648 |
| Soil depth | 72.51 | <0.0001 | 18.02 | 0.0002 | 76.39 | <0.0001 | 17.65 | 0.0002 | 32.01 | <0.0001 |
| Land cover × Soil depth | 16.47 | 0.0004 | 8.61 | 0.0066 | 16.17 | 0.0004 | 7.42 | 0.0110 | - | - |
| **(b) PF vs. PF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 22.17 | 0.0001 | 49.89 | <0.0001 | 20.04 | 0.0002 | 44.92 | <0.0001 | 6.32 | 0.0191 |
| Soil depth | 37.75 | <0.0001 | 10.18 | 0.0038 | 40.00 | <0.0001 | 11.15 | 0.0026 | 6.78 | 0.0153 |
| **(c) BF vs. BF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 0.08 | 0.7894 | 0.06 | 0.8217 | 0.12 | 0.7460 | 0.09 | 0.7749 | 0.16 | 0.6992 |
| Soil depth | 7.57 | 0.0284 | 8.87 | 0.0206 | 7.65 | 0.0279 | 8.90 | 0.0204 | 0.62 | 0.4545 |

Note: The degrees of freedom was 1 in all cases; F = *F*-value; P = *P*-value; NF = Natural forest; NF2C = Natural forest converted to cropland; PF= Plantation forest; PF2C = Plantation forest converted to cropland; BF = Bamboo forest; and, BF2C = Bamboo forest converted to cropland

In contrast, land cover had no significant effect on SOC ($p$=0.7894), $SOC_{st}$ ($p$=0.8217), TN ($p$=0.7460), and $TN_{st}$ ($p$=0.7749) in the BF vs. BF2C category. SOC ($p$=0.9825), $SOC_{st}$ ($p$=0.990), TN ($p$=0.9699), and $TN_{st}$ ($p$=0.9794) in BF were similar to BF2C. Cultivation of BF presented mixed results; that is, in the surface soils, SOC stocks increased by 1%, while TN stocks decreased by 0.6%. And in the subsurface soils, SOC and TN stocks reduced by about 13%. The absence of land-cover effect can be attributed to the establishment of croplands within BF less than 10 years ago. The sample sizes for BF and BF2C were also small ($n$=4), which may have not fully captured the variations within these land cover groups. However, soil depth had a significant effect on SOC ($p$=0.0284), $SOC_{st}$ ($p$=0.0206), TN ($p$=0.0279), and $TN_{st}$ ($p$=0.0204). Pairwise comparisons revealed differences in SOC ($p$=0.0157), $SOC_{st}$ ($p$=0.008), TN ($p$=0.0158), and $TN_{st}$ ($p$=0.008) between soil depths 0-15cm and 15-30cm.

The same trends were observed even when the entire topsoil (0-30cm) was considered. The highest stocks of SOC amounting to 127 Mg ha$^{-1}$ were in NF, which reduced by 46.8% when converted to croplands, and the lowest stocks amounting to 101.5 Mg ha$^{-1}$ were in BF, which reduced by 4.4% when converted to croplands. Similarly, the highest stocks of TN were in NF (12.7 Mg ha$^{-1}$), which reduced by 47% when converted to croplands, and the lowest stocks were in BF (10.3 Mg ha$^{-1}$), which reduced by 5.4% when converted to croplands. This coincides with the findings of previous studies in the tropics (Detwiler, 1986; Solomon et al., 2000; Bewketa and Stroosnidjer, 2003; Walker and Desanker, 2004; Lemenih et al., 2005; Enanga et al., 2011). Disruption of the balance between inputs and outputs of C and N in the soil system after forest conversion explains the decrease in SOC and TN stocks. Forest ecosystems have higher net primary productivity than agro-ecosystems; thus, their inputs of detritus to the soils are also higher (Smith, 2008; Eclesia et al., 2012). In the agro-ecosystems, the bulk of biomass is removed from the fields after harvest for use as food or fuel, which hinders accumulation of soil organic matter, and aggravates SOC and TN losses through erosion. In addition, frequent tillage disintegrates the soil aggregates, redistributes crop residues, and alters soil aeration, moisture, and temperature conditions. This accelerates microbial decomposition and oxidation of soil organic matter to $CO_2$, which is then emitted to the atmosphere (Follett, 2001; Murty et al., 2002; Lal, 2004; Powers, 2004; Wiesmeier et al., 2012).

These results highlight the impact of human activities on soils of the area. They have important implications for sustainable management of the croplands, forests (NF, PF, and BF), and soils at different depths in the area. For example, the reduction of SOC and TN concentrations and stocks after forest conversions calls for intervention measures that will enhance the storage of C and N. The decrease of these soil properties as soil depth increase, on the other hand, suggests that the measures should not only focus on enhancing C and N storage in the surface, but also in the subsurface soils. In a broader context, these findings, in addition to those presented in sections 3.1 and 3.2 contribute to the body of scientific literature describing the dynamics, drivers, and impacts of LULCC at the local, regional, and global levels.

### 4.4 The spatially-distributed estimates and patterns of SOC and TN stocks

Multiple linear regression results (Table 5) showed that elevation, silt content, TN concentration, and OLI band 11 had significant effects on SOC stocks and explained 72% of the spatial variability (adjusted $R^2$=0.72). Similarly, elevation, OLI band 11, and SOC concentration had significant effects on TN stocks and explained 71% of the spatial variability (adjusted $R^2$=0.71). Both MLR and GWR model outputs indicated that TN concentration had the largest magnitude of effect on SOC stocks, while SOC concentration had the largest magnitude of effect on TN stocks (Tables 5 & 6). Therefore, soil properties were more important than other environmental factors in controlling the observed patterns of SOC and TN stocks. This coincides with the conclusion made by Vågen and Winowiecki (2013) after mapping SOC to 30cm depth in four contrasting East African landscapes. However, unlike MLR models, GWR models showed that the magnitude of the effects of predictors varied with sampling location. That is, the relationships between the target variables (SOC and TN stocks) and environmental factors were spatially non-stationary.

**Table 5:** Parameter estimates of the MLR models

| | SOC stocks model | | | | | TN stocks model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | Estimate | SE | *t* value | Pr(>\|t\|) | VIF | Estimate | SE | *t* value | Pr(>\|t\|) | VIF |
| Intercept | 143.502 | 50.757 | 2.827 | 0.0053** | - | 16.741 | 5.131 | 3.263 | 0.0013** | - |
| Silt | 0.443 | 0.202 | 2.191 | 0.0298* | 1.531 | - | - | - | - | - |
| Band 11 | -0.003 | 0.001 | -2.360 | 0.0194* | 3.511 | -0.000 | 0.000 | -2.475 | 0.0143* | 3.489 |
| Elevation | -0.022 | 0.009 | -2.503 | 0.0133* | 3.613 | -0.002 | 0.001 | -2.305 | 0.0223* | 3.558 |
| TN | 178.200 | 12.269 | 14.524 | 0.0000*** | 2.471 | - | - | - | - | - |
| SOC | - | - | - | - | - | 1.597 | 0.106 | 15.103 | 0.0000*** | 1.807 |
| Adjusted $R^2$ | 0.72 | | | | | 0.71 | | | | |
| RMSE | 13.07 | | | | | 1.33 | | | | |
| Moran's I | 0.11 | | | | | 0.08 | | | | |

Significance codes: 0 '***' 0.001 '**' 0.01 '*'

**Table 6:** Parameter estimates of the GWR models

| | SOC stocks model | | | | | TN stocks model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | Mean | SD | Min. | Max. | Range | Mean | SD | Min. | Max. | Range |
| Intercept | 129.525 | 31.412 | 50.031 | 199.881 | 149.851 | 14.790 | 4.771 | 8.037 | 26.423 | 18.387 |
| Silt | 0.436 | 0.115 | 0.203 | 0.614 | 0.411 | - | - | - | - | - |
| Band 11 | -0.003 | 0.001 | -0.004 | -0.001 | 0.004 | -0.000 | 0.000 | -0.000 | -0.000 | 0.000 |
| Elevation | -0.021 | 0.007 | -0.041 | -0.004 | 0.037 | -0.002 | 0.001 | -0.005 | -0.001 | 0.004 |
| TN | 177.230 | 23.072 | 142.790 | 238.558 | 95.768 | - | - | - | - | - |
| SOC | - | - | - | - | - | 1.576 | 0.197 | 1.087 | 1.899 | 0.812 |
| Global adjusted $R^2$ | 0.73 | | | | | 0.72 | | | | |
| Global RMSE | 12.86 | | | | | 1.29 | | | | |
| Moran's I | 0.06 | | | | | 0.02 | | | | |

The different prediction maps of SOC and TN stocks produced MLR, MLRK, GWR, and GWRK models displayed similar spatial patterns of SOC and TN stocks. Hence, SOC and TN stocks responded similarly to the environmental factors. There was a general decrease of SOC and TN stocks from west to east (Figs. 12 & 13). The highest estimates of SOC and TN stocks occurred on the western and north-western parts, which according to the environmental data, had higher forest cover, elevations, and SOC and TN concentrations, but lower silt contents and surface temperatures. These *hotspots* were parts of the Logoman, Nessuiet, Kiptunga, and Baraget forests, which had not undergone deforestation. The lowest estimates, on the other hand, occurred on the eastern side where croplands had been established, including Teret, Nessuiet, Kapkembu, Tuiyotich, and Sururu locations. These *coldspots* were areas with higher crop cover, silt contents, and surface temperatures, but lower elevations, and SOC and TN concentrations. In the northern and south-eastern parts where crop cover was also high, the SOC and TN stocks were moderate to high.

The given characteristics of the *hotspots* of SOC and TN stocks on the western and north-western parts, in addition to the highly fertile *Andosols*, favours accumulation of SOC and TN stocks. For instance, the high rainfall and low temperatures associated with higher altitudes increase net primary productivity of the forests and decrease SOC turnover. The lower silt content relative to clay content in the forest soils is also an indication of the presence of organo-complexes, or allophane, imogolite, and ferrihydrite clay minerals, which stabilize organic matter and plant nutrients (Lemenih et al., 2005; Chaplot et al., 2010). The smaller pore spaces of clay particles also promote aggregation and physical protection of SOC. In contrast, the characteristics of the *coldspots* of SOC and TN stocks on the eastern side are unfavourable for accumulation of SOC and TN stocks. For example, the higher crop cover is attributed to the conversion of forests to croplands, which began in the mid-1990s. In these croplands, biomass removal after harvesting, erosive processes, and frequent tillage explain the lower SOC and TN stocks (Murty et al., 2002; Smith 2008; Eclesia et al., 2012; Wiesmeier et al., 2012). Thus, the *coldspots* of SOC and TN stocks also highlight the human-induced soil degradation, and sources of C and N emissions.

**Table 7:** Soil organic carbon and nitrogen stocks under different land cover types

| Land cover | Area | SOC stocks | | | | TN stocks | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Min. | Max. | Mean | Total | Min. | Max. | Mean | Total |
| | Ha | Mg $ha^{-1}$ | | | Tg | Mg $ha^{-1}$ | | | Tg |
| Forests | 32228.4 | 75.5 | 142.9 | 110.4 | 3.78 | 7.5 | 15.3 | 11.1 | 0.38 |
| Grasslands | 5509.4 | 66.7 | 129.8 | 103.5 | 0.57 | 6.7 | 12.6 | 10.4 | 0.06 |
| Croplands | 25828.1 | 62.9 | 126.9 | 95.2 | 2.46 | 6.5 | 12.2 | 9.6 | 0.25 |
| **Total** | **65565.9** | | | | **6.81** | | | | **0.69** |

Based on GWR method, which showed the lowest prediction error indices, forests stored the highest amounts of SOC and TN (i.e., 3.78 Tg C and 0.38 Tg N) followed by croplands (i.e., 2.46 Tg C and 0.25 Tg N) and grasslands (i.e., 0.57 Tg C and 0.06 Tg N) (1 Tg = $10^{12}$ g = 1 million tons) (Table 7). This is because forests covered the largest area (32,228 ha), while grasslands covered the smallest area (5,509 ha). Overall, the Eastern Mau Forest Reserve stored about 6.81 Tg and 0.69 Tg of SOC and TN, respectively. This accounts for 0.36% of the organic C stored in Kenyan soils to 30cm depth based on Batjes'

(2004) estimates of SOC stocks for Kenya. He further reported that *Andosols* of the humid and semi-humid regions in Kenya stored an average of 9.1 Kg C $m^{-2}$ (91 Mg C $ha^{-1}$) to 30cm depth, which slightly differs with the present findings (i.e., 10.3 Kg C $m^{-2}$, or 102.7 Mg C $ha^{-1}$). This can be attributed to the different spatial and temporal properties of data used in the two studies.

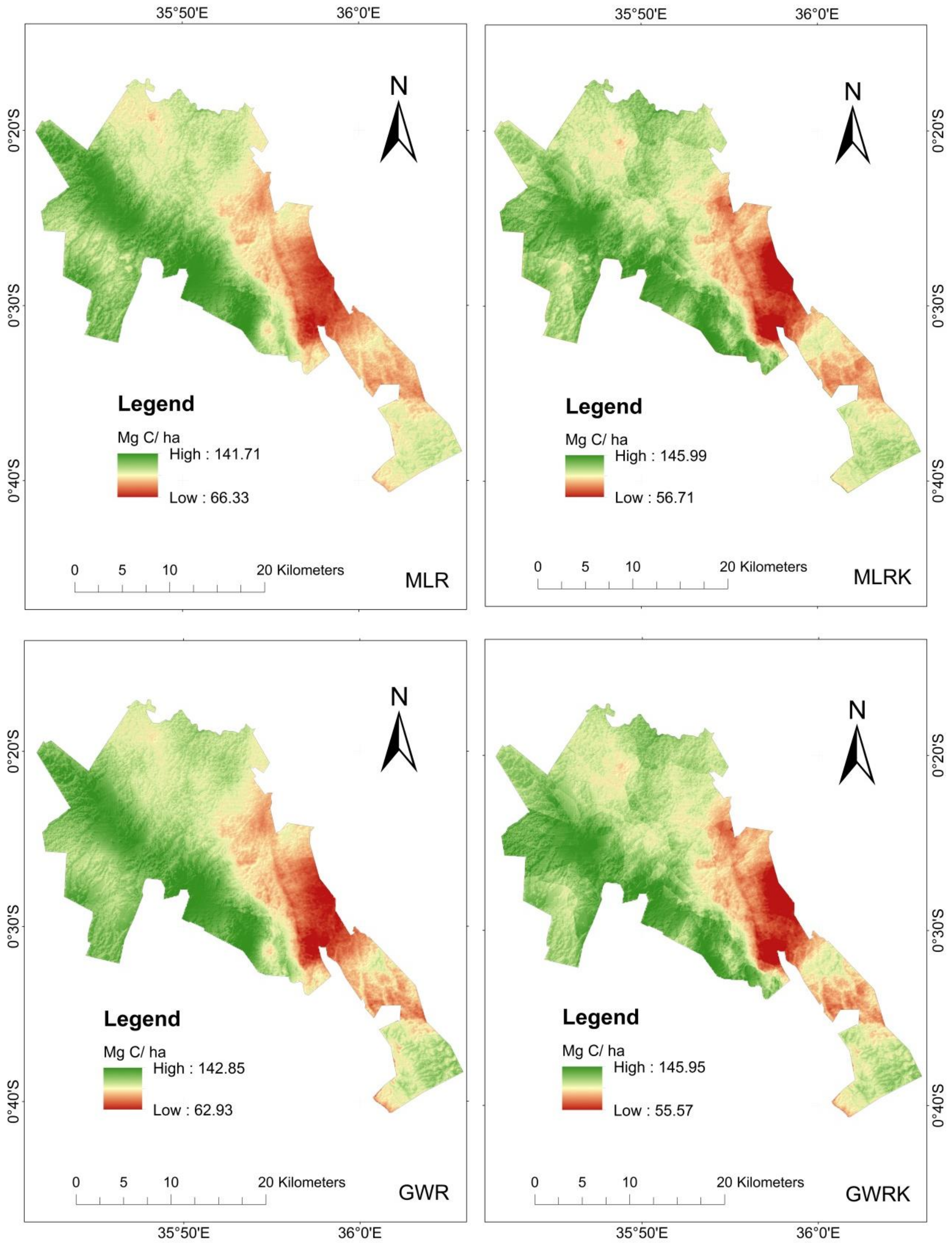



**Fig. 12:** Maps showing the spatial patterns of SOC stocks estimated using MLR, MLRK, GWR, and GWRK

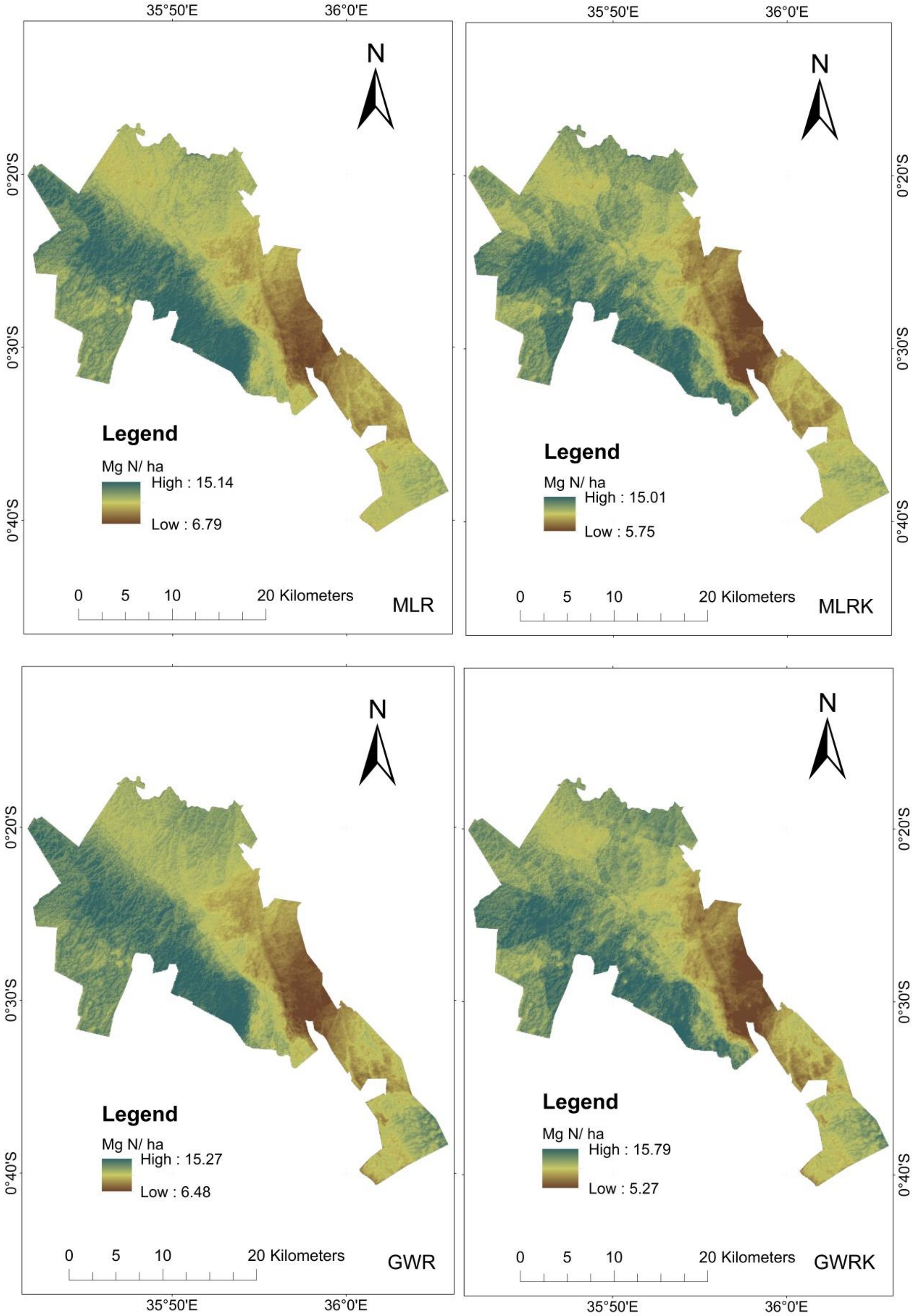



**Fig. 13:** Maps showing the spatial patterns of TN stocks estimated using MLR, MLRK, GWR, and GWRK

The fine-scale, spatially-exhaustive soil information produced for the Eastern Mau Forest Reserve for the first time is a fundamental step towards improved monitoring of soils, and informed formulation of spatially-targeted and sustainable land management policies.

Based on this information, for instance, the western and north-western parts would benefit from policies that promote conservation, while the eastern part from those that support accumulation of SOC and TN stocks. The information may also be useful for parameterization of other environmental models, and for simulation of impacts due to land use changes. In addition, the integrated approach of field sampling, GIS, remote sensing, and statistical analysis provides a cost-effective framework for deriving knowledge of soil processes, as well as multi-purpose soil information in other data-poor environments in Eastern Africa. Lastly, the resultant soil database can support the activities of other research programs; for example, soil carbon sequestration, REDD+, and ISRIC-World Soil Information programs, etc.

### 4.5 Limitations of the research

Analysis of the dynamics, drivers and impacts of land cover changes using remote sensing and GIS approach is a complex process with many potential factors that may introduce uncertainties to the outputs. Several sources of uncertainties, which also present opportunities for further research, were identified in this research. To commence with, the created land cover maps were not perfect as shown by the overall measures of accuracy (Paper I). Classification errors occurred mainly because of spectral confusion, which is common when classifying heterogeneous landscapes. The subsequent estimations of the magnitudes and rates of land cover changes, as well as the SOC and TN stocks under different land cover types were based on these land cover maps. Hence, the inherent classification errors may have influenced these estimates. Further, the reference data used for validating the land cover maps, and the auxiliary spatial data used for modelling (Paper II) were sourced from different databases in various formats and scales; thus, the data quality was not uniform. Errors in calibration and validation data can also bias model results. The uncertainties attached to the models developed in Paper IV are diverse. The first obvious source is errors from the field measurements and laboratory analysis. Besides, the soil properties used as predictors were products of interpolation by ordinary kriging. Hence, the interpolation errors may have been propagated into the subsequent prediction of SOC and TN stocks. Poor coverage of samples in some areas (e.g., thick impenetrable bamboo forests) may have also affected prediction accuracies around such areas. Lastly, some soil-forming factors (e.g., parent material and age) were not accounted for thanks to the absence of suitable data. The same applies to the land cover change models where some important explanatory factors were not captured owing to lack of spatial data and inability to quantify others (e.g., politics). Inclusion of these factors, if significant, may improve the predictive power of the models in future. Finally, like in most empirical studies, the models developed are site-specific and may be inapplicable outside the study area.

## 5.0 CONCLUSIONS, RECOMMENDATIONS AND OUTLOOK

In a nutshell, this research has demonstrated an integrated approach of field sampling, satellite remote sensing, GIS, and (spatial) statistical approaches to study the dynamics, drivers, and impacts of land cover changes in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve, Kenya. The research outputs fill the existing biophysical information gap and, thereby, contribute knowledge to support policy formulation for sustainable land management and climate change mitigation in the area. In view of the research questions and findings, the following specific conclusions are drawn:

### 5.1 Conclusions

a) The Lake Nakuru drainage basin, including the Eastern Mau Forest Reserve, transformed from a natural to human-dominated landscape over the 38-year study period. The forests-shrublands, which were dominant between 1973 and 2000, had been surpassed by croplands in 2011.

b) Forests-shrublands, grasslands, and croplands had higher magnitudes of change than built-up lands, bare lands, and water bodies. But, built-up lands had the highest annual rates of change.

c) The major land cover change processes were conversion of forests-shrublands and grasslands to croplands and built-up lands. Forests-shrublands were the biggest losers followed by grasslands, while croplands were the biggest gainers followed by built-up lands.

d) The spatial distribution of the land cover change *hotspots* varied with time. For example, the *hotspots* of forest-shrubland conversion occurred in the middle regions and northern side of Lake Nakuru between 1973 and 1985, but were concentrated on the western side between 1983 and 2000.

e) Climatic, topographic, soil, demographic, and accessibility factors (e.g., rainfall, proximity to road, town and river) determined the major land cover change processes (forest-shrubland conversion, grassland conversion, and agricultural expansion). These factors varied with the nature of land cover change process and time.

f) Soil properties, particularly SOC and TN concentrations and stocks, declined in response to the land cover changes and human impact in the Eastern Mau Forest Reserve. The concentrations and stocks of SOC and TN under the forests and cropland establishments were significantly different. Thus, the transformation from a natural to human-dominated landscape is increasing the risk of soil degradation, and restricting the ecosystem's capacity for SOC and TN storage.

g) The forests located on the western and north-western parts (i.e., Logoman, Nessuiet, Kiptunga, and Baraget forests) were the *hotspots*, while the croplands established on the eastern part (i.e., Teret, Nessuiet, Kapkembu, Tuiyotich, and Sururu locations) were the *coldspots* of SOC and TN stocks in the Eastern Mau Forest Reserve.

h) Climatic, edaphic, and topographic factors controlled the observed spatial patterns of SOC and TN stocks in the Eastern Mau Forest Reserve; however, soil properties, particularly TN and SOC concentrations, were the most important determinants.

## 5.2 Recommendations

Based on the research findings, formulation and implementation of spatially-targeted and time-specific policies that aim to restore and conserve the natural ecosystems, as well as enhance agricultural productivity for environmental sustainability and socio-economic welfare is recommended. For instance, the western and north-western parts where forests dominate need policies that will promote conservation, while the eastern part where croplands dominate requires those that will ensure both soil and environmental quality.

Furthermore, selection and adoption of best management practices (BMPs), which increase C and N inputs, and decrease decomposition rate, will be beneficial for the croplands. Agro-forestry systems where fast-growing, highly productive, deep-rooted, and N-fixing tree species are planted will, particularly, be useful. For the forest soils, long-term storage of C and N will be achieved through improved management and protection of the forests from further deforestation and degradation.

## 5.3 Outlook

The achievements and limitations of this research provide a number of opportunities for further research; hence, more LULCC research activities are envisaged in the Lake Nakuru drainage basin and Eastern Mau Forest Reserve. The future research directions include:

a) Re-evaluating and refining the land cover change models to explicitly account for other important geophysical and socio-economic drivers of land cover changes that were not considered at present, and to allow the projection of future patterns of land cover changes. This will need incorporation of additional detailed input environmental data as they become available with time.

b) Exploring the variations of the relationships among the drivers of land cover change processes within the local space. This will require application of local regression techniques, such as geographically weighted logistic regression, etc.

c) Refining the spatial predictions and models of SOC and TN stocks by: (i) incorporating other soil-forming factors (e.g., parent material) that may be significant in the modelling process as more data becomes available with time, (ii) analysing the sensitivity of model parameters to variations in the quality of multi-source data. This will shed some light on error propagation in the models, and (iii) application and evaluation of machine learning (data mining) algorithms, such as support vector regression, random forests, and artificial neural networks, etc.

d) Linking the generated research outputs with a variety of other environmental data in a spatially-explicit framework for integrated assessment of the impacts of land cover changes on ecosystem services, and formulation of holistic land management strategies. For example, within such a framework, the land cover maps and land cover change modelling results can be linked with species distribution data to identify, or predict the *hotspots* of biodiversity losses. This will, ultimately, put the Lake Nakuru drainage basin and Eastern Mau Forest Reserve on the road to integrated management of land resources.

## REFERENCES


Agresti, A. (2007). *An introduction to categorical data analysis*. New Jersey: John Wiley & Sons, Inc.

Aguiar, A.P.D., Câmara, G., Escada, M.I.S. (2007). Spatial statistical analysis of land use determinants in the Brazilian Amazonia: exploring intra-regional heterogeneity. *Ecological Modelling* 209, 169–188.

Amare, T., Hergarten, C., Hurni, H., Wolfgramm, B., Yitaferu, B., Selassie, Y.G. (2013). Prediction of soil organic carbon for Ethiopian highlands using soil spectroscopy. *ISRN Soil Science* 2013, 720589 (11 pp), http://dx.doi.org/10.1155/2013/720589.

Anderson, J.R., Harvey, E.E., Roach, J.T., Witmer, R.E. (1976). A land use and land cover classification system for use with remotely sensed data. US Geological Survey professional paper 964. United States Government Printing Office, Washington, D.C.

Ballabio, C. (2009). Spatial prediction of soil properties in temperate mountain regions using support vector regression. *Geoderma* 151, 338–350.

Baldyga, T., Miller, S.N., Driese, K.L., Gichaba, C.M. (2007). Assessing land cover change in Kenya's Mau Forest region using remotely-sensed data. *African Journal of Ecology* 46, 46–54.

Batjes, N.H. (2004). Soil carbon stocks and projected changes according to land use and management: a case study for Kenya. *Soil Use and Management* 20, 350-356.

Bewketa, W., Stroosnijder, L. (2003). Effects of agro-ecological land use succession on soil properties in Chemoga watershed, Blue Nile basin, Ethiopia. *Geoderma* 111, 85-98.

Blake, G.R. (1965). Bulk density. In: Black, C.A. (ed.), *Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling*. American society of agronomy, Inc., Madison, Wisconsin, USA.

Braimoh, A. K., Vlek, P. L. G. (2004). The impact of land-cover change on soil properties in northern Ghana. *Land Degradation and Development* 15, 65-74.

Braimoh, A.K., Vlek, P.L.G. (2005). Land cover change trajectories in northern Ghana. *Environmental Management* 36 (3), 356–373.

Bremner, J.M., Mulvaney, C.S. (1982). Nitrogen - total. In: Page, A.L. (ed.), *Methods of soil analysis, Part 2. Chemical and microbiological properties, second edition*. American society of agronomy, Inc., Madison, Wisconsin, USA.

Briassoulis, H. (2000). Analysis of land use change: theoretical and modeling approaches. [Online]. Available: http://www.rri.wvu.edu/WebBook/Briassoulis/contents.htm. [Accessed 2014, March 9].

Brown, S., Lugo, A.E. (1990). Effects of forest clearing and succession on the carbon and nitrogen content of soils in Puerto Rico and US Virgin Islands. *Plant and Soil* 124, 53-64.

Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A. (2014). Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty. *Geoderma* 213, 46-56.

Campbell, J.B. (2002). *Introduction to Remote Sensing*. London: Taylor & Francis.

Chaplot, V., Bouahom, B., Valentin, C. (2010). Soil organic carbon stocks in Laos: spatial variations and controlling factors. *Global Change Biology* 16, 1380-1393.

Chavez, P.S. (1996). Image-based atmospheric corrections: revisited and improved. *Photogrammetric Engineering and Remote Sensing* 62(9), 1025-1036.

Chen, G., Hay, G.J., Carvalho, L.M.T., Wulder, M.A. (2012). Object-based change detection. *International Journal of Remote Sensing* 33 (14), 4434-4457.

Chen, J., Chen, J., Tan, M., Gong, Z. (2002). Soil degradation: a global problem endangering sustainable development. *Journal of Geographical Sciences* 12(2), 243-252.

Chomitz, K.M., Gray, D.A. (1996). Roads, land use and deforestation: a spatial model applied to Belize. *The World Economic Review* 10 (3), 487-512.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing* 25(9): 1565-1596.

Daia, F., Zhou, Q., Lv, Z., Wang, X., Liu, G. (2014). Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecological Indicators* 45, 184–194.

Day, P.R. (1965). Particle fractionation and particle size analysis. In: Black, C.A. (ed.), *Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling*. American society of agronomy, Inc., Madison, Wisconsin, USA.

Demessie, A., Singh, B.R., Lal, R. (2013). Soil carbon and nitrogen stocks under chronosequence of farm and traditional agro-forestry uses in Gambo district, southern Ethiopia. *Nutr Cycl Agroecosys* 95, 365-375.

Detwiler, R.P. (1986). Land use change and the global carbon cycle: The role of tropical soils. *Biogeochemistry* 2, 67-93.

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* 204-205, 31-42.

Don, A., Schumacher, J., Freibauer, A. (2011). Impact of tropical land-use change on soil organic carbon stocks - A meta-analysis, *Global Change Biology* 17, 1658–1670.

Dorji, T., Odeh, I.O.A., Field, D.J., Baillie, I.C. (2014). Digital soil mapping of soil organic carbon stocks under different land use and land cover types in montane ecosystems, Eastern Himalayas. *Forest Ecology and Management* 318, 91-102.

Eaton, J.M., McGoff, N.M., Byrne, K.A., Leahy, P., Kiely, G. (2008). Land cover change and soil organic carbon stocks in the Republic of Ireland 1851–2000. *Climatic Change* 91, 317–334.

Eclesia, R.P., Jobbagy, E.G., Jackson, R.B., Biganzoli, F., Piñeiro, G. (2012). Shifts in soil organic carbon for plantation and pasture establishment in native forests and grasslands of South America. *Global Change Biology* 18, 3237–3251.

Enanga, E.M., Shivoga, W.A., Maina-Gichaba, C., Creed, I.F. (2011). Observing changes in riparian buffer strip soil properties related to land use activities in the River Njoro watershed, Kenya. *Water Air Soil Pollut* 218, 587–601.

Evrendilek, F., Celik, I., Kilic, S. (2004). Changes in soil organic carbon and other physical soil properties along adjacent Mediterranean forest, grassland, and cropland ecosystems in Turkey. *Journal of Arid Environments* 59, 743–752.

Fan, F., Weng, Q., Wang, Y. (2007). Land use and land cover change in Guangzhou, China from 1998 to 2003 based on Landsat TM /ETM+ imagery. *Sensors* 7, 1323-1342

Follett, R.F. (2001). Soil management concepts and carbon sequestration in cropland soils. *Soil and Tillage Research* 61, 77-92.

Foody, G.M. (2002) Status of land cover classification accuracy assessment. *Remote Sensing of Environment* 80, 185– 201.

Fortier, J., Rogan, J., Woodcock, C.E., Runfola, D.M. (2011). Utilizing temporally invariant calibration sites to classify multiple dates and types of satellite imagery. *Photogrammetric Engineering and Remote Sensing* 77(2), 181-189.

Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. *Geographically weighted regression: The analysis of spatially varying relationships*. John Wiley & Sons Inc., England.

Geist, H.J., Lambin, E.F. (2001) What drives tropical deforestation? A meta-analysis of proximate and underlying causes of deforestation based on subnational case study evidence. (LUCC Report Series; 4). [Online]. Available: https://www.pik-potsdam.de/members/cramer/teaching/0607/Geist_2001_LUCC_Report.pdf . [Accessed 2014, March 9].

Geist, H.J., Lambin, E.F. (2002). Proximate causes and underlying driving forces of tropical deforestation. *BioScience* 52 (2), 143-150.

Girmay, G., Singh, B.R. (2012). Changes in soil organic carbon stocks and soil quality: Land use system effects in northern Ethiopia. *Acta Agriculturae Scandinavica, Section B &#x2013, Soil & Plant Science* 62 (6), 519-530.

Government of Kenya (2009). Report of the prime minister's task force on the conservation of the Mau forest complex. [Online]. Available: http://www.maurestoration.go.ke/. [Accessed 2012, December 20].

Grimm, R., Behrens, T., Märker, 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.

Grunwald, S. (2009). Multi-criteria characterization of recent soil mapping and modelling approaches. *Geoderma* 152, 195-207.

Heistermann, M., Müller, C., Ronneberger, K. (2006). Land in sight? Achievements, deficits and potentials of continental to global scale land use modelling. *Agriculture, Ecosystems and Environment* 114, 141–158.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G. (2007). About regression-kriging: From equations to case studies. *Computers & Geosciences* 33, 1301-1315.

Hengl, T., Heuvelink, G.B.M., Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma* 120, 75-93.

Hesslerová, P., Pokorný, J. (2010). Effect of Mau forest clear-cut on temperature distribution and hydrology of catchment of Lakes Nakuru and Naivasha: preliminary study. In: Vymazal, J. (Ed.), *Water and Nutrient Management in Natural and Constructed Wetlands*. Springer, Netherlands, pp. 263.

Holling, C.S. (2001). Understanding the complexity of economic, ecological, and social systems. *Ecosystems* 4, 390-405.

Houghton, R.A., Goodale, C.L. (2004). Effects of land-use change on the carbon balance of terrestrial ecosystems. *Ecosystems and Land Use Change Geophysical Monograph*, Series 153.

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D. (2013). Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing* 80, 91-106.

Jaetzold, R., Schmidt, H., Hornetz, B., Shisanya, C. (2010). Farm management handbook of Kenya, Vol. II. Natural conditions and farm management information, second edition, Part B Central Kenya, Subpart B1a Southern Rift Valley Province. Ministry of Agriculture, Kenya and German Agency for Technical Cooperation (GTZ), Nairobi.

Jafarian, Z., Kavian, A. (2013). Effects of land use change on soil organic carbon and nitrogen. *Communications in Soil Science and Plant Analysis* 44 (1-4), 339-346.

Jenny, H. (1941). *Factors of soil formation - a system of quantitative pedology*. New York: McGraw-Hill.

Karunaratne, S.B., Bishop, T.F.A., Baldock, J.A., Odeh, 1.O.A. (2014). Catchment scale mapping of measureable soil organic carbon fractions. *Geoderma* 219-220, 14-23.

Kibichii, S., Shivoga, W.A., Muchiri, M., Miller, S.N. (2007). Macro-invertebrate assemblages along a land-use gradient in the upper River Njoro watershed of Lake Nakuru drainage basin, Kenya. *Lakes & Reservoirs: Research and Management* 12, 107–117.

Krupnik, T.J. (2004). Linking farmer, forest and watershed: agricultural systems and natural resources management along the upper Njoro River, Kenya. University of California International and Area Studies Digital Collection, Article #3. [Online]. Available: http://repositories.cdlib.org/uciaspubs/articles/3 [Accessed 2012, December 27].

Kumar, S., Lal, R. (2011). Mapping the organic carbon stocks of surface soils using local spatial interpolator. *Journal of Environmental Monitoring* 13, 3128-3135.

Kumar, S., Lal, R., Liu, D. (2012). A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma* 189–190, 627-634.

Kumar, S., Lal, R., Liu, D. (2013). Estimating the spatial distribution of organic carbon density for the soils of Ohio, USA. *Journal of Geographical Sciences*, 23(2), 280-296.

Lacoste, M., Minasny, B., McBratney, A., Michot, D. (2014). High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. *Geoderma* 213, 296-311.

Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma* 123, 1-22.

Lal, R. (2008). Carbon sequestration. *Phil. Trans. R. Soc. B* 363, 815–830.

Lambin, E.F. (1997). Modelling and monitoring land-cover change processes in tropical regions. *Progress in Physical Geography* 21, 375–393.

Lambin, E.F., Geist, H.J., Lepers, E. (2003). Dynamics of land use and land cover changes in tropical regions. *Annual Reviews of Environmental Resources* 28, 205-241.

Lambin, E.F., Rounsevell, M.D.A., Geist, H.J. (2000). Are agricultural land use models able to predict changes in land-use intensity? *Agriculture, Ecosystems and Environment* 82, 321–331.

Lemenih, M., Karltun, E., Olsson, M. (2005). Assessing soil chemical and physical property responses to deforestation and subsequent cultivation in smallholders farming system in Ethiopia. *Agriculture, Ecosystems and Environment* 105, 373–386.

Li, Q., Yue, T., Wang, C., Zhang, W., Yu, Y., Li, B., Yang, J., Bai, G. (2013). Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach. *Catena* 104, 210-218.

Lillesand, T.M., Kiefer, R.W., Chipman, J.W. (2008). *Remote sensing and image interpretation, sixth edition.* John Wiley and sons, USA.

Lu, D., Li, G., Moran, E. (2014). Current situation and needs of change detection techniques. *International Journal of Image and Data Fusion* 5(1), 13-38.

Lu, D., Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* 28(5), 823-870.

Lu, D., Mausel, P., Brondizio, E., Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing* 25(12), 2365-2401.

Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M. (2009). Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138-152.

Martin, M.P., Orton, T.G., Lacarce, E., Meersmans, J., Saby, N.PA., Paroissien, J.B., Jolivet, C., Boulonne, L., Arrouays, D. (2014). Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale. *Geoderma*, http://dx.doi.org/10.1016/j.geoderma.2014.01.005.

Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., Arrouays, D. (2011). Spatial distribution of soil organic carbon stocks in France. *Biogeosciences* 8, 1053-1065.

Mas, J. F. (1999). Monitoring land-cover changes: a comparison of change detection techniques. *International Journal of Remote Sensing* 20(1), 139 – 152.

McBratney, A.B., Santos, M.LM., Minasny, B. (2003). On digital soil mapping. *Geoderma* 117, 3-52.

McCall, G.J.H. (1967). Geology of the Nakuru-Thomson’s falls-Lake Hannington area: degree sheet No. 35, S.W. Quarter and 43 N.W. Quarter, Report No. 78. Government Printer, Nairobi.

McKenzie, N.J., Ryan, P.J. (1999). Spatial prediction of soil properties using environmental correlation. *Geoderma* 89, 67-94.

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, 1-13.

Millennium Ecosystem Assessment (2005). *Ecosystems and human well-being: Synthesis*. Island Press, Washington, DC.

Mishra, U., Lal, R., Liu, D., van Meirvenne, M. (2010). Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil Science Society of America Journal* 74, 906–914.

Montgomery, D.C, Peck, E.A., Vining, G.G. (2006a). *Introduction to linear regression analysis, fourth edition*. New Jersey: John Wiley & Sons, Inc.

Montgomery, D.C. (2006b). *Design and analysis of experiments, seventh edition*. New Jersey: John Wiley & sons, Inc.

Müller, D., Mburu, J. (2009). Forecasting hotspots of forest clearing in Kakamega forest, western Kenya. *Forest Ecology and Management* 257, 968–977.

Muñoz-Rojas, M., Jordán, A, Zavala, L.M, de la Rosa, D., Abd-Elmabod, S.K., Anaya-Romero, M. (2012). Impact of land use and land cover changes on organic carbon stocks in Mediterranean soils (1956–2007). *Land Degradation and Development*, doi: 10.1002/ldr.2194.

Murty, D., Kirschbaum, M.F., McMurtrie, R.E., McGilvray, H. (2002). Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology* 8, 105-123.

Mwetu, K.K., Mutua, B.M., Kundu, P.M., Furst, J., Loiskandl, W. (2009). Effects of deforestation and climate variability on river discharge in the Njoro River catchment, Kenya, Proceedings of the Sumawa Mau forest complex conference, Egerton University, 27-29$^{th}$ April 2009.

Nelson, D.W., Sommers, L.E. (1982). Total carbon, organic carbon and organic matter. In: Page, A.L. (ed.), *Methods of soil analysis, Part 2, Chemical and microbiological properties, second edition*. American society of agronomy, Inc., Madison, Wisconsin, USA.

Odada, E.O., Raini, J., Ndetei, R. (2006). Lake Nakuru: experience and lessons learned brief. [Online]. Available : http://www.worldlakes.org/uploads/18_Lake_Nakuru_27February2006.pdf [Accessed 2014, January 10].

Okalebo, J.R., Gathna, K.W., Woomer, P.L. (2002). Laboratory methods for soil and plant analysis: A working manual, second edition. Tropical Soil Biology and Fertility Programme, Nairobi.

Osher, L.J., Matson, P.A., Amundson, R. (2003). Effect of land use change on soil carbon in Hawaii. *Biogeochemistry* 65, 213–232.

Osman, K.T. (2014). *Soil degradation, conservation and remediation*. Springer, Netherlands.

Overmars, K.P., Verburg, P.H. (2005). Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science* 19 (2), 125–152.

Powers, J.S. (2004) Changes in soil carbon and nitrogen after contrasting land-use transitions in north-eastern Costa Rica. *Ecosystems* 7, 134–146.

Raini, J. A. (2009). Impact of land use changes on water resources and biodiversity of Lake Nakuru catchment basin, Kenya. *African Journal of Ecology* 47(1), 39-45.

Rogan, J., Chen, D. (2004). Remote sensing technology for mapping and monitoring land-cover and land-use change. *Progress in Planning* 61, 301–325.

Scull, P., Franklin, J., Chadwick, O.A., McArthur, D. (2003). Predictive soil mapping: a review. *Progress in Physical Geography* 27, 2, 171–197.

Shivoga, W. A., Muchiri, M., Kibichii, S., Odanga, J., Miller, S. N., Baldyga, T. J. (2007). Influences of land use/ cover on water quality in the upper land middle reaches of river Njoro, Kenya. *Lakes & Reservoirs: Research and Management* 12, 97-105.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing* 10(6), 989 – 1003.

Smith, P. (2004). Soils as carbon sinks: the global context. *Soil Use and Management* 20, 212-218.

Smith, P. (2008). Land use change and soil organic carbon dynamics. *Nutr Cycl Agroecosyst* 81, 169–178.

Solomon, D., Lehmann, J., Zech, W. (2000). Land use effects on soil organic matter properties of chromic luvisols in semi-arid northern Tanzania: Carbon, nitrogen, lignin and carbohydrates. *Agriculture, Ecosystems and Environment* 78, 203–213.

Szymanowski, M., Kryza, M. (2012). Local regression models for spatial interpolation of urban heat island–an example from Wrocław, SW Poland. *Theor Appl Climatol* 108, 53-71.

Turner, B. L. II, Skole, D., Moss, R. (eds.) (1993). *Relating land use and global land-cover change: A proposal for an IGBP-HDP core project.* Report from the IGBP-HDP Working Group on Land-Use/Land-Cover Change. Joint publication of the International Geosphere-Biosphere Programme (Report No. 24) and the Human Dimensions of Global Environmental Change Programme (Report No. 5). Stockholm: Royal Swedish Academy of Sciences.

UNEP (2009). Kenya: Atlas of our changing environment. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP). [Online]. Available: http://www.unep.org/dewa/africa/kenyaatlas/. [Accessed 2013, August 28].

Vågen, T.G., Winowiecki, L.A. (2013). Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters* 8, 015011 (9pp), doi:10.1088/1748-9326/8/1/015011.

Verburg, P.H., Kok, K., Pontius, R.G., Veldkamp, A. (2006). Modelling land-use and land-cover change. In: Lambin**,** E.F., Geist**,** H.J. (eds.), *Land use and land cover change*: *local processes and global impacts*. Springer, Heidelberg, pp. 117-135.

Walker, S.M., Desanker, P.V. (2004). The impact of land use on soil carbon in Miombo woodlands of Malawi. *Forest Ecology and Management* 203, 345–360.

Wanjogu, S.N, Kibe, J.M., Wagate, P.N., Mwangi, J.A. (2010). Soil characteristics and land degradation in Lake Nakuru catchment. Kenya Soil Survey, Nairobi.

Wiesmeier, M., Spörlein, P., Geuß, U., Hangen, E., Haug, S., Reischl, A., Schilling, B., von Lützow, M., Kögel-Knabner, I. (2012). Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biology* 18, 2233–2245.

Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J., Zhu, B. (2008). Storage, patterns and controls of soil organic carbon in the Tibetan grasslands, *Global Change Biology*, 14: 1592-1599.

Yimer, F., Ledin, S., Abdelkadir, A. (2007). Changes in soil organic carbon and total nitrogen contents in three adjacent land use types in the Bale Mountains, south-eastern highlands of Ethiopia. *Forest Ecology and Management* 242, 337–342.

Zhang, C., Tang, Y., Xu, X., Kiely, G. (2011). Towards spatial geochemical modelling: Use of geographically weighted regression for mapping soil organic carbon contents in Ireland. *Applied Geochemistry* 26, 1239-1248.

## ERRATA

**1.** An incorrect entry appears in column 2 (Area, km$^2$) of Table 4 in **Paper I**, which is the same as Table 1 on page 20 of the introductory part of the thesis. The correct entry should be as highlighted in red colour in the table below. The sum of the area under the six land cover types for each year (i.e., 1973, 1985, 2000, and 2011) should, ideally, be the same. However, the slight differences can be attributed to the effects of partitioning the image subsets for each year into several smaller segments for separate classification.

Areal extent, magnitude and the annual average rates of land cover changes

| | Area (km$^2$) | | | | Magnitude of change (km$^2$) | | | | Average rate of change p.a (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Land cover class | 1973 | 1985 | 2000 | 2011 | 73 - 85 | 85 - 00 | 00 - 11 | 73 - 11 | 73 - 85 | 85 - 00 | 00 - 11 | 73 - 11 |
| Forests-shrublands | 1067 | 893 | 797 | 639 | -174 | -96 | -158 | -428 | -1 | -1 | -2 | -1 |
| Grasslands | 589 | 531 | 421 | 331 | -58 | -110 | -90 | -258 | -1 | -1 | -2 | -1 |
| Croplands | 293 | 521 | 714 | 953 | 228 | 193 | 239 | 660 | 6 | 2 | 3 | 6 |
| Built-up lands | 4 | 5 | 18 | 28 | 1 | 13 | 10 | 24 | 2 | 17 | 5 | 16 |
| Bare lands | 2 | 7 | 7 | 2 | 3 | 0 | -5 | -2 | 6 | 0 | -6 | -1 |
| Water bodies | 42 | 40 | 40 | 43 | -2 | 0 | 3 | 1 | 0 | 0 | 1 | 0 |

**2.** Incorrect entries also appear in columns 10 ($SOC_{st}$, Mg ha$^{-1}$, 0-30 cm depth) and 12 ($TN_{st}$, Mg ha$^{-1}$, 0-30 cm depth) of Table 3 in **Paper III**. The correct entries should be as highlighted in red colour in the table below. The values in column 8 ($TN_{st}$, Mg ha$^{-1}$, 15-30 cm depth) have also been rounded off to one decimal place like in the other columns.

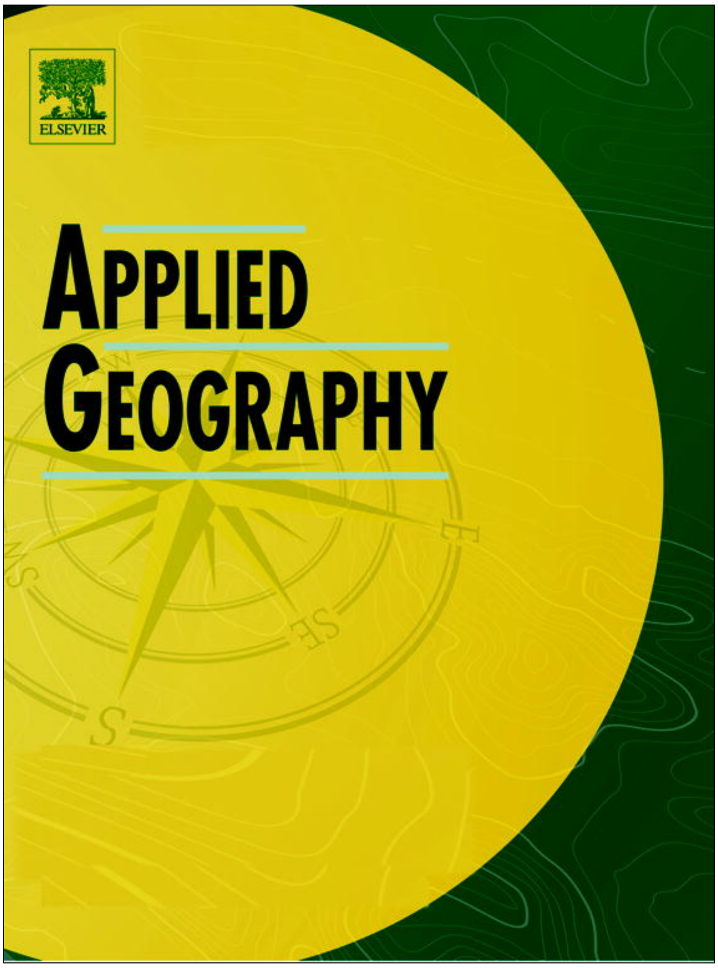
Soil organic carbon and total nitrogen stocks under different land cover types and soil depths (Means and standard deviations)

| | Soil depth (0 - 15 cm) | | | | Soil depth (15 - 30 cm) | | | | Soil depth (0 - 30 cm) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Land cover | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change |
| NF | 71.6±17.9 | | 7.1±1.8 | | 55.7±11.3 | | 5.6±1.2 | | 127.3±26.5 | | 12.7±2.6 | |
| NF2C | 35.4±10.8 | -50.6 | 3.5±1.1 | -50.4 | 32.5±7.4 | -41.8 | 3.2±0.7 | -42.7 | 67.9±15.5 | -46.8 | 6.7±1.6 | -47.0 |
| PF | 54.6±3.9 | | 5.4±0.4 | | 49.9±7.0 | | 4.9±0.7 | | 104.5±8.9 | | 10.3±1.0 | |
| PF2C | 39.3±11.2 | -28.0 | 3.9±1.1 | -27.6 | 32.3±6.5 | -35.3 | 3.2±0.6 | -36.0 | 71.6±14.2 | -31.48 | 7.1±1.5 | -31.6 |
| BF | 61.6±27.3 | | 6.3±2.8 | | 39.9±9.0 | | 4.1±0.9 | | 101.5±32.3 | | 10.4±3.3 | |
| BF2C | 62.2±20.1 | 1.0 | 6.2±2.0 | -0.6 | 34.8±10.2 | -12.7 | 3.5±1.0 | -13.1 | 97.0±19.8 | -4.48 | 9.7±1.9 | -5.4 |

Note: NF = Natural forest; NF2C = Natural forest converted to cropland; PF= Plantation forest; PF2C = Plantation forest converted to cropland; BF = Bamboo forest; and, BF2C = Bamboo forest converted to cropland

# Paper I

**Were, K.O.,** Dick, Ø.B., Singh, B.R. (2013). Remotely sensing the spatial and temporal land cover changes in the Eastern Mau Forest Reserve and Lake Nakuru drainage basin, Kenya. *Applied Geography*, 41, 75-86.


ELSEVIER
APPLIED
GEOGRAPHY

Applied Geography 41 (2013) 75–86

Contents lists available at SciVerse ScienceDirect

Applied Geography

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

# Remotely sensing the spatial and temporal land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya

K.O. Were [a,b,*], Ø.B. Dick [a], B.R. Singh [a]

[a] Norwegian University of Life Sciences, Institute of Mathematical Sciences and Technology, P.O. Box 5003, NO-1432 Aas, Norway
[b] Kenya Agricultural Research Institute, Kenya Soil Survey, P.O. Box 14733-00800, Nairobi, Kenya

*Keywords:*
Remote sensing
Land cover
Hybrid classification
Spatial reclassification
Change detection
Eastern Mau
Lake Nakuru
Kenya

ABSTRACT

This study aimed at characterizing land cover dynamics for four decades in Eastern Mau forest and Lake Nakuru basin, Kenya. The specific objectives were to: (i) identify and map the major land cover types in 1973, 1985, 2000 and 2011; (ii) detect and determine the magnitude, rates and nature of the land cover changes that had occurred between these dates, and; (iii) establish the spatial and temporal distribution of these changes. Land cover types were discriminated through partitioning, hybrid classification and spatial reclassification of multi-temporal Landsat imagery. The land cover products were then validated and overlaid in post-classification comparison to detect the changes between 1973 and 2011. The accuracies of the land cover maps for 1973, 1985, 2000 and 2011 were 88%, 95%, 80% and 89% respectively. Six land cover classes, namely forests-shrublands, grasslands, croplands, built-up lands, bare lands and water bodies, were mapped. Forests-shrublands dominated in 1973, 1985 and 2000 covering about 1067 km$^2$, 893 km$^2$ and 797 km$^2$ respectively, but were surpassed by croplands (953 km$^2$) in 2011. Bare lands occupied the least area that varied between 2 km$^2$ and 7 km$^2$ during this period. Overall, forests-shrublands and grasslands decreased by 428 km$^2$ and 258 km$^2$ at the annual average rates of 1% each, whereas croplands and built-up lands expanded by 660 km$^2$ and 24 km$^2$ at the annual rates of 6% and 16% respectively. The key hotspots of these changes were distributed in all directions of the study area, but at different times. Therefore, policies that integrate restoration and conservation of natural ecosystems with enhancement of agricultural productivity are strongly recommended. This will ensure environmental sustainability and socio-economic well-being in the area. Future research needs to assess the impacts of the land cover changes on ecosystem services and to project the future patterns of land cover changes.



## Introduction

Land cover is dynamic and varies at different spatial and temporal scales (Cihlar, 2000); yet, its role in the structure and functioning of the earth system is fundamental. The array of ecosystem services it offers include, provisioning services (e.g. food), regulating services (e.g. climate regulation), cultural services (e.g. recreation and ecotourism) and supporting services (e.g. biogeochemical cycling) (Millennium Ecosystems Assessment, 2005). Thus, land cover modifications or conversions through human or nature's agency can have profound impacts on climate, hydrological and biogeochemical cycles, biodiversity, soil quality and human well-being (Foody, 2002; Lambin, Geist, & Lepers, 2003; Overmars & Verburg, 2005; Potter et al., 2007). This justifies the importance of land use and land cover change research in the context of global environmental change and sustainable development.

In Eastern Africa, land cover is constantly changing, especially in the major watersheds, due to various biophysical and societal factors. Lake Nakuru drainage basin, including the entire Eastern Mau forest reserve, in the Kenyan Rift Valley system is among the hotspots where such land cover changes (LCC) have rapidly occurred over the last 3 decades (Baldyga, Miller, Driese, & Gichaba, 2007; Daniels & Bassett, 2002). This is an important study area because Eastern Mau forest is part of the largest closed-canopy montane forest ecosystem in Eastern Africa. The forest is also among the 5 important water catchment areas in Kenya and a major sink of $CO_2$, which is the main driver of global warming and climate change. Lake Nakuru drainage basin has been transformed from a sparsely populated and densely vegetated area to a highly populated, rapidly urbanizing and extensively cultivated area. In 1970, 47% of the area was under natural vegetation, but by 2003, 49% of this had

* Corresponding author. Norwegian University of Life Sciences, Institute of Mathematical Sciences and Technology, P.O. Box 5003, NO-1432 Aas, Norway. Tel.: +47 966 563 62; fax: +47 649 654 01.
*E-mail address:* kenwerez@yahoo.com (K.O. Were).

0143-6228/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.apgeog.2013.03.017

been cleared and the number of small-scale farms had grown (Daniels & Bassett, 2002; UNEP, 2009). Baldyga et al. (2007) also reported that between 1986 and 2003, about one-fifth of the forests in the upper catchment of Njoro River were lost. The loss is partly attributed to illegal encroachments in 1990s and ill-advised political decision to excise about 353 km$^2$ of Eastern Mau forest reserve for human settlement in 2001 (Government of Kenya, 2009).

The demographic and attendant landscape changes, especially forest losses, have had ramifications on water quality (Shivoga et al., 2007), hydrological regime (Mwetu, Mutua, Kundu, Fürst, & Loiskandl, 2009), temperature distribution (Hesslerová & Pokorný, 2010) and biodiversity (Raini, 2009) (See Fig. 1). Therefore, it is imperative to make decisions for mitigating the adverse environmental impacts of the on-going LCC, promoting restoration and sound management of ecosystems in the area. This is indispensable for realization of sustainable development, Millennium Development Goals, Vision 2030 and the National Climate Change Response Strategy in Kenya. However, effective decision-making hinges on availability of the past and present land cover information. The paucity of land cover data in Kenya renders remote sensing the only practical means of providing complete, spatially-explicit, accurate, consistent, quantitative and cost-effective time-series data for systematic mapping, monitoring and analyses of the spatial and temporal land cover dynamics using image processing and Geographic Information Systems (GIS). Space-borne electromagnetic sensors have continuously acquired Earth surface data since the launch of Landsat 1 in 1972 and, thereafter, many other Earth observing systems (e.g. SPOT and ALOS). Several studies have demonstrated the utility of such remotely sensed data in monitoring LCC in different environments including watersheds (Olang, Kundu, Bauer, & Fürst, 2011; Wasige, Groen, Smaling, & Jetten, 2013), mountainous regions (Aguirre-Gutiérrez, Seijmonsbergen, & Duivenvoorden, 2012), urban areas (Dewan & Yamaguchi, 2009; Stefanov, Ramsey, & Christensen, 2001; Yang & Lo, 2002; Yang, Xian, Klaver, & Deal, 2003; Yuan, Sawaya, Loeffelholz, & Bauer, 2005; Wu & Zhang, 2012), estuarine areas (Yang & Liu, 2005), wetlands (Mwita et al., 2013), forests (Gao & Liu, 2012; Laurin et al., 2013; Lung & Schaab, 2010; Pellikka, Lötjönen, Siljander, & Lens, 2009; Schmitt-Harsh, 2013), savannas (Romero-Ruiz et al., 2012), dry lands (Diouf & Lambin, 2001; Tsegaye, Moe, Vedeld, & Aynekulu, 2010; Muriuki et al., 2011), river deltas (Abd El-Kawy et al., 2011, Dewidar, 2004; Seto et al., 2002; Weng, 2002), coastal zones (Kolios & Stylios, 2013; Rodriguez-Galiano & Chica-Olmo, 2012; Shalaby & Tateishi, 2007) and agricultural areas (Shalaby & Ali, 2010).

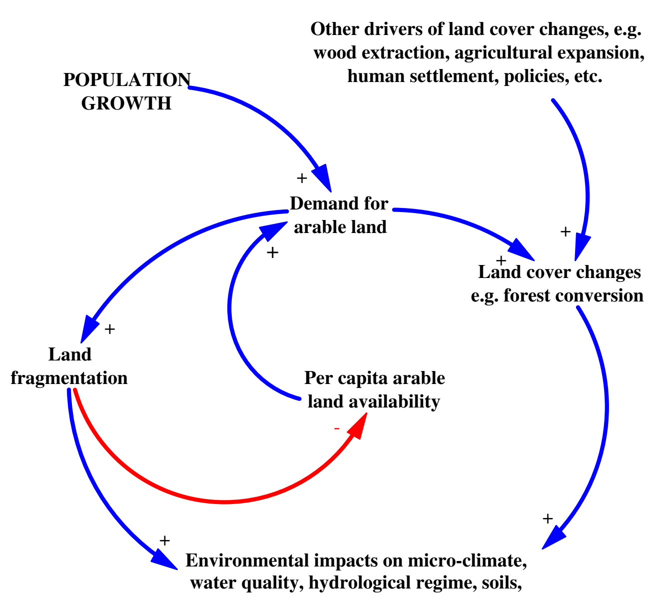



**Fig. 1.** Causal Loop Diagram illustrating the problem. The positive (+) sign denotes an increasing effect while the negative (−) sign denotes a decreasing effect.

The successful application of remote sensing in LCC research is due to the subsequent development of image classification and change detection techniques, improvements on the spatial and spectral properties of optical data, and open accessibility to Landsat archives. Coppin, Jonckheere, Nackaerts, Muys, and Lambin (2004), Lu, Mausel, Brondizio, and Moran (2004), Mas (1999), and Singh (1989) have provided detailed reviews of change detection methods ranging from composite analysis and image differencing, to image ratioing, image regression, linear data transformation, post classification comparison, change vector analysis, neural networks, multi-temporal spectral mixture analysis, multi-dimensional temporal feature space analysis and temporal trajectory analysis. Similarly, Campbell (2002), Lillesand, Kiefer, and Chipman (2008) and Lu and Weng (2007) have explained some of the existing classifiers including the parametric (e.g. maximum likelihood), non-parametric (e.g. artificial neural network), sub-pixel (e.g. spectral mixture analysis), object-oriented (e.g. extraction and classification of homogeneous objects (ECHO)), textural and contextual classifiers.

This paper presents the results of a study that aimed at characterizing land cover and its dynamics for four decades in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya, using remote sensing techniques. The specific objectives were to: (i) identify and map the major land cover types in 1973, 1985, 2000 and 2011; (ii) detect and determine the magnitude, rates and nature of the land cover changes that occurred between these dates, and; (iii) establish the spatial and temporal distribution of the changes. The outputs formed the basis for spatially distributed modelling and assessment of the impact of LCC on terrestrial carbon storage in Eastern Mau forest reserve.

## Materials and methods

### *Study area*

Lake Nakuru drainage basin, including the entire Eastern Mau forest reserve, is located along the Kenyan Rift Valley system, between latitudes 0° 10′–0° 45′ S and longitudes 35° 40′–36° 5′ E (Fig. 2). It covers about 2000 km$^2$ with the altitude varying from 1750 to 3090 m above sea level. The Menengai crater borders it to the north, Ol Doinyo Eburru volcano and Kiambogo hills to the south, Soysambu hills to the east and the Mau escarpment to the west. Makalia, Enjoro, Naishi, Lamuriak, Enderit and Ngosorr river systems drain down this area into Lake Nakuru, whereas Nessuiet and Rongai rivers flow into Lake Bogoria and Baringo respectively. Its main geomorphological features are mountains and major scarps, hills and minor scarps, plateaus, volcanic footridges, uplands, volcanic and lacustrine plains and bottomlands. Geologically, the area is characterized by volcanic rocks (e.g. basalts, trachyte, phonolites, pumice tuff and lavas) and associated sediments of tertiary and quaternary age (McCall, 1967). The soils are also of volcanic origin, which comprise andosols, planosols, vertisols, nitisols, regosols, calcisols, solonetz or phaeozems (Wanjogu, Kibe, Wagate, & Mwangi, 2010). The climate varies from cool and humid to hot and humid depending on the altitude and topography. Higher areas at Mau escarpment receive substantial rainfall (∼2000 mm), which decreases notably (∼700 mm) on the lower areas around Lake Nakuru. The rainfall pattern is bimodal with the long rains falling between March and May and short rains between November and December due to the seasonal north-south movement of the inter-tropical convergence zone (Odada, Raini, &

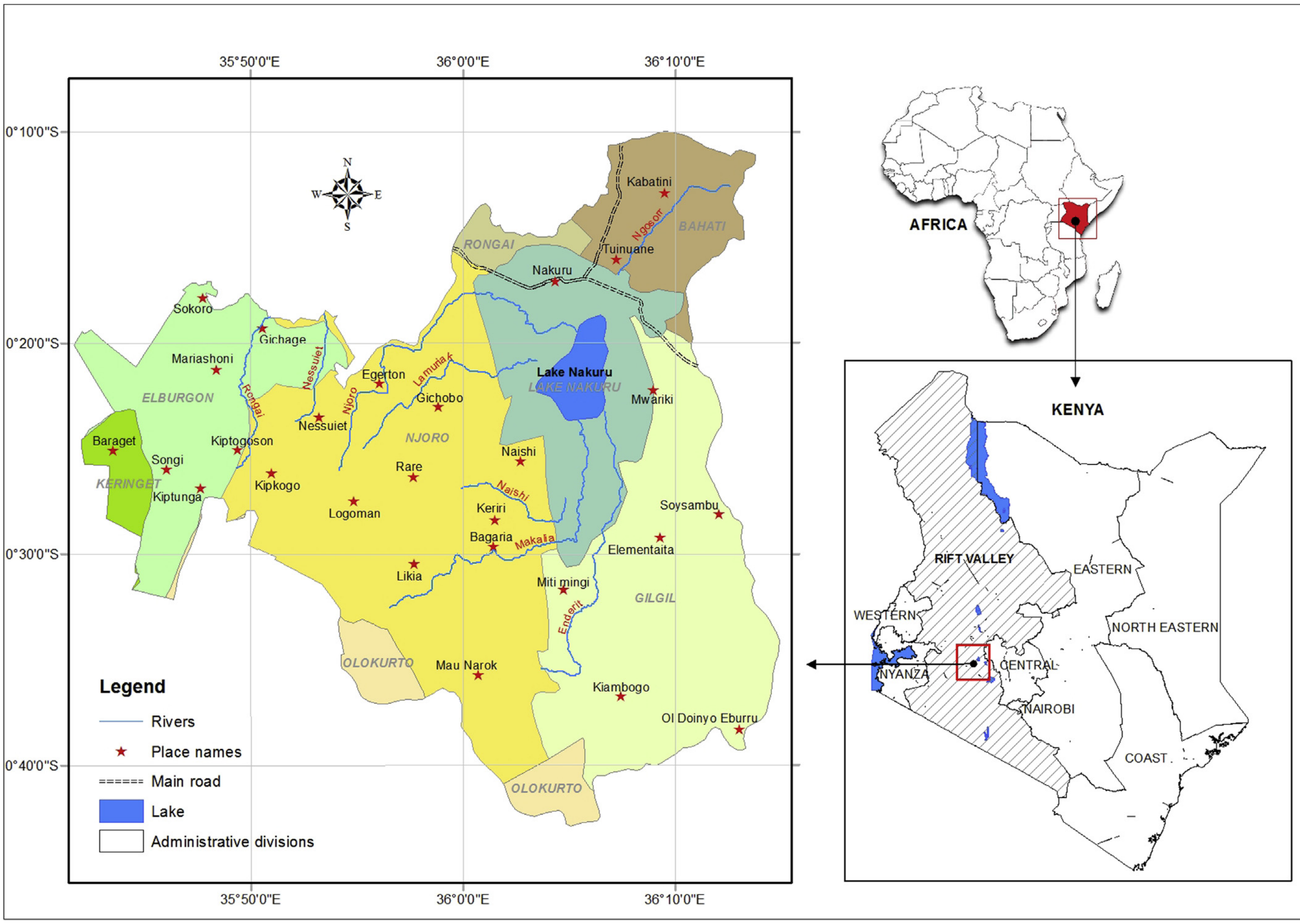



**Fig. 2.** Geographical location of the study area.

Ndetei, 2006). Floristically, indigenous and exotic tree species, e.g. *Pinus patula*, *Cupressus lusitanica*, *Eucalyptus* spp., *Prunus africana*, *Arundinaria alpina*, *Juniperus procera*, *Olea europaea* ssp. *africana*, *Olea hochstetteri*, *Podocarpus latifolius* and *Dombeya torrida*, and grass species, e.g. *Cynodon dactylon*, *Digitaria scalarum* and *Pennistum clandestinum*, cover the area.

### *Data*

The data sources were satellite-based remote sensors, fieldwork and existing spatial databases. These are summarized in Table 1.

#### *Remotely sensed data*

Landsat 1 Multispectral Scanner System (MSS), Landsat 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper plus (ETM+) imagery, acquired in January 1973, 1985, 2000 and 2011 were obtained from the USGS archive (http://earthexplorer.usgs.gov/). These had been processed to level L1T (i.e. terrain-corrected). Landsat data were selected due to their open accessibility, historical record and suitable processing levels, cloud coverage, swath, and spatial, spectral and temporal resolutions.

#### *Field data*

Fieldwork was conducted between July and August 2012 to collect data for identifying the land cover types on Landsat imagery, validating the extracted land cover map for 2011, and understanding land use/cover history of the area. The biophysical land attributes (e.g. percent tree cover) were recorded and georeferenced using Garmin eTrex 30 handheld GPS device at 450 sampling locations. Random sampling strategy was used to objectively select the sampling units into the sample. The sampling units were randomly generated using ArcGIS 10.1 with the scale of 30 m selected to coincide with the spatial resolution of the Landsat imagery. Interviews were also conducted with key informants, particularly, the local administrators, forest managers, farmers, elders and community group leaders.

#### *Ancillary data*

Existing digital spatial data comprising the topographical maps, Google Earth imagery and thematic layers (i.e. Africover land cover map, administrative boundaries, towns, villages, roads, forests, protected areas, rivers and elevation) were also obtained and prepared to facilitate fieldwork, image classification, post-classification processing and validation.

### *Data preparation and analysis*

Image pre-processing, classification scheme design, image classification, post-classification processing, spatial reclassification, accuracy assessment and change detection were performed.

#### *Image pre-processing*

This involved rectifying the radiometric and geometric distortions of the satellite data prior to classification. Firstly, the

**Table 1**
Data characteristics.

| Data type | Date of acquisition | Spatial scale | Source |
|---|---|---|---|
| Remotely sensed data: | | | |
| Landsat 7 ETM+ (Path 161 row 060) | January 2000 | Multi-spectral: 30×30 m<br>Panchromatic: 15×15 m | USGS (http://earthexplorer.usgs.gov/) |
| Landsat 5 TM (Path 161 row 060) | January 1985 & 2011 | Multi-spectral: 30×30 m | USGS (http://earthexplorer.usgs.gov/) |
| Landsat 1 MSS (Path 181 row 060) | January 1973 | Multi-spectral: 60×60 m | USGS (http://earthexplorer.usgs.gov/) |
| Ancillary data: | | | |
| Land cover (Africover) | 2003 | 1: 200,000 | FAO (http://www.africover.org/) |
| Digital topographical maps (Sheet 118/4, 119/1, 119/3, 132/2 &133/1) | 1975, 1974, 1997, 1975 &1975 | 1: 50,000 | Survey of Kenya |
| Google Earth imagery | – | – | Google Inc. |
| GIS thematic layers (administrative, rivers, towns, villages & roads) | – | – | ILRI (http://www.ilri.org/gis) |
| Field data | July–August 2012 | Sampling unit: 30×30 m | Field survey |

downloaded Landsat TM and ETM+ data for each date were unzipped and 6 bands (excluding the thermal band) were stacked to form multi-band images using ERDAS imagine 2011. Then, the multi-band images were reprojected to the Universal Transverse Mercator grid (Zone 37S, WGS 84 ellipsoid and datum) using the nearest neighbour resampling method, geometrically co-registered and subsets prepared. Finally, the image-based COST technique for atmospheric correction coupled with a radiometric model was applied to compensate for the systematic and random sensor noise, as well as the atmospheric effects. This implemented the concept of dark object subtraction (Chavez, 1996) to remove the atmospheric effects in Landsat's bands 1–5 and 7. The raw digital numbers were first converted to at-satellite spectral radiance and then to atmospherically-corrected reflectance with values ranging between 0 and 1. The image-based COST technique was adopted because atmospheric profile data during the satellite overpass, which are needed for absolute radiometric calibration, were not available.

*Classification scheme design*

Due to lack of a standard land cover classification system for remote sensing applications in Kenya, a classification nomenclature was developed (Table 2). The land cover classes were defined based on the percentage of biophysical cover noted at the sampling sites during fieldwork and also on modification of the definitions used by Anderson, Harvey, Roach, and Witmer (1976) and the USGS' National Land Cover Database (NCLD) 2006 (http://www.mrlc.gov/nlcd06_leg.php). The main considerations were the spatial resolution of Landsat data being used and possibility of the classification system to interface with the others.

*Image classification, post-classification processing and spatial reclassification*

Partitioning, hybrid classification and spatial reclassification approach was used to discriminate land cover types on the image subsets using ERDAS imagine 2011. Firstly, the subset for each date was further subdivided into a number of spectrally distinct segments, using the area of interest (AOI) tool, for separate classification. This was to minimize classification errors due to spectral confusion within the highly heterogeneous and fragmented scenes. Unsupervised classification using ISODATA algorithm was executed to define 15 - 30 spectral clusters in each segment depending on the complexity. The resultant clusters were assigned to the 6 land cover classes (Table 2) based on ancillary and field data, and analyst's knowledge of the area. Where there was misclassification of pixels, additional spectral signatures were extracted and merged with the

**Table 2**
Land cover classification scheme.

| Land cover class | Definition | Anderson's equivalent class | NCLD's equivalent class |
|---|---|---|---|
| Forests-shrublands | Areas covered by (i) >60% natural or planted woody vegetation, which are >6 m tall and have a crown density of >40%, and; (ii) >60% natural or planted woody vegetation, which are <6 m tall and have a crown density of >30%. The latter also includes woody vegetation with sparse foliage cover (10–30%) and stunted growth (<5 m tall) found on the drier parts of Lake Nakuru basin. | Forest land | Forest, shrubland |
| Grasslands | Areas dominated by >60% grasses, grass-likes or herbs, often mixed with sparse trees, shrubs or scrubs (<20%), and are managed either by agronomic, forestry or ecological principles. | Rangeland | Grassland/herbaceous |
| Croplands | Areas where growing herbaceous crops account for >60% of the cover, or fields have been ploughed for planting crops. | Agricultural land | Cultivated crops |
| Built-up lands | Areas characterized by >60% constructed or impervious materials (e.g. asphalt, concrete, buildings). | Built-up land | Developed |
| Bare lands | Areas characterized by >60% soils (gravel, sand, silt, clay), rock outcrops, quarries or dry salts, with or without vegetation (<10%). | Barren land | Barren land |
| Water bodies | Open areas covered with water (e.g. Lake) | Water | Open water |

signatures from ISODATA clustering. These were then classified using the maximum likelihood algorithm. See Campbell (2002) and Lillesand et al. (2008) for an explanation of the algorithms. The land cover classes in each classified segment were recoded and, thereafter, the reclassified segments for each date were mosaicked. In post-classification processing, ancillary data, visual appraisal and GIS functions (e.g. on-screen digitization, extraction by AOIs, reclassification and mosaicking) were integrated to improve the accuracy of the land cover maps. Lastly, a $3 \times 3$ majority filter was applied to reduce noise on the final seamless land cover maps.

*Accuracy assessment*

The quality of the land cover maps was evaluated both qualitatively and quantitatively. Firstly, each land cover map and the corresponding Landsat data were displayed on-screen and visually inspected. Then the uncertainties related to the land cover products of 1973, 1985, 2000 and 2011 were quantified by comparing them with the topographical maps published in 1974, temporally-invariant land cover data, Africover land cover map produced in 2003 (using Landsat TM images acquired in 1999) and ground data collected in 2012 respectively. The temporally-invariant land cover data were created using the steps described by Fortier, Rogan, Woodcock, and Runfola (2011). These included: (i) overlaying the imagery for 1973 (MSS) and 1985 (TM); (ii) identifying and digitizing polygons at the centre of invariant sites (e.g. established urban areas and large forests), and; (iii) assigning land cover labels using the topographical maps. The Africover land cover map was also reclassified based on the classes shown on Table 2. Three sets of 600 random feature points were thereafter created on the temporally-invariant land cover data, reclassified Africover land cover map and topographical maps. The associated land cover attributes were also extracted to these points. These sets of random feature points, together with the 450 ground data, were compared with the corresponding pixels on the land cover products of the respective years to ascertain the number of correctly and incorrectly classified pixels for each land cover class. The results were presented on conventional error matrices, from which statistical measures of map accuracy (i.e. Kappa statistics, overall-, producer's- and user's accuracy) were computed (Campbell, 2002; Congalton, 1991; Foody, 2002).

*Land cover change detection*

The validated land cover maps for 1973 (resampled to 30 m), 1985, 2000 and 2011 were overlaid in post-classification comparison to detect the pixel by pixel land cover changes between 1973–1985, 1985–2000, 2000–2011 and 1973–2011. The outputs were cross-tabulation matrices showing the pathways, and change maps showing the spatial patterns, of the LCC. The change maps were as accurate as the product of the overall accuracies of the individual land cover maps that produced them (Singh, 1989; Yuan et al., 2005). Such outputs explain the popularity of post-classification comparison method for change detection. It also reduces the impact of radiometric and atmospheric differences because the imagery are classified independently (Coppin et al., 2004; Lu et al., 2004; Singh, 1989).

## Results

*Land cover classification and accuracy assessment*

The spatial patterns of the six major land cover types in Lake Nakuru drainage basin and Eastern Mau forest reserve in 1973, 1985, 2000 and 2011 are presented in Fig. 3. Table 4 shows that forests-shrublands were dominant in 1973, 1985 and 2000 covering about 1067 km$^2$, 893 km$^2$ and 797 km$^2$ respectively, but were surpassed by croplands (953 km$^2$) in 2011. Bare lands occupied the least area that varied between 2 km$^2$ and 7 km$^2$ during this period. Fig. 3 also shows that in 1973, only a few croplands were located in the northern and southern parts, which by 1985 had diffused centrally to cover Bagaria, Keriri, Naishi, Pwani mutukanio, Miti mingi, Lare and Gichobo. The maps of 2000 and 2011 further reveal a westward pattern of cropland expansion towards the Eastern Mau forest. Thus, the forest-shrublands, which dominated the western and middle regions in 1973, show a general westward pattern of contraction. On the northern side of Lake Nakuru, the built-up lands (i.e. Nakuru town and its environs) exhibit a west to east growth pattern from 1973 to 2011. While on the eastern side, around Lake Nakuru National Park, Soysambu and Elementaita, the grasslands show a relatively stable pattern of dominance throughout the period. This is because most of the land is protected and dedicated to wildlife conservation and ranching. The water bodies, mainly Lake Nakuru, also indicate this stability.

Table 3 shows that the overall cartographic accuracies of the 1973, 1985, 2000 and 2011 land cover maps were 88%, 95%, 80% and 89%, with overall Kappa statistics of 82%, 93%, 72% and 84% respectively. Therefore, the accuracies of the resulting land cover change maps (Fig. 5), which are the products of overall accuracies of the individual land cover maps, were 84% for 1973–1985, 76% for 1985–2000, 71% for 2000–2011 and 78% for 1973–2011. In addition, the user's and producer's accuracies achieved for all classes, except for grasslands, were above 70%.

*Land cover change detection*

The magnitudes and the annual average rates of change for the land cover types are shown in Table 4 and Fig. 4. Fig. 4 reveals that most of the land cover changes were uni-directional except for the bare lands and water bodies. Forests-shrublands, grasslands and croplands had higher magnitudes of change compared to the built-up lands, bare lands and water bodies during the three periods (Table 4). More specifically, the forests-shrublands and grasslands decreased by 428 km$^2$ and 258 km$^2$ at the annual average rates of 1% each, while croplands and built-up lands expanded by 660 km$^2$ and 24 km$^2$ at the annual average rates of 6% and 16% respectively. Judging by the annual average rates of change, built-up lands were the most dynamic having grown at the rates of 2%, 17% and 5% respectively within the three periods. The built-up lands encompasses Nakuru town and its environs, where the population has been growing rapidly at the rate of 5.6% per annum (Mubea & Menz, 2012).

Table 5 comprise the land cover change matrices showing the nature of the land cover changes. The major diagonals represent the amounts of each land cover type that did not change (persistence) at a given time while the off-diagonals indicate the gains, losses and trajectories of the conversions. For example, Table 5 (d) reveals that out of the 1067 km$^2$ of forests-shrublands in 1973, 530 km$^2$ were stable while 92 km$^2$, 441 km$^2$, 2 km$^2$ and 1 km$^2$ were lost to grasslands, croplands, built-up lands and bare lands respectively. The total forests-shrublands losses were 536 km$^2$. Some of the 639 km$^2$ of forests-shrublands in 2011 were gained from grasslands (87 km$^2$) and croplands (22 km$^2$); hence, the total gains were 109 km$^2$. Similarly, out of the 589 km$^2$ of grasslands in 1973, 230 km$^2$ remained unchanged while 87 km$^2$, 250 km$^2$, 21 km$^2$ and 1 km$^2$ changed to forests-shrublands, croplands, built-up lands and bare lands respectively. Some of the 331 km$^2$ of grasslands in 2011 were gained from forests-shrublands (92 km$^2$), croplands (7 km$^2$) and bare lands (1 km$^2$). Croplands increased markedly from 293 km$^2$ in 1973 to 953 km$^2$ after gaining 441 km$^2$ and 250 km$^2$ from forests-shrublands and grasslands respectively. Only 262 km$^2$ of the croplands in 1973 were persistent while the rest were

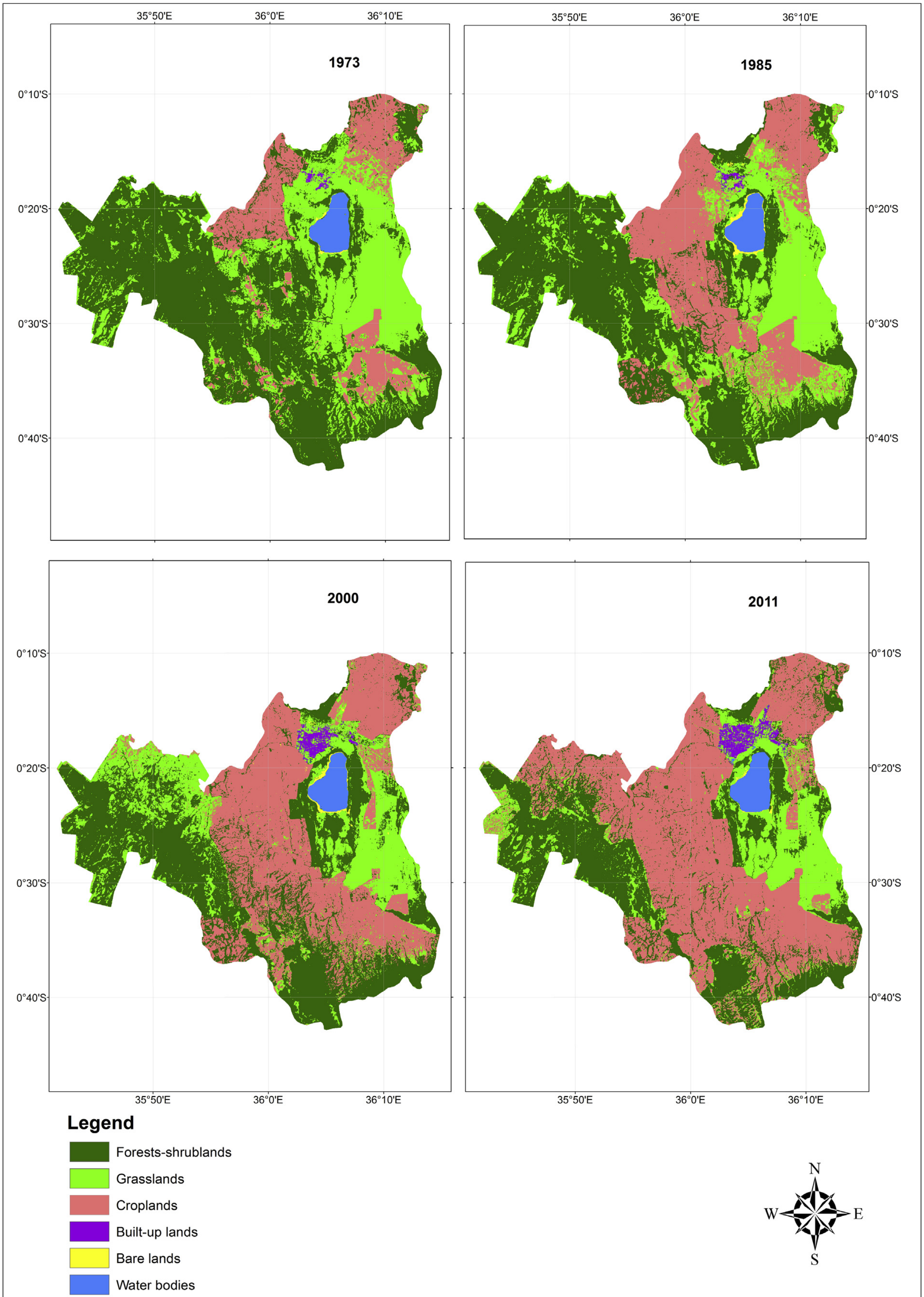



**Fig. 3.** Land cover classification maps for Eastern Mau forest reserve and Lake Nakuru drainage basin.

**Table 3**
Accuracy statistics (in %) for the land cover classifications.

| Land cover class | 1973 | | 1985 | | 2000 | | 2011 | |
|---|---|---|---|---|---|---|---|---|
| | Producer's | User's | Producer's | User's | Producer's | User's | Producer's | User's |
| Forests-shrublands | 91 | 95 | 100 | 88 | 81 | 89 | 92 | 85 |
| Grasslands | 96 | 62 | 99 | 97 | 81 | 57 | 53 | 59 |
| Croplands | 71 | 99 | 83 | 100 | 78 | 87 | 91 | 90 |
| Built-up lands | 79 | 100 | 95 | 100 | 73 | 100 | 84 | 100 |
| Bare lands | 100 | 100 | 100 | 100 | 85 | 86 | 100 | 100 |
| Water bodies | 100 | 100 | 100 | 100 | 100 | 91 | 100 | 100 |
| Overall accuracy | 88 | | 95 | | 80 | | 89 | |
| Overall Kappa | 82 | | 93 | | 72 | | 84 | |

**Table 4**
Areal extent, magnitude and the annual average rates of land cover changes.

| Land cover class | Area ($km^2$) | | | | Magnitude of change ($km^2$) | | | | Average rate of change p.a (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1973 | 1985 | 2000 | 2011 | 73–85 | 85–00 | 00–11 | 73–11 | 73–85 | 85–00 | 00–11 | 73–11 |
| Forests-shrublands | 1067 | 893 | 797 | 639 | −174 | −96 | −158 | −428 | −1 | −1 | −2 | −1 |
| Grasslands | 589 | 531 | 421 | 331 | −58 | −110 | −90 | −258 | −1 | −1 | −2 | −1 |
| Croplands | 293 | 521 | 714 | 953 | 228 | 193 | 239 | 660 | 6 | 2 | 3 | 6 |
| Built-up lands | 4 | 5 | 18 | 28 | 1 | 13 | 10 | 24 | 2 | 17 | 5 | 16 |
| Bare lands | 4 | 7 | 7 | 2 | 3 | 0 | −5 | −2 | 6 | 0 | −6 | −1 |
| Water bodies | 42 | 40 | 40 | 43 | −2 | 0 | 3 | 1 | 0 | 0 | 1 | 0 |

replaced by forests-shrublands (22 $km^2$), grasslands (7 $km^2$) and built-up lands (2 $km^2$). Lastly, the growth of built-up lands from 4 $km^2$ in 1973 to 28 $km^2$ in 2011 came at the expense of forests-shrublands (2 $km^2$), grasslands (21 $km^2$) and croplands (2 $km^2$). These changes reflect the land cover dynamics, but it is also acknowledged that some may be attributed to classification and data errors. For instance, even though changes from croplands to forests-shrublands and grasslands may be explained by the agro-forestry practices and agricultural cycles in parts of the area, classification errors cannot also be ruled out.

The spatial patterns of the land cover changes are displayed on Fig. 5. It shows that between 1973 and 1985, the major hotspots of the forest-shrubland conversions were distributed in the mid regions (e.g. Gichobo, Naishi, Keriri and Bagaria) and northern side of Lake Nakuru (e.g. Menengai). Between 1985 and 2000, these conversions shifted to the western side spanning Baraget, Mariashoni, Sokoro, Gichage, Nessuiet, Logoman, Likia and Sururu. In the last period, forest-shrubland conversion sites not only occurred in the west, but also in the south (e.g. Kiambogo) and around Lake Nakuru National park. Similarly, between 1973 and 1985, grassland transitions were located in the mid and southern regions. Between 1985 and 2000, these conversions spread to other areas in the west (e.g. Teret and Likia), east (e.g. Mwariki), north (e.g. Tuinuane and Nakuru) and south (e.g. Kiambogo and Miti mingi). In the final period, 2000–2011, the western part, particularly Mariashoni, Gichage, Nessuiet and Sokoro, became the hotspots of grassland

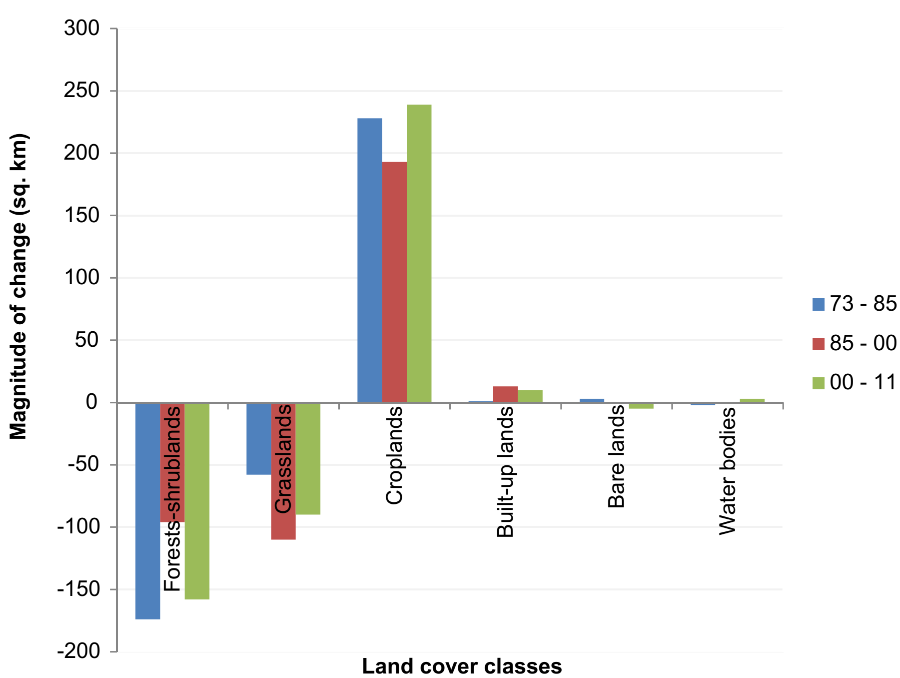



**Fig. 4.** Comparative illustration of the magnitude of the land cover changes.

**Table 5**
Nature of the land cover changes from 1973 to 2011 (area in km$^2$).

(a) 1973–1985

| | | 1973 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Fs | G | C | Bt | B | W | Total | Gain |
| 1985 | Forests-shrublands (Fs) | **781** | 102 | 10 | 0 | 0 | 0 | 893 | 111 |
| | Grasslands (G) | 140 | **352** | 37 | 1 | 0 | 0 | 531 | 179 |
| | Croplands (C) | 144 | 131 | **246** | 0 | 0 | 0 | 521 | 275 |
| | Built-up lands (Bt) | 0 | 2 | 0 | **2** | 0 | 0 | 5 | 3 |
| | Bare lands (B) | 1 | 2 | 0 | 0 | **2** | 2 | 7 | 5 |
| | Water bodies (W) | 0 | 0 | 0 | 0 | 0 | **40** | 40 | 0 |
| Total | | 1067 | 589 | 293 | 4 | 2 | 42 | 1997 | |
| Loss | | 285 | 237 | 47 | 2 | 0 | 2 | | |

(b) 1985–2000

| | | 1985 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Fs | G | C | Bt | B | W | Total | Gain |
| 2000 | Forests-shrublands (Fs) | **658** | 106 | 32 | 0 | 1 | 0 | 797 | 139 |
| | Grasslands (G) | 133 | **262** | 25 | 0 | 1 | 0 | 421 | 159 |
| | Croplands (C) | 99 | 153 | **461** | 0 | 0 | 0 | 714 | 253 |
| | Built-up lands (Bt) | 2 | 9 | 3 | **5** | 0 | 0 | 18 | 13 |
| | Bare lands (B) | 1 | 1 | 0 | 0 | **5** | 0 | 7 | 2 |
| | Water bodies (W) | 0 | 0 | 0 | 0 | 0 | **40** | 40 | 1 |
| Total | | 893 | 531 | 521 | 5 | 7 | 40 | 1997 | |
| Loss | | 235 | 269 | 59 | 0 | 2 | 0 | | |

(c) 2000–2011

| | | 2000 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Fs | G | C | Bt | B | W | Total | Gain |
| 2011 | Forests-shrublands (Fs) | **542** | 56 | 41 | 0 | 0 | 0 | 639 | 97 |
| | Grasslands (G) | 95 | **209** | 22 | 3 | 2 | 0 | 331 | 122 |
| | Croplands (C) | 157 | 146 | **650** | 0 | 0 | 0 | 953 | 304 |
| | Built-up lands (Bt) | 1 | 10 | 1 | **15** | 0 | 0 | 28 | 13 |
| | Bare lands (B) | 0 | 0 | 0 | 0 | **1** | 0 | 2 | 1 |
| | Water bodies (W) | 1 | 0 | 0 | 0 | 3 | **40** | 43 | 3 |
| Total | | 797 | 421 | 714 | 18 | 7 | 40 | 1997 | |
| Loss | | 255 | 213 | 65 | 3 | 5 | 0 | | |

(d) 1973–2011

| | | 1973 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Fs | G | C | Bt | B | W | Total | Gain |
| 2011 | Forests-shrublands (Fs) | **530** | 87 | 22 | 0 | 0 | 0 | 639 | 109 |
| | Grasslands (G) | 92 | **230** | 7 | 1 | 1 | 0 | 331 | 101 |
| | Croplands (C) | 441 | 250 | **262** | 0 | 0 | 0 | 953 | 691 |
| | Built-up lands (Bt) | 2 | 21 | 2 | **3** | 0 | 0 | 28 | 25 |
| | Bare lands (B) | 1 | 1 | 0 | 0 | **1** | 0 | 2 | 2 |
| | Water bodies (W) | 1 | 0 | 0 | 0 | 0 | **42** | 43 | 2 |
| Total | | 1067 | 589 | 293 | 4 | 2 | 42 | 1997 | |
| Loss | | 536 | 359 | 31 | 1 | 1 | 0 | | |

Note: Gain is the sum of the off-diagonal values in a row, loss is the sum of the off-diagonal values in a column.

conversions. Finally, the spatial distribution pattern of cropland change was rather patchy. Between 1973 and 1985, the changes were located in the southern and northern parts, but in the next two periods, patches of conversions appeared on all sides of the area.

## Discussion

### *Land cover classification and accuracy assessment*

The classification results have shown that forests-shrublands dominated Lake Nakuru Basin and Eastern Mau Forest Reserve in 1973, 1985 and 2000, but were overtaken by croplands between 2000 and 2011. This clearly indicates the on-going transition from natural to human-dominated environment and from timber production to crop production (agriculture) as the major land use the area.

The overall accuracies of the land cover maps for 1973, 1985, 2000 and 2011 were above 80% and, hence, met the target accuracy threshold of 80–85% for thematic mapping in satellite remote sensing (Treitz & Rogan, 2004). These maps were also 70–80% better than what was expected from random assignment of pixels to classes based on their overall Kappa coefficients. The user's accuracies imply that over 70% of each classified land cover type (except for grasslands) could be reliably located on the ground by the end-users. Similarly, the producer's accuracies mean that over 70% of each land cover type (except for grasslands) was correctly classified by the analyst. Despite the per-class variations, the accuracy levels are generally good considering the complexity of the study area. Thus, the land cover maps qualify for other applications depending on the level of accuracy desired by the users. For instance, for change detection studies as this one, some of the differences observed over time would be spurious because the resultant change maps would be just as accurate as the product of the overall accuracies of the bi-temporal land cover maps that produced them (Foody, 2002).

The classification errors are mainly attributed to the spectral confusion between croplands and grasslands and, between forests-shrublands and croplands. These were compounded by the heterogeneity of the landscape and agricultural dynamics that caused conversions from grass to crops and vice versa at different periods. This explains the low accuracy values obtained for grasslands. Despite these challenges, the partitioning, hybrid classification and spatial reclassification technique that was used proved its utility by producing satisfactory classification results for the study area. It can be replicated in mapping other complex landscapes in Kenya. But the ultimate success of such projects lies in application of suitable techniques that integrate spectral, spatial and contextual information with GIS functions and human intelligence.

Application of temporally-invariant land cover data to evaluate the quality of the land cover map for 1985 satisfactorily resolved the problem of missing historical ground data for validation. This complements the independent visual interpretation of unclassified satellite imagery that has been employed in other studies (Biro, Pradhan, Buchroithner, & Makeschin, 2011) to extract information for validating historical land cover maps. Reference data *per se* are never perfect (Congalton, 1991; Foody, 2010); hence, it is also instructive to highlight some of the uncertainties associated with the ones used here. Firstly, the Africover land cover map was vector-based, produced through manual per-field classification of Landsat imagery, whereas the corresponding classified map was raster-based, produced through automatic per-pixel classification of Landsat imagery. The coarseness of the former's spatial unit for comparison could, for example, increase its degree of agreement with the latter. Moreover, the classification scheme of the Africover land cover map and the one used in the study had different levels of detail, so the former was reclassified to harmonize the map legends. Reclassification of mixed polygons that had been assigned 2 different land cover codes was a potential source of bias. Thus, only pure land cover polygons were included in the reference sample in order to retain the value of direct comparison of the maps. Lastly, the timing of Landsat image acquisition in 2011 and fieldwork did not coincide. The ground data were collected in July and August 2012, while the imagery was acquired in January 2011. Thus, some areas were bare during image acquisition but fallow, cultivated or with grass during fieldwork. The appropriate adjustments were made though using the crop calendar, land use history, ancillary spatial data and analyst's experience and knowledge.

### *Land cover changes*

The land cover change results from 1973 to 2011 have been presented in terms of the magnitude, annual average rates, nature and spatial distribution. The most important land cover changes

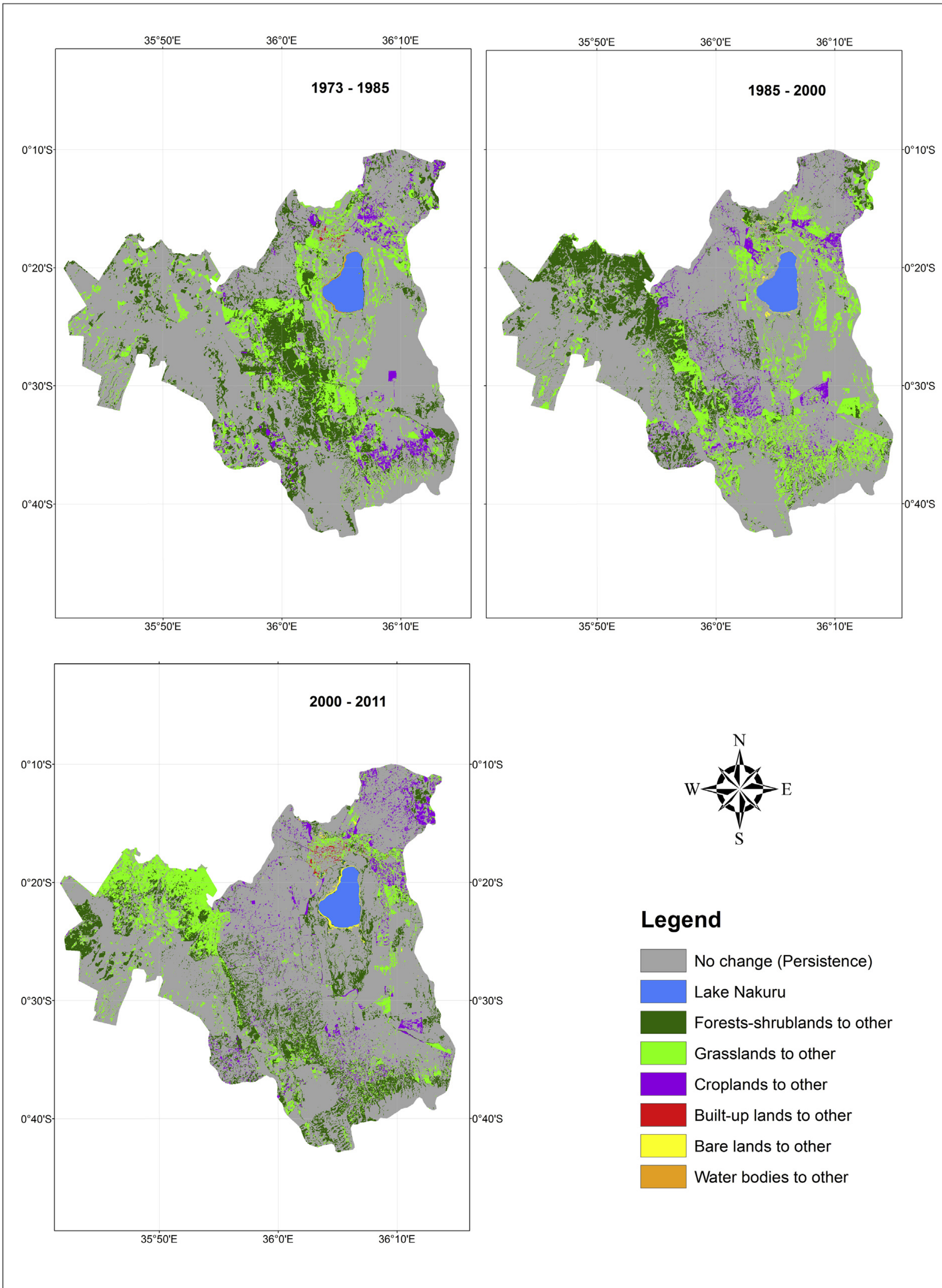



**Fig. 5.** Maps showing the spatial distribution of the land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin from 1973 to 2011.

revealed during this 38-year period were forest-shrubland and grassland conversions in favour of croplands and built-up lands. This concurs with the reports given by Odada et al. (2006) and UNEP (2009). Similar trends were also observed in the upper catchment of Njoro River (Baldyga et al., 2007), which falls within the study area, and also in the neighbouring Lake Baringo catchment (Kiage, Liu, Walker, Lam, Huh, 2007). Such losses of natural elements in the landscape affect the climate (Hesslerová & Pokorný, 2010; Otieno & Anyah, 2012), soil properties (Biro et al., 2011; Braimoh & Vlek, 2004), hydrology and water quality (Mati, Mutie, Gadain, Home, & Mtalo, 2008; Palamuleni, Ndomba, & Annegarn, 2011; Shivoga et al., 2007) and biodiversity (Raini, 2009). Therefore, holistic and sustainable environmental strategies that blend restoration and conservation of the natural ecosystems with enhancement of agricultural productivity should be formulated and implemented.

Furthermore, the land cover change maps showed that the forest-shrubland conversion hotspots, which should be targeted for restoration, developed in the mid, northern, western and southern regions at different times. These are consistent with the historical accounts of land uses and the biophysical and societal processes that have operated in these parts. For example, Eastern Mau in the west was covered by indigenous forests and inhabited by the *Ogieks*, who were hunter-gatherers and bee keepers, in the precolonial era (Krupnik, 2004). In 1900s, the colonial government began felling the indigenous trees and replacing them with fast-growing exotic species to meet the rising industrial and domestic demand for wood. Consequently, the *shamba* system, which allowed farmers to grow food crops in small plots where trees had been cut and concurrently plant and nurture tree seedlings, was introduced. But its abuse, particularly cultivation of natural forests in the ensuing years, led to its ban in 1987. In 1980s, logging was mostly done by commercial enterprises (e.g. Timsales Ltd). Deforestation in 1990s, which occurred at the annual average rate of about 1% (Table 4), followed illegal encroachments and excisions of parts of the Eastern Mau forest (Odada et al., 2006). In theory, the government exerted powers provided by the Forest Act 1942 (Cap. 385), to excise about 353 $km^2$ plantation forests to resettle the 1990s victims of ethnic clashes from Molo, Likia, Mauche and Njoro, and about 3000 *Ogiek* families that lived in the indigenous forests (Government of Kenya, 2009). But in reality, patronage politics, where votes were rewarded with forest resources, especially with the advent of multi-partyism in 1990s (Klopp, 2012), ensured that mostly people from particular districts (i.e. Koibatek, Baringo, Bomet, Kericho, Bureti and Transmara) were allocated land in Mariashoni, Nessuiet, Teret, Likia, Baraget and Sururu forests. The weak policy, legal and institutional framework at the time meant, for instance, that environmental impact assessments were not conducted for the forest excisions and settlement schemes, and that stakeholders did not participate in forest management. Clear-cutting and burning mostly preceded cultivation of the allotted forest lands in late 1990s and early 2000s. And upon depletion of trees, the adjacent indigenous forests obviously became the target for illegal logging and charcoal burning in 2000s (Fig. 6).

Like in Eastern Mau, Gichobo, Lare, Bagaria and Naishi locations, in the mid regions, were scarcely populated and mostly covered by woody vegetation in early 1970s. Then in mid-1970s, the government resettled evictees from other forests (e.g. Sabatia and Maji mazuri) in these areas. Workers from the nearby forests (e.g. Nessuiet, Logoman and Likia) and white settlers' farms, and people from other areas (e.g. Central Province) also benefited from these settlement schemes. Clearance of natural vegetation preceded extensive cultivation in 1980s. Additionally, vegetation and large-scale white settlers' farms dominated Kiambogo location in the south prior to 1973. But due to post-independence changes, some of the farms were subdivided by the government and allocated to the farm workers, women groups and local people in 1973, 1978 and 1983. This initiated the fragmentation of this rural landscape.


**Fig. 6.** (a) Croplands expansion; (b) charcoal burning, and; (c) illegal logging in Eastern Mau forest. Source: Were, K.O.

## Conclusions

In conclusion, the major land cover types in Lake Nakuru drainage basin and Eastern Mau forest reserve in 1973, 1985, 2000 and 2011 were forests-shrublands, grasslands, croplands, built-up lands, bare lands and water bodies. Partitioning, hybrid classification and spatial reclassification technique used for the discrimination of these land cover types on Landsat imagery thus provides a promising alternative for classification of the complex tropical African landscapes. Within the 38-year period, notable land cover transformations were detected. The gross loss was highest for forests-shrublands followed by grasslands, while the gross gain was highest for croplands followed by built-up lands. Obviously, the spatial scale of human activities widened with time. The main hotspots of these land cover changes occurred in all directions, at different times, depending on the biophysical and societal processes in operation. Therefore, policies that target both restoration and conservation of natural ecosystems, as well as enhancement of agricultural productivity are recommended for environmental sustainability and socio-economic well-being in the area. Our outputs would provide a good base of geospatial information for such policy formulation. Future research needs to assess the impacts of the land cover changes on ecosystem services (e.g. carbon storage) and to project the patterns of future land cover changes.

## Acknowledgements

This work was funded by the Research Council of Norway through the Norwegian University of Life Sciences. We greatly thank Mr. P. Owenga for technical support, Mr. E. Thairu for excellent driving under difficult field conditions, and the two anonymous reviewers for improving the manuscript.

## References

Abd El-Kawy, O. R., Rød, J. K., Ismail, H. A., & Suliman, A. S. (2011). Land use and land cover change detection in the western Nile delta of Egypt using remote sensing data. *Applied Geography, 31*, 483–494.
Aguirre-Gutiérrez, J., Seijmonsbergen, A. C., & Duivenvoorden, J. F. (2012). Optimizing land cover classification accuracy for change detection, a combined pixel-based and object-based approach in a mountainous area in Mexico. *Applied Geography, 34*, 29–37.
Anderson, J. R., Harvey, E. E., Roach, J. T., & Witmer, R. E. (1976). *A land use and land cover classification system for use with remotely sensed data. US geological survey professional paper 964*. Washington, D.C: United States Government Printing Office.
Baldyga, T., Miller, S. N., Driese, K. L., & Gichaba, C. M. (2007). Assessing land cover change in Kenya's Mau forest region using remotely-sensed data. *African Journal of Ecology, 46*, 46–54.
Biro, K., Pradhan, B., Buchroithner, M., & Makeschin, F. (2011). Land use/land cover change analysis and its impact on soil properties in the northern part of Gadarif region, Sudan. *Land Degradation and Development*, http://dx.doi.org/10.1002/ldr.1116.
Braimoh, A. K., & Vlek, P. L. G. (2004). The impact of land-cover change on soil properties in northern Ghana. *Land Degradation and Development, 15*, 65–74.
Campbell, J. B. (2002). *Introduction to remote sensing*. London: Taylor & Francis.
Chavez, P. S. (1996). Image-based atmospheric corrections: revisited and improved. *Photogrammetric Engineering and Remote Sensing, 62*(9), 1025–1036.
Cihlar, J. (2000). Land cover mapping of large areas from satellites: status and research priorities. *International Journal of Remote Sensing, 21*, 1093–1114.
Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35–46.
Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing, 25*(9), 1565–1596.
Daniels, R., & Bassett, T. J. (2002). The spaces of conservation and development around Lake Nakuru national park, Kenya. *The Professional Geographer, 54*, 481–490.
Dewan, A. M., & Yamaguchi, Y. (2009). Land use and land cover change in Greater Dhaka, Bangladesh: using remote sensing to promote sustainable urbanization. *Applied Geography, 29*, 390–401.
Dewidar, K. M. (2004). Detection of land use/land cover changes for the northern part of the Nile delta (Burullus region), Egypt. *International Journal of Remote Sensing, 25*(20), 4079–4089.
Diouf, A., & Lambin, E. F. (2001). Monitoring land-cover changes in semi-arid regions: remote sensing data and field observations in the Ferlo, Senegal. *Journal of Arid Environments, 48*, 129–148.
Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment 80*, 185–201.
Foody, G. M. (2010). Assessing the accuracy of land cover change with imperfect ground reference data. *Remote Sensing of Environment, 114*, 2271–2285.
Fortier, J., Rogan, J., Woodcock, C. E., & Runfola, D. M. (2011). Utilizing temporally invariant calibration sites to classify multiple dates and types of satellite imagery. *Photogrammetric Engineering and Remote Sensing, 77*(2), 181–189.
Gao, G., & Liu, Y. (2012). Deforestation in Heilongjiang province of China, 1896–2000: severity, spatio-temporal patterns and causes. *Applied Geography, 35*, 345–352.
Government of Kenya. (2009). *Report of the prime minister's task force on the conservation of the Mau forest complex*. Available online: http://www.maurestoration.go.ke/. Accessed 20.12.12.
Hesslerová, P., & Pokorný, J. (2010). Effect of Mau forest clear cut on temperature distribution and hydrology of catchment of Lakes Nakuru and Naivasha: preliminary study. In J. Vymazal (Ed.), *Water and nutrient management in natural and constructed wetlands* (pp. 263–273). London: Springer Science+Business Media B.V.
Kiage, L. M., Liu, K.-B., Walker, N. D., Lam, N., & Huh, O. K. (2007). Recent land-cover/use change associated with land degradation in the Lake Baringo catchment, Kenya, East Africa: evidence from Landsat TM and ETM+. *International Journal of Remote Sensing, 28*(19), 4285–4309.
Klopp, J. M. (2012). Deforestation and democratization: patronage, politics and forests in Kenya. *Journal of Eastern African Studies, 6*(2), 351–370.
Kolios, S., & Stylios, C. D. (2013). Identification of land cover/land use changes in the greater area of the Preveza peninsula in Greece using Landsat satellite data. *Applied Geography, 40*, 150–160.
Krupnik, T. J. (2004). *Linking farmer, forest and watershed: agricultural systems and natural resources management along the upper Njoro River, Kenya*. University of California International and Area Studies Digital Collection. Article #3. Available online: http://repositories.cdlib.org/uciaspubs/articles/3. Accessed 27.12.12..
Lambin, E. F., Geist, H. J., & Lepers, E. (2003). Dynamics of land use and land cover changes in tropical regions. *Annual Reviews of Environmental Resources, 28*, 205–241.
Laurin, G. V., Liesenberg, V., Chen, Q., Guerriero, L., Del Frate, F., Bartolini, A., et al. (2013). Optical and SAR sensor synergies for forest and land cover mapping in a tropical site in West Africa. *International Journal of Applied Earth Observation and Geoinformation, 21*, 7–16.
Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote sensing and image interpretation* (6th ed.). USA: John Wiley and sons.
Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*(12), 2365–2401.
Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing, 28*(5), 823–870.
Lung, T., & Schaab, G. (2010). A comparative assessment of land cover dynamics of three protected forest areas in tropical eastern Africa. *Environmental Monitoring and Assessment, 161*, 531–548.
McCall, G. J. H. (1967). *Geology of the Nakuru-Thomson's falls–Lake Hannington area: Degree sheet No. 35, S.W. quarter and 43 N.W. quarter*. Report No. 78. Nairobi: Government Printer.
Mas, J. F. (1999). Monitoring land-cover changes: a comparison of change detection techniques. *International Journal of Remote Sensing, 20*(1), 139–152.
Mati, B. M., Mutie, S., Gadain, H., Home, P., & Mtalo, F. (2008). Impacts of land-use/cover changes on the hydrology of the trans-boundary Mara River, Kenya/Tanzania. *Lakes & Reservoirs: Research and Management, 13*, 169–177.
Millennium Ecosystem Assessment. (2005). *Ecosystems and human well-being: Synthesis*. Washington, DC: Island Press.
Mubea, K., & Menz, G. (2012). Monitoring land use change in Nakuru (Kenya) using multi-sensor satellite data. *Advances in Remote Sensing, 1*, 74–84.
Muriuki, G., Seabrook, L., McAlpine, C., Jacobson, C., Price, B., & Baxter, G. (2011). Land cover change under unplanned human settlements: a study of the Chyulu hills squatters, Kenya. *Landscape and Urban Planning, 99*(2), 154–165.
Mwetu, K. K., Mutua, B. M., Kundu, P. M., Fürst, J., & Loiskandl, W. (2009). Effects of deforestation and climate variability on river discharge in the Njoro river catchment, Kenya. In *Proceedings of the Sumawa Mau forest complex conference, Egerton University, 27–29th April 2009*.
Mwita, E., Menz, G., Misana, S., Becker, M., Kisanga, D., & Boehme, B. (2013). Mapping small wetlands of Kenya and Tanzania using remote sensing techniques. *International Journal of Applied Earth Observation and Geoinformation, 21*, 173–183.
Odada, E. O., Raini, J., & Ndetei, R. (2006). Lake Nakuru: experience and lessons learned brief. Available online: http://www.ilec.or.jp/eg/lbmi/pdf/18_Lake_Nakuru_27February2006.pdf. Accessed 27.11.12.
Olang, L. O., Kundu, P., Bauer, T., & Fürst, J. (2011). Analysis of spatio-temporal land cover changes for hydrological impact assessment within the Nyando river basin of Kenya. *Environmental Monitoring and Assessment, 179*, 389–401.
Otieno, V. O., & Anyah, R. O. (2012). Effects of land use changes on climate in the Greater Horn of Africa. *Climate Research, 52*, 77–95.
Overmars, K. P., & Verburg, P. H. (2005). Analysis of land use drivers at the watershed and household level: linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science, 19*(2), 125–152.

Palamuleni, L. G., Ndomba, P. N., & Annegarn, H. J. (2011). Evaluating land cover change and its impact on hydrological regime in Upper Shire river catchment, Malawi. *Regional Environmental Change, 11*, 845–855.

Pellikka, P. K. E., Lötjönen, M., Siljander, M., & Lens, L. (2009). Airborne remote sensing of spatio-temporal change (1955–2004) in indigenous and exotic forest cover in the Taita hills, Kenya. *International Journal of Applied Earth Observation and Geoinformation, 11*, 221–232.

Potter, C., Genovese, V., Gross, P., Boriah, S., Steinbach, M., & Kumar, V. (2007). Revealing land cover change in California with satellite data. *EOS, Transactions, American Geophysical Union, 88*(26), 269.

Raini, J. A. (2009). Impact of land use changes on water resources and biodiversity of Lake Nakuru catchment basin, Kenya. *African Journal of Ecology, 47*(1), 39–45.

Rodriguez-Galiano, V., & Chica-Olmo, M. (2012). Land cover change analysis of a Mediterranean area in Spain using different sources of data: multi-seasonal Landsat images, land surface temperature, digital terrain model and texture. *Applied Geography, 35*, 208–218.

Romero-Ruiz, M. H., Flantua, S. G. A., Tansey, K., & Berrio, J. C. (2012). Landscape transformations in savannas of northern South America: land use/land cover changes since 1987 in the Llanos Orientales of Colombia. *Applied Geography, 32*, 766–776.

Schmitt-Harsh, M. R. (2013). Landscape change in Guatemala: driving forces of forest and coffee agroforest expansion and contraction form 1990 to 2010. *Applied Geography, 40*, 40–50.

Seto, K. C., Woodcock, C. E., Song, C., Huang, X., Lu, J., & Kaufmann, R. K. (2002). Monitoring land use change in the Pearl river delta using Landsat TM. *International Journal of Remote Sensing, 23*(10), 1985–2004.

Shalaby, A., & Ali, R. R. (2010). Agricultural land monitoring in Egypt using NOAA-AVHRR and SPOT vegetation data. *Nature and Science, 8*(11), 275–278.

Shalaby, A., & Tateishi, R. (2007). Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the northwestern coastal zone of Egypt. *Applied Geography, 27*(1), 28–41.

Shivoga, W. A., Muchiri, M., Kibichi, S., Odanga, J., Miller, S. N., Baldyga, T. J., et al. (2007). Influences of land use/cover on water quality in the upper land middle reaches of river Njoro, Kenya. *Lakes & Reservoirs: Research and Management, 12*, 97–105.

Singh, A. (1989). Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing, 10*(6), 989–1003.

Stefanov, W. L., Ramsey, M. S., & Christensen, P. R. (2001). Monitoring urban land cover change: an expert system approach to land cover classification of semi-arid to arid urban centers. *Remote Sensing of Environment, 77*, 173–185.

Treitz, P., & Rogan, J. (2004). Remote sensing for mapping and monitoring land cover and land use change: an introduction. *Progress in Planning, 61*(4), 269–279.

Tsegaye, D., Moe, S. R., Vedeld, P., & Aynekulu, E. (2010). Land-use/cover dynamics in northern Afar rangelands, Ethiopia. *Agriculture, Ecosystems and Environment, 139*, 174–180.

UNEP. (2009). Kenya: atlas of our changing environment. *Division of Early Warning and Assessment (DEWA)*.

Wanjogu, S. N., Kibe, J. M., Wagate, P. N., & Mwangi, J. A. (2010). *Soil characteristics and land degradation in Lake Nakuru catchment*. Nairobi: Kenya Soil Survey.

Wasige, J. E., Groen, T. A., Smaling, E., & Jetten, V. (2013). Monitoring basin-scale land cover changes in Kagera Basin of Lake Victoria using ancillary data and remote sensing. *International Journal of Applied Earth Observation and Geoinformation, 21*, 32–42.

Weng, Q. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *Journal of Environmental Management, 64*, 273–284.

Wu, K., & Zhang, H. (2012). Land use dynamics, built-up land expansion patterns, and driving forces analysis of the fast-growing Hangzhou metropolitan area, eastern China (1978–2008). *Applied Geography, 34*, 137–145.

Yang, X., & Lo, C. P. (2002). Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. *International Journal of Remote Sensing, 23*(9), 1775–1798.

Yang, L., Xian, G., Klaver, J. M., & Deal, B. (2003). Urban land-cover change detection through sub-pixel imperviousness mapping using remotely sensed data. *Photogrammetric Engineering and Remote Sensing, 69*(9), 1003–1010.

Yang, X., & Liu, Z. (2005). Using satellite imagery and GIS for land-use and land-cover change mapping in an estuarine watershed. *International Journal of Remote Sensing*, http://dx.doi.org/10.1080/01431160500219224.

Yuan, F., Sawaya, K. E., Loeffelholz, B. C., & Bauer, M. E. (2005). Land cover classification and change analysis of the twin cities (Minnesota) metropolitan area by multi-temporal Landsat remote sensing. *Remote Sensing of Environment, 98*, 317–328.

# Paper II

**Were, K.O.,** Dick, Ø.B., Singh, B.R. (2014). Exploring the geophysical and socio-economic determinants of land cover changes in the Eastern Mau Forest Reserve and Lake Nakuru drainage basin, Kenya. *GeoJournal*, doi: 10.1007/s10708-014-9525-2.

# *Exploring the geophysical and socio-economic determinants of land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya*

**Kennedy Were, Øystein B. Dick & Bal R. Singh**

**GeoJournal**
Spatially Integrated Social Sciences and Humanities

ISSN 0343-2521

GeoJournal
DOI 10.1007/s10708-014-9525-2


Springer

GeoJournal
DOI 10.1007/s10708-014-9525-2

# Exploring the geophysical and socio-economic determinants of land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya

**Kennedy Were · Øystein B. Dick · Bal R. Singh**



**Abstract** Understanding the linkages between the biogeophysical and socio-economic processes that operate at different spatial and temporal scales is important for land cover change mitigation. This study analysed several factors that explained the forest-shrubland conversions, grassland conversions and cropland expansions in Lake Nakuru drainage basin and Eastern Mau forest reserve in Kenya from 1985 to 2011. Logistic regression models were developed using a combination of remote sensing-based land cover data, and geographical information systems-based geophysical and socio-economic data (i.e., temperature, rainfall, elevation, slope, aspect, topographic wetness, curvature, soil pH, soil cation exchange capacity (CEC), population density and distance to road, river and town). The results were varied; for example, in the period 1985–2000, forest-shrubland conversions were linked to distance to road), distance to town, soil pH, soil CEC, rainfall, topographic wetness, curvature and aspect. The same factors, in addition to slope and distance to river, also determined the likelihood of forest-shrubland conversions in the period 2000–2011. Overall, significance of the determining factors varied depending on time and nature of land cover change. For example, topographical factors influenced grassland conversions in the period 1985–2000, while soil-related factors did not. But in the period 2000–2011, the converse was true. Therefore, policies for restoration, conservation and sustainable management of critical ecosystems (e.g., forests) should be spatially targeted and time-specific. These results broaden our knowledge of land cover dynamics in this locality, and provide a base for effective environmental policy formulation, planning and management.

**Keywords** Land cover change · Driving factors · Spatio-temporal analysis · Logistic regression · GIS · Eastern Mau · Lake Nakuru

K. Were · Ø. B. Dick · B. R. Singh
Institute of Mathematical sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway

K. Were (✉)
Kenya Soil Survey, Kenya Agricultural Research Institute, P.O. Box 14733-00800, Nairobi, Kenya
e-mail: kenwerez@yahoo.com

## Introduction

Land cover change has become the focus of geographical research and discourse due to the unprecedented rates and magnitude of human alterations of the Earth's surface (Odada et al. 2009). The rising human population and the associated demand for food, fuel, water and shelter has been the key driver of these changes. Global extents of croplands, pastures, plantations and urban areas have expanded at the expense

Published online: 29 January 2014

of natural vegetation over the years. For instance, FAO (2011) reported that the world's forests shrunk by an average of 16 million ha per year in 1990s, but this reduced to 13 million ha per year in the last decade. In Africa, the net forest loss was estimated at 4 million ha per year between 1990 and 2000, but this reduced to 3.4 million ha per year between 2000 and 2010. Such land cover changes have significantly impacted on food security, climate, biodiversity, biogeochemical cycles, water availability, soil quality and human welfare (Foody 2002; Overmars and Verburg 2005; Heistermann et al. 2006; Potter et al. 2007). To mitigate these impacts through effective policies, it is important to understand the interactions between the biogeophysical and socio-economic processes, which operate at various spatial and temporal scales, leading to the land cover changes and land degradation (Overmars and Verburg 2005).

Theories, observations and models are the essential tools for enhancing our understanding of land cover change processes. Thus, modelling and prediction have been mainstreamed in land cover change research. The different modelling techniques that have been applied to analyse the processes, drivers and consequences of land cover changes can be grouped into: (1) empirical–statistical models, where the observed land cover changes and explanatory variables are analysed using multivariate statistics (e.g., logistic regression); (2) stochastic models, where the processes that move in a sequence of steps, through a set of states, are described (e.g., Markov chain models); (3) optimisation models, where economic techniques of optimal allocation of resources are applied (e.g., linear programming and general equilibrium models); (4) dynamic, process-based simulation models, where the biophysical and socio-economic processes are simulated and run systematically using some differential equations (e.g., cellular automata models), and; (5) integrated models where existing models are coupled with existing tools for spatially explicit evaluation and allocation of land resources (Lambin et al. 2000; Aspinall 2004; Heistermann et al. 2006; van Dessel et al. 2011).

Regression-based empirical models have been the most widely used in land cover change research for explanation of different processes of change and prediction of future changes (Millington et al. 2007). In particular, logistic regression has been used to explain the driving mechanisms of urban growth (Braimoh and Onishi 2007; Hu and Lo 2007; Huang et al. 2009; Wu et al. 2009; Li et al. 2013); deforestation, forest regrowth, expansion and fires (Chomitz and Gray 1996; Schneider and Pontius 2001; Munroe et al. 2004; Chowdhury 2006; Mertens et al. 2008; Crk et al. 2009; Wyman and Stein 2010; Badia et al. 2011; Müller et al. 2011, 2012; Muriuki et al. 2011; Schmitt-Harsh 2013); cropland abandonment, degradation and expansion (Jasinski et al. 2005; Overmars and Verburg 2005; Gellrich et al. 2007a, b; Rutherford et al. 2008; Serra et al. 2008; Lakes et al. 2009; López and Sierra 2010; Dubovyk et al. 2013; Prishchepov et al. 2013); and, grassland conversions and degradation (Li et al. 2012; Monteiro et al. 2011). Logistic regression has been popular because it: (1) can model binary and non-normally distributed response variables (e.g. land cover change); (2) does not assume linear relationships between the response and explanatory variables; (3) incorporates both categorical and continuous explanatory variables, and; (4) can be used directly for spatial prediction of future land cover changes (Millington et al. 2007; Serra et al. 2008; Martinez et al. 2011; Li et al. 2013). Regression-based land cover change modelling has also been stimulated by the improved availability of geographical information systems (GIS) and remotely sensed data, through government agencies, private enterprises, non-governmental organisations and communities.

In Kenya, the major watersheds have experienced rapid land cover changes for long periods. Lake Nakuru drainage basin, including Eastern Mau forest reserve in the Rift Valley system, is one of the hotspots that has been in a state of flux over the last three decades (Daniels and Bassett 2002; Baldyga et al. 2007). It is an important study area because Eastern Mau forest forms part of the largest closed-canopy montane forest ecosystem in Eastern Africa. The forest is also among the five important water catchment areas in Kenya and offers various ecosystem services. For instance, it is a major sink of $CO_2$, which is the major driver of global warming and climate change; hence, its role in the implementation of REDD[1] programme in Kenya is incontestable. Lake Nakuru drainage basin has been transformed from a sparsely populated and densely vegetated area to a highly populated, rapidly urbanizing and extensively

[1] Reduction of Emissions from Deforestation and forest Degradation.

cultivated area. Were et al. (2013) studied the dynamics of land cover patterns in this area from 1973 to 2011 and found that forests-shrublands and grasslands had decreased by about 428 and 258 km$^2$, respectively, at the annual average rates of 1 % each, while croplands and built-up lands had expanded by 660 and 24 km$^2$, at the annual rates of 6 and 16 %, respectively. Other studies in the area also revealed that these land cover changes had affected the water quality (Shivoga et al. 2007), hydrological regime (Mwetu et al. 2009), temperature distribution (Hesslerová and Pokorný 2010) and biodiversity (Kibichii et al. 2007; Raini 2009). However, no studies have been designed to analyse the linkage between the land cover change processes and the surrounding socio-economic and biogeophysical conditions. This is largely attributed to shortage of appropriate spatial data and analytical tools. Such analyses are of great importance for the development of effective strategies to alleviate the impacts of the on-going land cover changes, restore and sustainably manage land resources in the area.

This study built on the work of Were et al. (2013) by applying binary logistic regression and geographic information techniques in a spatially explicit framework to analyse the geophysical and socio-economic factors that related to and explained the variations in the three major land cover change processes (i.e., forest-shrubland conversions, grassland conversions and cropland expansions) observed in Lake Nakuru drainage basin between 1985 and 2011. It was hypothesized that different geophysical and socio-economic factors with complex relationships could explain the variations in the foregoing land cover conversions.

## Materials and methods

### Study area

The study area is Lake Nakuru drainage basin and Eastern Mau forest reserve, which covers about 2,000 km$^2$. It lies along the Kenyan Rift Valley system, between latitudes 0°10′–0°45′S and longitudes 35°40′–36°5′E (Fig. 1), with the altitude ranging from 1,750 to 3,090 m above sea level. Makalia, Enjoro, Naishi, Lamuriak, Enderit and Ngosorr river systems drain down the area into Lake Nakuru, whereas Nessuiet and Rongai rivers flow into Lake Bogoria and Baringo, respectively. The landforms include mountains and major scarps, hills and minor scarps, plateaus, volcanic footridges, uplands, volcanic and lacustrine plains and bottomlands. The major soils are Andosols, Planosols, Vertisols, Nitisols, Regosols, Calcisols, Solonetz and Phaeozems (Wanjogu et al. 2010), of which the parent materials originated from volcanic rocks (e.g., basalts, trachyte, phonolites, pumice tuff and lavas) and associated sediments of tertiary–quaternary age (McCall 1967). The climate varies from cool and humid to hot and humid depending on the altitude and topography. Higher areas at Mau escarpment receive substantial rainfall (~1,200 mm), which decreases notably (~700 mm) on the lower areas around Lake Nakuru. The rainfall pattern is bimodal with the long rains falling between March and May and short rains between November and December due to the seasonal north–south movement of the inter-tropical convergence zone (Odada et al. 2006). The vegetation comprises grasslands and scrublands in the lower parts, acacia trees along the lakeshore, riverine vegetation along the river courses, and forests in the higher areas. The major land use systems, which also contribute to the human and national economy, are agriculture, ranching, pastoralism, forestry, urban and industrial centres, and tourism and wildlife conservation. Land ownership is varied with the government mainly owning the national park and forest reserves, subsistence farmers owning the small-scale farms, and commercial farmers/ranchers leasing the large-scale farms/ranches. Lake Nakuru drainage basin is also an important centre for a constantly growing human population. Currently, it has over 1.5 million inhabitants with over 300,000 living in the rapidly expanding Nakuru Municipality (www.opendata.go.ke/). The rest live in small towns, market centres and rural settlements.

### Data sources, geoprocessing and geodatabase development

The data used in this study are summarized in Table 1, while details of data preparation and creation of a geodatabase are provided in the following subsections.

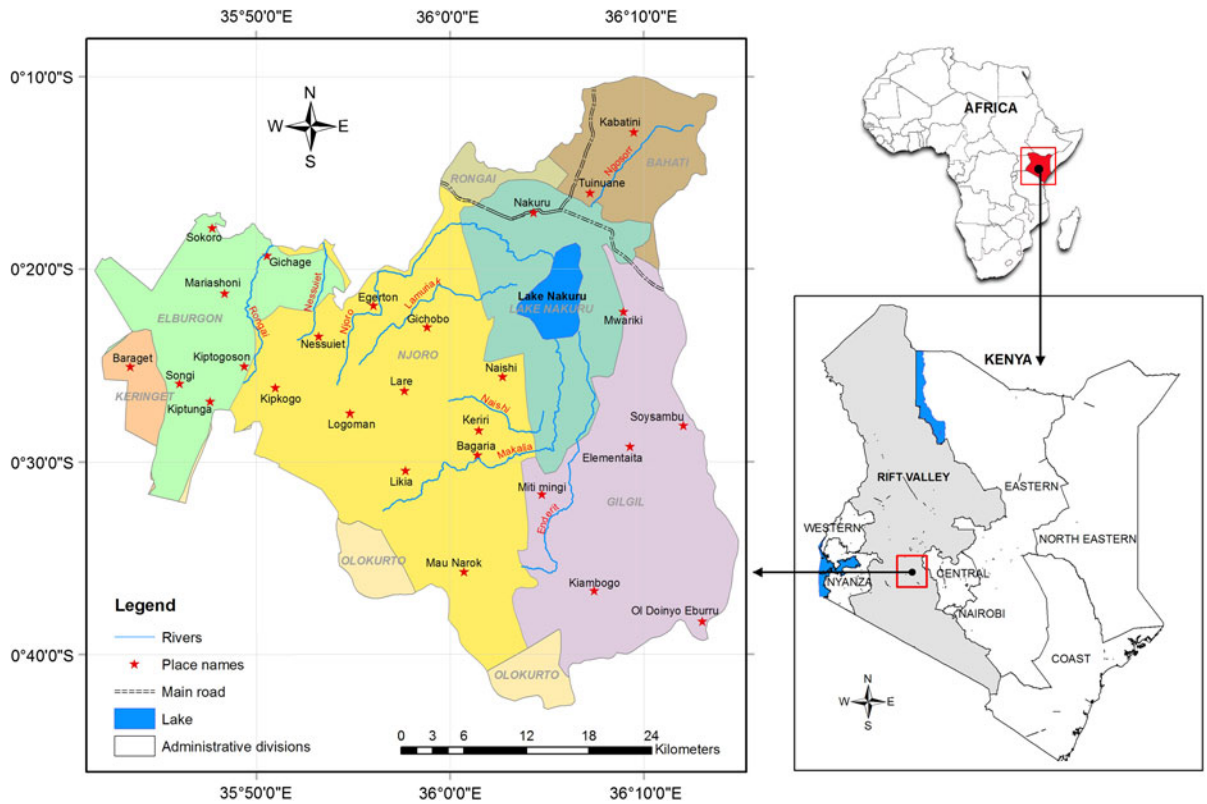



**Fig. 1** Geographical location of the study area

*Response variables*

The land cover change maps of the study area formerly produced by Were et al. (2013) for the periods 1985–2000 and 2000–2011 were used. These maps were produced through classification of multi-temporal Landsat TM (1985 and 2011) and ETM+ (2000) imagery using partitioning, hybrid classification and spatial reclassification methods. These yielded three land cover maps (i.e., for 1985, 2000 and 2011) each showing the six main land cover types; namely, forests-shrublands (i.e., woody vegetation), grasslands (i.e., herbaceous vegetation), croplands, built-up lands, bare lands and water bodies. The land cover maps were validated and overlaid in post-classification comparison to detect the land cover changes between 1985–2000 and 2000–2011. The accuracies of the maps for 1985, 2000 and 2011 were 95, 80 and 89 %, respectively; hence, the accuracies of the resultant land cover change maps were 76 % for 1985–2000 and 71 % for 2000–2011 (i.e., products of overall accuracies of the respective land cover maps). The three important processes of land cover changes that were detected are: (1) forest-shrubland conversions to other land cover types; (2) grassland conversions to other land cover types, and; (3) cropland expansions from other land cover types. These formed the basis for preparing the maps of response variables (i.e., the presence or absence of land cover change). Since the land cover change information was dichotomous (i.e., change vs no-change), binary maps for the three significant land cover change processes at each period were produced. The maps highlighted areas of: (1) forest conversions versus stable forests; (2) grassland conversions versus stable grasslands, and; (3) conversions to croplands versus stable croplands. The map pixels that represented areas of change were coded 1, while those that represented areas of no-change were coded 0. Overall, six binary maps of response variables were created.

**Table 1** Summary of the data used for logistic regression modelling

| Variable | Unit | Data type | Proxy for | Source |
|---|---|---|---|---|
| *Response* | | | | |
| Forest-shrub conversions | 0–1 | Binary | | Classification of multi-temporal Landsat imagery (Were et al. 2013) |
| Grassland conversions | 0–1 | Binary | | |
| Cropland expansions | 0–1 | Binary | | |
| *Geophysical* | | | | |
| 1. Rainfall | mm | Continuous | Water availability | www.worldclim.org |
| 2. Temperature | °C | Continuous | Warmth | www.worldclim.org |
| 3. Elevation | m | Continuous | Climatic elements, land form | DEM (http://srtm.csi.cgiar.org) |
| 4. Slope | ° | Continuous | Drainage, erosion hazard | DEM |
| 5. Aspect | ° | Continuous | Exposure to the Sun | DEM |
| 6. Curvature | – | Continuous | Drainage | DEM |
| 7. Topographic wetness index | – | Continuous | Drainage, soil moisture | DEM |
| 8. Soil pH | 0–14 | Continuous | Soil quality | Kenya Soil Survey |
| 9. Soil CEC | cmol/kg | Continuous | Soil quality | Kenya Soil Survey |
| *Socio-economic* | | | | |
| 10. Distance to road | km | Continuous | Land accessibility | www.ilri.org/gis; Survey of Kenya |
| 11. Distance to town | km | Continuous | Market accessibility | www.ilri.org/gis; Survey of Kenya |
| 12. Distance to river | km | Continuous | Water accessibility | www.ilri.org/gis; Survey of Kenya |
| 13. Population density (1989 and 2009) | Persons/sq. km | Continuous | Population pressure | Kenya National Bureau of statistics via www.opendata.go.ke |

*Explanatory variables*

A suite of thirteen potential explanatory variables (Table 1) were selected a priori based on existing theories of land use, fieldwork experience, data availability and literature review (Chomitz and Gray 1996; Geist and Lambin 2002; Lambin et al. 2003; Braimoh and Vlek 2005; Aguiar et al. 2007). These were categorized as either geophysical or socio-economic variables according to the natural or accessibility conditions of the environment. Policy, governance and institutional factors (e.g. land tenure, political actions, etc.) had also influenced processes (e.g., deforestation) that drive land cover changes in the area. But, they were not included in the analyses due to lack of spatial dataset that represented them. The geophysical variables were rainfall, temperature, soil cation exchange capacity (CEC), soil pH, elevation, slope, aspect, topographic wetness index (TWI) and curvature. Climate data were obtained from www.worldclim.org, soil data from Kenya Soil Survey and digital elevation model (DEM) from CGIAR consortium for spatial information (http://srtm.csi.cgiar.org/). Primary and secondary terrain attributes (i.e., slope and aspect, TWI and curvature) were extracted from the DEM.

The socio-economic variables comprised distance to road, distance to river, distance to town and population density. Feature datasets including rivers, roads and towns were obtained from the International Livestock Research Institute database (www.ilri.org/gis), while population density data for 1989 and 2009 census were from the Kenya open database (www.opendata.go.ke/). Additional towns, rivers and roads were also digitized on-screen from digital topographical maps. Proximity to the road, river or town was calculated as Euclidean distance.

Each feature dataset was rasterized; thereafter, all the raster maps were clipped according the extents of the study area, and transformed to the Universal Transverse Mercator projection system, Zone 37S, WGS 84 ellipsoid and datum (Fig. 2). Lastly, the data

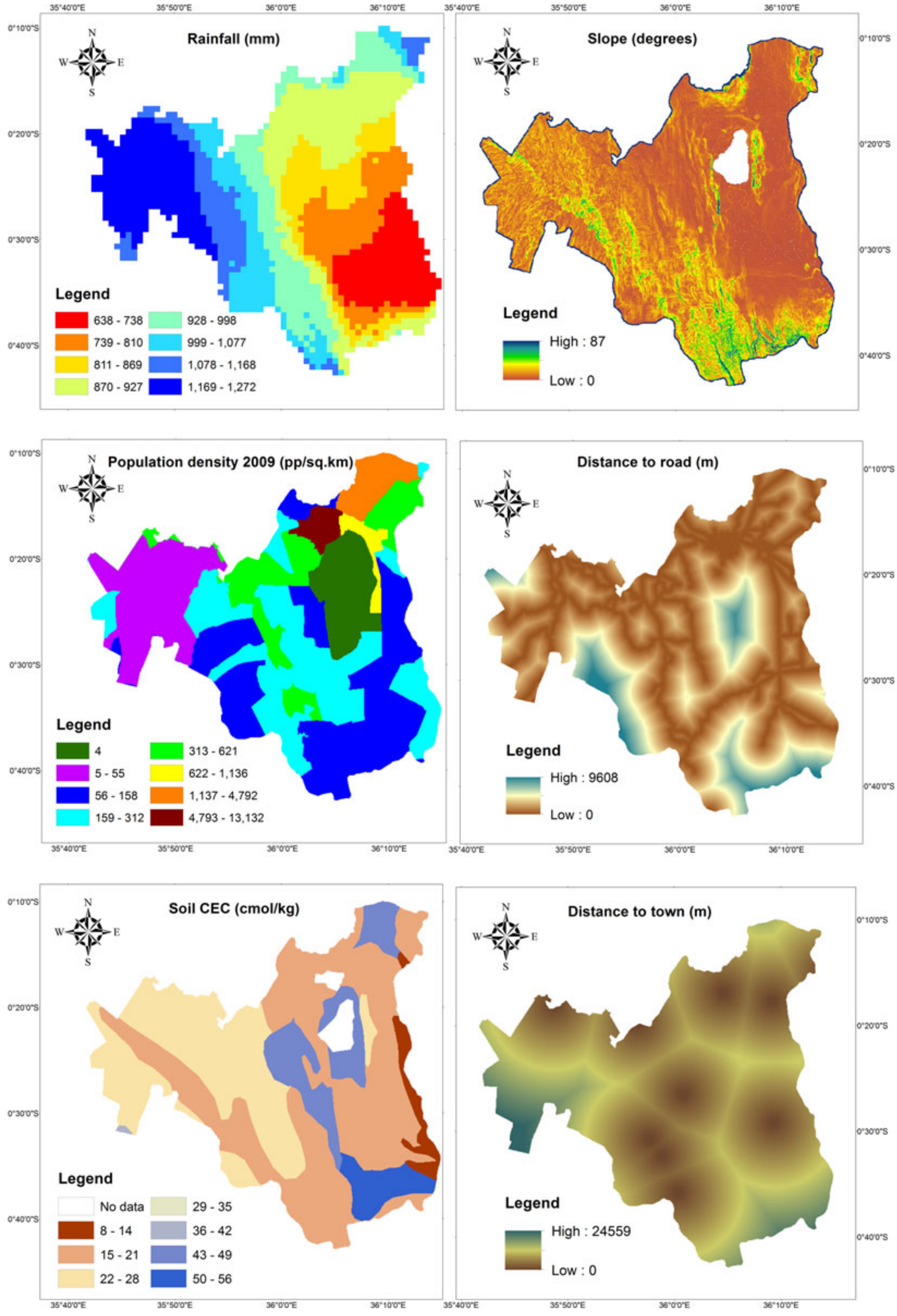



**Fig. 2** Maps showing the spatial distribution of some of the explanatory variables

were resampled to 100 m, and a geodatabase built. The geoprocessing tasks were performed using ArcGIS® 10.1, SAGA and ERDAS imagine® 2011 software.

## Statistical modelling framework

### *Sampling design*

Spatially balanced random points were sampled from the six binary maps of response variables; thus, six sets of sample points were created. The total sample points per set were equivalent to 0.5 % of the pixels constituting the respective binary map. Thereafter, the sample points, coupled with the maps of explanatory and response variables, were overlaid in order to extract attribute values from the map pixels to the corresponding sample points. The attributes also included the X and Y coordinates for each sample point to ensure that the subsequent analyses were spatially explicit, and also to facilitate the testing for spatial autocorrelation in the model residuals. Each set of sample points was cleaned (e.g., by deleting spurious values) and, subsequently, used for model training.

### *Exploratory data analysis*

Once the dataset for model training were prepared, pairwise Pearson's product-moment correlation analysis was conducted to detect potential multi-collinearity. Multi-collinearity can cause inefficient parameter estimates and inaccurate measures of statistical significance. If the Pearson's correlation coefficient between two explanatory variables exceeded 0.8, only one of them was retained for model building (Menard 2002). An explanatory variable was also excluded from a model if its variance inflation factor (VIF) exceeded 10 (Montgomery et al. 2006). Descriptive statistics, i.e., the means and standard deviations of the explanatory variables, were also estimated prior to model building.

### *Multiple logistic regression modelling*

The response variable was categorical and dichotomous; so, binary logistic regression was the most suited method to model the probability of occurrence of each land cover change process given the set of explanatory variables in Table 1 (Müller et al. 2011). The formula for logistic regression model is:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \propto + \sum_{i=1}^{n} \beta_i x_i + \varepsilon \tag{1}$$

where $\pi(x)$ is the probability that the response variable ($y$) equals 1, $\propto$ is the constant (or intercept), $\beta$ is the vector of estimated coefficients of the explanatory variables $x_i$, and $\varepsilon$ is the error term. The ratio $\frac{\pi(x)}{1-\pi(x)}$ is called the *odds*, while $\log\left(\frac{\pi(x)}{1-\pi(x)}\right)$ is called the *log-odds* or *logit* transformation of $\pi(x)$. After back-transformation, the response variables are expressed as conditional probabilities, in the interval [0, 1], as follows:

$$\widehat{\pi}(x) = \frac{e^{\propto + \sum_{i=1}^{n} \beta_i x_i}}{1 + e^{\propto + \sum_{i=1}^{n} \beta_i x_i}} \tag{2}$$

Full models were fitted using the maximum likelihood estimator and later reduced by step-wise backward elimination method of variable selection. Several full models were tested in the process. The significance of the model parameters and logistic regression models were determined by *Wald statistic* and *likelihood ratio statistic* respectively, while the six simplest models with adequate fit to the data were selected based on Akaike information criterion (AIC) (Agresti 2007). The simplest model was the one that minimized:

$$AIC = -2(loglikelihood - number\ of\ parameters\ in\ the\ model) \tag{3}$$

Model adequacy was also checked using residual plots, measures of influence and leverage (e.g., Cook's $D$), Pearson $\chi^2$ statistic and pseudo-$R^2$ (i.e., Nagelkerke-$R^2$). Spatial autocorrelation in the residuals of the models was measured using the Global Moran's $I$ (Overmars et al. 2003). The Moran's $I$ values ranged between $-1$ and $+1$ with positive values indicating positive autocorrelation and vice versa, and values close to 0 signifying low spatial autocorrelation (i.e., low clustering of similar residuals). Lastly, the discriminatory power of the models was summarized by the area under the receiver operating characteristic (ROC) curve (AUC), based on predicted conditional probabilities from tenfold cross-validation procedure. Using this unbiased procedure, the sample points were split into 10 equal-sized folds; thereafter, each fold was randomly selected for testing the models fitted to the other ninefolds of the data. This way, each observation was predicted once (as an out-of-sample prediction), and the prediction errors calculated. The ROC curve plotted the sensitivity (or true positive rate) against 1-specificity (or false positive rate) for a range of cut-offs that could be applied to interpret $\widehat{\pi}(x)$ as actual land cover change. The AUC showed the probability of the models ranking a randomly chosen positive instance (i.e., change) higher than a randomly chosen negative instance (i.e., no-change). A measure of 0.5 indicated random performance, while a measure of 1 indicated perfect performance. Detailed explanations of ROC curves have been provided elsewhere (Fawcett 2006; Pontius and Schneider 2001). The statistical procedures were conducted using R version 2.15.0 (R Development Core Team 2012) and Minitab® 16.

## Results

### Exploratory data analysis

Correlation between the explanatory variables was generally low and never exceeded 0.7, except for rainfall and temperature (−0.75), elevation and rainfall (0.79), and elevation and temperature (−0.98) (Table 2). The highly correlated variables were not used simultaneously in model parameterization, in order to avoid multi-collinearity problems. Absence of multi-collinearity is also evidenced by the low VIFs (<10) of the variables that were used to estimate the parameters of the models (Tables 3, 4).

### Models of land cover changes between 1985 and 2000

Table 3 presents the logit models of forest-shrubland conversions, grassland conversions and cropland expansions for the period 1985–2000. It provides the parameter estimates, odds ratios and VIF of the explanatory variables in each model. The table also gives the strength of spatial autocorrelation, goodness-of-fit, significance and discriminatory power of the models. Figure 3 illustrates the performance of the three models.

**Table 2** Pearson's correlation coefficients between the explanatory variables

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Rainfall | 1.00 | | | | | | | | | | | | | |
| 2. Temperature | **−0.75** | 1.00 | | | | | | | | | | | | |
| 3. Elevation | **0.79** | **−0.98** | 1.00 | | | | | | | | | | | |
| 4. Slope | −0.06 | −0.08 | 0.06 | 1.00 | | | | | | | | | | |
| 5. Aspect | 0.03 | −0.02 | 0.04 | −0.00 | 1.00 | | | | | | | | | |
| 6. Curvature | −0.01 | −0.00 | 0.01 | 0.66 | 0.00 | 1.00 | | | | | | | | |
| 7. TWI | −0.12 | 0.28 | −0.29 | −0.64 | −0.01 | −0.28 | 1.00 | | | | | | | |
| 8. Soil pH | −0.38 | 0.41 | −0.43 | 0.03 | 0.00 | 0.00 | 0.01 | 1.00 | | | | | | |
| 9. Soil CEC | −0.24 | 0.29 | −0.30 | −0.06 | −0.04 | −0.00 | 0.07 | 0.60 | 1.00 | | | | | |
| 10. DIST_road | −0.18 | −0.18 | 0.14 | 0.17 | 0.02 | 0.02 | −0.21 | 0.13 | 0.10 | 1.00 | | | | |
| 11. DIST_town | 0.17 | −0.28 | 0.29 | 0.09 | 0.16 | 0.01 | −0.13 | −0.04 | 0.10 | 0.27 | 1.00 | | | |
| 12. DIST_river | −0.17 | 0.02 | −0.03 | 0.07 | 0.09 | 0.07 | −0.03 | −0.12 | −0.18 | 0.18 | 0.08 | 1.00 | | |
| 13. Pop. density '89 | −0.16 | 0.34 | −0.32 | −0.05 | 0.02 | 0.00 | 0.09 | 0.09 | −0.03 | −0.19 | −0.23 | −0.05 | 1.00 | |
| 14. Pop. density '09 | −0.05 | 0.14 | −0.12 | −0.03 | −0.00 | −0.01 | 0.05 | 0.02 | −0.10 | −0.14 | -0.16 | −0.06 | 0.67 | 1.00 |

Highly correlated explanatory variables ($r > 0.8$) are in bold

**Table 3** Summary statistics of the logistic regression models for 1985–2000

| Explanatory variable | Forest-shrubland conversions | | | Grassland conversions | | | Cropland expansions | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF |
| (Intercept) | −4.216*** | – | – | 6.362*** | – | – | 10.938*** | – | – |
| Rainfall | 0.002*** | 1.00 | 1.37 | −0.003*** | 1.00 | 1.61 | ni | – | – |
| Temperature | ni | – | – | ni | – | – | −0.851*** | 0.43 | 1.95 |
| Elevation | ni | – | – | ni | – | – | ni | – | – |
| Slope | ns | – | – | 0.021· | 1.02 | 2.65 | ns | – | – |
| Aspect | −0.002*** | 1.00 | 1.02 | −0.002*** | 1.00 | 1.06 | 0.002*** | 1.00 | 1.05 |
| Curvature | 0.069*** | 1.07 | 1.08 | −0.121** | 0.89 | 1.74 | ns | – | – |
| TWI | 0.355*** | 1.43 | 1.14 | −0.203*** | 0.82 | 1.95 | −0.075* | 0.93 | 1.28 |
| Soil pH | −0.287*** | 0.75 | 1.90 | ns | – | – | 0.497*** | 1.64 | 1.43 |
| Soil CEC | 0.061*** | 1.06 | 1.87 | ns | – | – | −0.034*** | 0.97 | 1.33 |
| DIST_road | −0.342*** | 0.71 | 1.09 | −0.161*** | 0.85 | 1.32 | 0.505*** | 1.66 | 1.27 |
| DIST_town | −0.183*** | 0.83 | 1.23 | 0.039 ** | 1.04 | 1.17 | 0.036* | 1.04 | 1.48 |
| DIST_river | ns | – | – | −0.317*** | 0.73 | 1.42 | 0.362*** | 1.44 | 1.36 |
| POP_89 | ns | – | – | ns | – | – | 0.001*** | 1.00 | 1.26 |
| *N* | 3,870 | | | 1,890 | | | 2,902 | | |
| Pr > LR $\chi^2$ | 0.00 | | | 0.00 | | | 0.00 | | |
| Pr > Pearson $\chi^2$ | 0.52 | | | 0.41 | | | 0.71 | | |
| Nagelkerke $R^2$ | 0.38 | | | 0.20 | | | 0.36 | | |
| AUC | 0.81 | | | 0.72 | | | 0.80 | | |
| Moran's I | 0.16 | | | 0.18 | | | 0.19 | | |

Significance level 0 = ***, 0.001 = **, 0.01 = * and 0.05 = .; ni = not included due to high correlation with either elevation, temperature or rainfall; ns = not significant; VIF = variance inflation factor; *N* = number of observations; LR = likelihood ratio, and; AUC = area under curve

*Forest-shrubland conversions model*

According to Table 3, the effects of rainfall, soil pH, soil CEC, TWI, aspect, curvature and distance to road and town on forest-shrubland conversions were significant, while the effects of slope, population density and distance to river were not. An increase in TWI, curvature and soil CEC by 1 unit multiplied the odds of forest-shrubland conversions by 1.43, 1.07 and 1.06, respectively, while an increase in distance to road, distance to town and soil pH by 1 unit multiplied the odds by 0.71, 0.83 and 0.75, respectively. Even though the effects of rainfall and aspect were significant, the magnitude of their effects was negligible. Their odds ratios (i.e., 1 each) also indicate that increasing them by 1 unit neither increased nor decreased the likelihood of forest-shrubland conversions. The *P* value for the likelihood ratio statistic is small (0.00) and the *P* value for the Pearson $\chi^2$ statistic is large (0.52); thus, there is evidence that the forest-shrubland conversions model was significant and an adequate fit to the data. The high AUC value (0.81) indicates very good discrimination between the presence and absence of forest-shrubland conversions by the model, while the low Moran's I (0.16) indicates very weak positive spatial autocorrelation in the response. The Nagelkerke-$R^2$ shows that the model explained 38 % of variability in the presence of forest-shrubland conversions.

*Grassland conversions model*

The results reveal that grassland conversions were explained largely by factors related to climate, drainage and accessibility; soil and demographic factors were not important (Table 3). The probability of grassland conversions decreased by 27 %, 15 %, 18 % and 11 % for every 1 unit increase in distance to river, distance to road, TWI and curvature, respectively, and increased by 4 % and 2 % for every 1 unit increase in distance to town and slope, respectively. Just like in

**Table 4** Summary statistics of the logistic regression models for 2000–2011

| Explanatory variable | Forest-shrubland conversions | | | Grassland conversions | | | Cropland expansions | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF | Coeff. | Odds ratio | VIF |
| (Intercept) | 4.714*** | – | – | 9.461*** | – | – | −7.558*** | – | – |
| Rainfall | −0.002*** | 1.00 | 1.54 | ni | – | – | 0.008*** | 1.01 | 1.30 |
| Temperature | ni | – | – | −0.471*** | 0.62 | 1.86 | ni | – | – |
| Elevation | ni | – | – | ni | – | – | ni | – | – |
| Slope | −0.020** | 0.98 | 1.55 | ns | – | – | ns | – | – |
| Aspect | −0.002*** | 1.00 | 1.05 | ns | – | – | 0.002*** | 1.00 | 1.05 |
| Curvature | 0.039* | 1.04 | 1.49 | ns | – | – | −0.097*** | 0.91 | 1.08 |
| TWI | ns | – | – | ns | – | – | −0.297*** | 0.74 | 1.35 |
| Soil pH | −0.391*** | 0.68 | 2.61 | −0.306* | 0.74 | 3.12 | 0.827*** | 2.29 | 2.05 |
| Soil CEC | 0.054*** | 1.06 | 2.71 | 0.061*** | 1.06 | 3.70 | −0.009*** | 0.99 | 1.80 |
| DIST_road | −0.278*** | 0.76 | 1.22 | −0.528*** | 0.59 | 1.40 | −0.119*** | 0.89 | 1.53 |
| DIST_town | −0.086*** | 0.92 | 1.34 | −0.041* | 0.96 | 1.44 | 0.060*** | 1.06 | 1.72 |
| DIST_river | 0.077** | 1.08 | 1.26 | −0.244*** | 0.78 | 1.48 | 0.217*** | 1.24 | 1.29 |
| POP_09 | ns | – | – | ns | – | – | −0.003*** | 1.00 | 1.13 |
| *N* | 3,408 | | | 1,618 | | | 3,898 | | |
| Pr > LR $\chi^2$ | 0.00 | | | 0.00 | | | 0.00 | | |
| Pr > Pearson $\chi^2$ | 0.13 | | | 0.08 | | | 0.76 | | |
| Nagelkerke $R^2$ | 0.21 | | | 0.49 | | | 0.43 | | |
| AUC | 0.74 | | | 0.86 | | | 0.84 | | |
| Moran's I | 0.19 | | | 0.13 | | | 0.14 | | |

Significance level 0 = ***, 0.001 = **, 0.01 = * and 0.05 = .; ni = not included due to high correlation with either elevation, temperature or rainfall; ns = not significant; VIF = variance inflation factor; *N* = number of observations; LR = likelihood ratio, and; AUC = area under curve

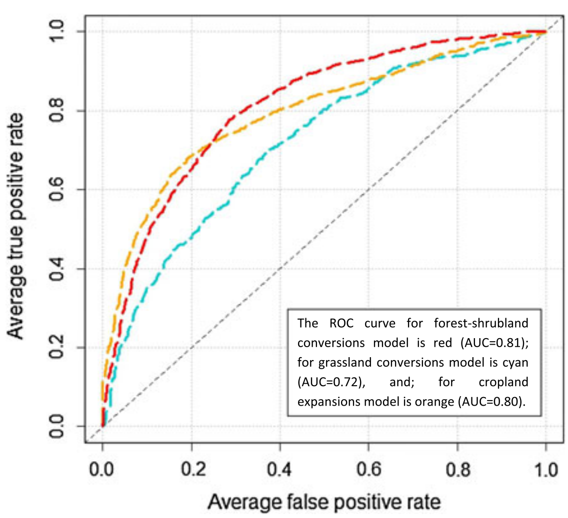



**Fig. 3** ROC curves illustrating the performance of the models for 1985–2000

the forest-shrubland conversions model, rainfall and aspect were significant, but the magnitude of their effects was rather small. The grassland conversions model was highly significant as shown by the likelihood ratio statistic (*P* value = 0.00), and fitted the data well as shown by the Pearson $\chi^2$ statistic (*P* value = 0.41). It was also successful in discriminating the actual grassland conversions given the high value of AUC (0.72), and exhibited positive spatial autocorrelation with low intensity given the low value of Moran's statistic (0.18). The Nagelkerke-$R^2$ indicates that the model accounted for 20 % of variability in the presence of grassland conversions.

*Cropland expansions model*

Climatic, demographic, accessibility, drainage and soil factors were significant determinants of the presence of cropland expansions (Table 3). For 1 unit increase in distance to road, distance to river, distance to town and soil pH, the croplands had 1.66, 1.44, 1.04 and 1.64 more chance of expansion, respectively. By contrast, the chance was only 0.43, 0.93 and 0.97 for each additional unit of temperature, TWI and soil CEC respectively. Despite their significance, the magnitude of the effects of population density and aspect was very small; their odds ratios also attest to this. The *P* values of the likelihood ratio statistic (0.00) and Pearson $\chi^2$ statistic (0.71) show that the cropland conversions model was highly significant with a good fit to the data. Its power to discriminate the presence of cropland expansions was very good (AUC = 0.80), while the strength of residual spatial autocorrelation was very weak (Moran's I = 0.19). The Nagelkerke-$R^2$ shows that the model explained 36 % of variability in the presence of cropland expansions.

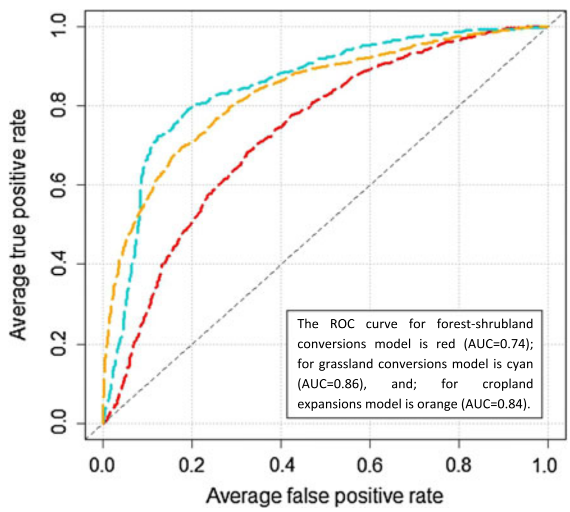



**Fig. 4** ROC curves illustrating the performance of the models for 2000–2011

## Models of land cover changes between 2000 and 2011

Table 4 provides the resultant models of the three major land cover processes for the period 2000–2011, while Fig. 4 displays the results of their performance.

### *Forest-shrubland conversions model*

Table 4 shows that rainfall, soil pH, soil CEC, slope, aspect, curvature and distance to road, town and river were important in determining the occurrence of forest-shrubland conversions in the second period. Contrary to the previous period, distance to river and slope were also significant determinants, while TWI was not. The signs of the effects of the factors that were significant in both periods (i.e., rainfall, soil properties, aspect, curvature, and distance to road and town) did not change except for rainfall (Tables 3, 4). However, the magnitudes of the effects of soil CEC, curvature and distance to road and town decreased, while for soil pH increased over time. Distance to river, soil CEC and curvature had an increasing effect on the odds of forest-shrubland conversions; that is, they multiplied the odds by 1.08, 1.06 and 1.04, respectively, for every 1 unit increase. On the contrary, soil pH, slope, distance to road and distance to town had a decreasing effect on the odds; that is, they multiplied the odds by 0.68, 0.98, 0.76 and 0.92, respectively, for every 1 unit increase. Rainfall and aspect were significant, but like in the previous period, the likelihood of forest-cropland conversions was not affected by their increases as evidenced by the small coefficients and odds ratios of 1 each. The *P* values of the likelihood ratio statistic and Pearson $\chi^2$ statistic were 0.00 and 0.13, respectively; hence, there was no evidence to question the significance and adequacy of the model. The AUC value of 0.74 indicated that its power to discriminate the actual cropland expansions was satisfactory, while Moran's statistic of 0.19 indicated that spatial autocorrelation was positive but very weak. The Nagelkerke-$R^2$ shows that the model accounted for 21 % of variability in the presence of forest-shrubland conversions.

### *Grassland conversions model*

The significant factors for explaining grassland conversions in the second period included: temperature, soil pH, soil CEC and distance to road, river and town (Table 4). Unlike before, soil properties were significant, whereas curvature, aspect and slope were not significant. The signs of the effects of the factors that were significant in both periods (i.e., accessibility factors) remained the same except for distance to town (Tables 3, 4). However, the magnitudes of the effects of distance to river decreased, while for distance to road increased over time. Soil CEC increased the likelihood of grassland conversions by 6 % for every 1 unit increase, while temperature, soil pH, distance to road, distance to town and distance to river decreased the likelihood by 38 %, 26 %, 41 %, 4 % and 22 %, respectively, for every 1 unit increase. The $P$ values of the likelihood ratio statistic and Pearson $\chi^2$ statistic were 0.00 and 0.08, respectively; hence, the model was significant and a good fit to the data. The high AUC value (0.86) also implied that its discriminatory power was very good, while the low Moran's I (0.13) indicated positive spatial autocorrelation with low intensity. The Nagelkerke-$R^2$ shows that the model explained 49 % of variability in the presence of grassland conversions.

*Cropland expansions model*

Rainfall, soil pH, soil CEC, aspect, curvature, TWI, population density and distance to road, river and town were the important variables for explaining cropland expansions in the second period (Table 4). Curvature also became an important factor during this period. The signs of the effects of TWI, aspect, soil pH, soil CEC and, distance to town and river were the same as in the first period, while for population density and distance to road changed (Tables 3, 4). Further, the magnitudes of the effects of soil CEC and distance to river decreased, while for TWI, soil pH and distance to town increased over time. Soil pH, rainfall, distance to town and distance to river increased the odds of cropland expansions 2.29, 1.01, 1.06 and 1.24 times, respectively, for every 1 unit increase. Conversely, TWI, curvature, soil CEC and distance to road decreased the probability by 26 %, 9 %, 1 % and 11 %, respectively, for every 1 unit increase. Though population density and aspect were significant, the odds ratios and coefficients show that their effects were small. The small $P$ value of the likelihood ratio statistic (0.00) shows that the model of cropland expansion was highly significant, while the large $P$ value of the Pearson $\chi^2$ statistic (0.76) shows that it fitted the data well. Its discriminatory power was also very good given the high value of AUC value (0.84), and the strength of spatial structuring was very weak given the low Moran's coefficient (0.14). The Nagelkerke-$R^2$ reveals that the model explained about 43 % of variability in the presence of cropland expansions.

## Discussion

The results revealed that forest-shrubland conversions, grassland conversions and cropland expansions in Lake Nakuru drainage basin and Eastern Mau forest reserve stemmed from a combination of geophysical and socio-economic factors. The logit model of forest-shrubland conversions between 1985 and 2000 indicated that as rainfall and soil CEC increased, the probability of forest-shrubland conversions increased, but as distance to road and town increased, the converse was true (Table 3). This was because most of the forest-shrubland conversions that occurred at this time were on higher areas on the western side (Were et al. 2013), which were characterized by high rainfall, low temperatures and fertile soils. The areas included: Baraget, Mariashoni, Sokoro, Gichage, Nessuiet, Logoman, Likia and Sururu. The good road network, especially on the western side, provided access to inputs and markets for charcoal, logs, fuel wood and timber in Nakuru, Njoro and Elburgon towns. As pointed out by Chomitz and Gray (1996), proximity to roads also lowered the cost of migration, land access and land clearance for subsistence farming. In the second period, the signs of the effects of the explanatory variables on probability of forest-shrubland conversions were maintained except for rainfall (Table 4). The change in the direction of rainfall effect was expected because the forest-shrubland conversions had diffused to the southern parts (e.g., Kiambogo) and around Lake Nakuru National park where rainfall was relatively lower. These findings are comparable to those of Müller and Mburu (2009) who modelled the *hotspots* of forest clearance in Kakamega forest, Kenya using artificial neural networks.

In theory, many factors could potentially explain the forest-shrubland conversions, but in practice, only a few factors could be captured by the models due to lack of spatial data (especially of socio-economic nature) and inability to quantify some variables (e.g., political interference). This might partly explain the low Nagelkerke-$R^2$ of the models. But again, Hosmer and Lemeshow (2000, in Gellrich et al. 2007a) maintained that low $R^2$ values are the norm for logistic regression models; hence, they should be interpreted with caution and not strictly by the standards of goodness-of-fit in linear regression analysis. Some of the socio-economic and political processes that influenced the forest-shrubland conversions, but lacked proxies in the models include: (a) government actions (e.g., excision of about 353 $km^2$ of Eastern Mau forests in 2001 for human settlement); (b) patronage politics where votes from certain communities and political patrons were rewarded with allocation of forest resources, especially with the advent of multi-partyism in 1990s (Klopp 2012); (c) fluctuations of prices at the international oil markets leading to high costs of alternative energy (e.g., liquefied petroleum gas); (d) poverty (e) shortage of human and financial resources for sustainable management and monitoring of forests; (f) technological advancements (e.g., use of power saws, tractors and lorries for logging and transportation); (g) poor environmental governance manifested through abuse of power, impunity and corruption (h) weak policy, legal and regulatory framework (e.g., the forest excisions and resettlement schemes in 1990s were not subjected to environmental impact assessments until the enactment of the Environmental Management and Co-ordination Act 1999), and; (i) Individual behaviour (i.e., some people practiced or abetted illegal logging and charcoal burning because of greed and ignorance (or unconcern) about the problems of deforestation). It was assumed that omission of these factors did not compromise the validity of the models; however, their inclusion in further work as the data become available is encouraged in order to gain a holistic understanding of the land cover change processes.

Moreover, grassland conversions model for the period 1985–2000 (Table 3) showed that soil quality factors were not significant determinants of the probability of grassland conversions. This is because, though the largest share of grasslands were converted to croplands (Were et al. 2013), most of the conversions occurred in the lower areas with lower soil CEC. This implies that soil fertility was more dependent on fertilization than natural soil richness in these areas. In the second period, soil quality factors became significant since the grassland conversions were mostly present in the higher areas with higher soil CEC. Accessibility factors were also important in determining the likelihood of grassland conversions in both periods. Between 1985 and 2000, the likelihood of grassland conversions decreased with increasing distance to road and river, but increased with increasing distance to town. This suggests that water and land accessibility were more important in agricultural land use decision-making than market accessibility at that time. The small-scale farmers did not attach potential land rents to most of the agricultural land uses as theorized by von Thünen (see Chomitz and Gray 1996). This is quite common in developing countries where subsistence is the overriding goal of agricultural production. In case of surpluses, the middlemen often collect and transport the produce from the farm-gates to the market. By contrast, between 2000 and 2011, market accessibility also played an important role in agricultural land use decision-making according to the model. This is because the bulk of grasslands at this time were formerly part of Eastern Mau forest on the western side (i.e., Mariashoni, Gichage, Nessuiet and Sokoro), which had been converted partly due to their proximity to road and town (Table 3). Their proximity to Elburgon, Njoro and Nakuru towns, among other factors, attracted enterprising individuals who hired these parcels of land and cultivated food crops (e.g., potatoes and maize) for sale. Few studies have analyzed the determinants of grassland conversions in the region. Serneels and Lambin (2001) who conducted a similar study found that soil quality and market accessibility explained rangeland conversions between 1985 and 1995 in Narok district, Kenya. This deviates from the findings here for the period 1985–2000. But, such deviations can be attributed to the different contexts and spatial scales of analyses.

Additionally, distance to road was among the significant factors that explained cropland expansions for the periods 1985–2000 and 2000–2011. However, the direction of the effect during the first period was unexpected. A negative estimate was expected because the total gains from forest-shrubland and grassland conversions were highest for croplands, and increasing distance to road decreased the chances of occurrence of these conversions (Table 3). The same applies to the effects of soil properties during both periods. It is acknowledged that such unexpected signs in the model parameters may have emanated from the omission of other important explanatory variables as mentioned earlier, data errors (e.g., classification and measurement errors) and data inconsistencies (i.e., data of different types, sources and scale). For instance, van Dessel et al. (2011) found that the significance and signs of parameters in the logistic regression models they calibrated varied with the accuracy level of the land cover classifications. These limitations may have also caused the weak spatial autocorrelation detected in the residuals (Dormann

2007a). Though spatial autocorrelation also affects coefficients and inference in geographical modelling (Overmars et al. 2003; Dormann 2007b), it was neither filtered nor incorporated because the quantitative evidence (Tables 3, 4) was considered too weak to bias the results. Generally, it is important to control these uncertainties when modelling in order to achieve realistic results.

The revealed relationships between the three land cover change processes and the geophysical and socio-economic factors extend our knowledge of land cover dynamics in Lake Nakuru drainage basin and Eastern Mau forest reserve where empirical evidence of the underlying causal factors is scarce. This knowledge can be applied, for instance, in spatial predictions of possible future trends of land cover changes. Such predictions, in addition to the knowledge, are beneficial for environmental policy makers, planners and managers since they can: (1) inform selection of priority areas for targeted policies or detailed analyses in an effective and efficient manner; and, (2) be linked with biophysical data, e.g., species distribution or carbon storage data, to identify *hotspots* of biodiversity or carbon losses following land cover conversions. In future, this modelling approach can be improved through incorporation of spatial non-stationarity and other important environmental factors, as they become available, for detailed modelling of land cover changes in other parts of Kenya.

## Conclusions and recommendations

Using logistic regression, remote sensing and GIS-based data, this study analysed the geophysical and socio-economic factors that related to and explained the variations in the three major land cover change processes (i.e., forest-shrubland conversions, grassland conversions and cropland expansions) in Lake Nakuru basin and Eastern Mau forest reserve between 1985 and 2011. The analysis revealed that a combination of climatic, topographic, soil quality, demographic and accessibility factors determined the likelihood of these land cover change processes as hypothesized. It also highlighted that the relationships among these factors varied depending on the land cover change process and time. This calls for spatially targeted and time-specific policies for conservation and sustainable management of critical ecosystems (e.g., forests). The findings improve our understanding of land cover dynamics in this locale and provide a basis for spatial prediction of land cover change risks, as well as effective environmental policy formulation, planning and management. This is also critical in climate change mitigation, especially in the context of REDD and carbon sequestration. Such analysis may further assist planning and management efforts in other areas experiencing dramatic land cover changes. In future, studies should incorporate other important environmental data as they become available, for holistic appreciation of the drivers and refinement of the current models. The effects of the drivers per se may be spatially non-stationary; thus, application of local regression methods (e.g., geographically weighted logistic regression) to analyse the spatially varying relationships would also be interesting. Finally, future studies should also go further and analyse the impacts of land cover change on ecosystems services by linking the model outputs with pertinent biophysical data, e.g., species distribution or carbon stocks data to identify the *hotspots* of biodiversity or carbon sequestration, respectively.

**Acknowledgments** This study was funded by the Research Council of Norway through the Norwegian University of Life Sciences. We highly appreciate the anonymous reviewers for improving the quality of the manuscript through their constructive criticism.

## References

Agresti, A. (2007). *An introduction to categorical data analysis*. New Jersey: Wiley.

Aguiar, A. P. D., Câmara, G., & Escada, M. I. S. (2007). Spatial statistical analysis of land use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. *Ecological Modelling, 209*(2–4), 169–188.

Aspinall, R. (2004). Modelling land use change with generalized linear models: A multi-model analysis of change between 1860 and 2000 in Gallatin Valley, Montana. *Journal of Environmental Management, 72*(1–2), 91–103.

Badia, A., Serra, P., & Modugno, S. (2011). Identifying dynamics of fire ignition probabilities in two representative Mediterranean wildland–urban interface areas. *Applied Geography, 31*(3), 930–940.

Baldyga, T., Miller, S. N., Driese, K. L., & Gichaba, C. M. (2007). Assessing land cover change in Kenya's Mau forest region using remotely sensed data. *African Journal of Ecology, 46*(1), 46–54.

Braimoh, A. K., & Onishi, T. (2007). Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy, 24*(2), 502–515.

Braimoh, A. K., & Vlek, P. L. G. (2005). Land cover change trajectories in northern Ghana. *Environmental Management, 36*(3), 356–373.

Chomitz, K. M., & Gray, D. A. (1996). Roads, land use and deforestation: A spatial model applied to Belize. *The World Economic Review, 10*(3), 487–512.

Chowdhury, R. R. (2006). Landscape change in the Calakmul Biosphere Reserve, Mexico: Modeling the driving forces of smallholder deforestation in land parcels. *Applied Geography, 26*(2), 129–152.

Crk, T., Uriarte, M., Corsi, F., & Flynn, D. (2009). Forest recovery in a tropical landscape: What is the relative importance of biophysical, socioeconomic, and landscape variables? *Landscape Ecology, 24*(5), 629–642.

Daniels, R., & Bassett, T. J. (2002). The spaces of conservation and development around Lake Nakuru national park, Kenya. *The Professional Geographer, 54*(4), 481–490.

Dormann, C. F. (2007a). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography, 30*(5), 609–628.

Dormann, C. F. (2007b). Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography, 16*(2), 129–138.

Dubovyk, O., Menz, G., Conrad, C., Kan, E., Machwitz, M., & Khamzina, A. (2013). Spatio-temporal analyses of cropland degradation in the irrigated lowlands of Uzbekistan using remote sensing and logistic regression modelling. *Environmental Monitoring and Assessment, 185*(6), 4775–4790.

FAO. (2011). *Status of the world's forests 2011*. ISBN 978-92-5-106750-5. http://www.fao.org/docrep/013/i2000e/i2000e00.htm. Accessed 4 May 2013.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*(1), 185–201.

Geist, H. J., & Lambin, E. F. (2002). Proximate causes and underlying driving forces of tropical deforestation. *BioScience, 52*(2), 143–150.

Gellrich, M., Baur, P., Koch, B., & Zimmermann, N. E. (2007a). Agricultural land abandonment and forest regrowth in the Swiss mountains: A spatially explicit economic analysis. *Agriculture, Ecosystems & Environment, 118*(1-4), 93–108.

Gellrich, M., Baur, P., & Zimmermann, N. E. (2007b). Natural forest regrowth as a proxy variable for agricultural land abandonment in the Swiss mountains: A spatial statistical model based on geophysical and socio-economic variables. *Environmental Modelling and Assessment, 12*(4), 269–278.

Heistermann, M., Müller, C., & Ronneberger, K. (2006). Land in sight? Achievements, deficits and potentials of continental to global scale land use modelling. *Agriculture, Ecosystems & Environment, 114*(2-4), 141–158.

Hesslerová, P., & Pokorný, J. (2010). Effect of Mau forest clear cut on temperature distribution and hydrology of catchment of Lakes Nakuru and Naivasha: preliminary study. In: J. Vymazal (Ed.), *Water and nutrient management in natural and constructed wetlands*. doi:10.1007/978-90-481-9585-5_19.

Hu, Z., & Lo, C. P. (2007). Modelling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems, 31*(6), 667–688.

Huang, B., Zhang, L., & Wu, B. (2009). Spatio-temporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science, 23*(3), 379–398.

Jasinski, E., Morton, D., & DeFries, R. (2005). Physical landscape correlates of the expansion of mechanized agriculture in Mato Grosso, Brazil. *Earth Interactions, 9*(16), 1–18.

Kibichii, S., Shivoga, W. A., Muchiri, M., & Miller, S. N. (2007). Macro-invertebrate assemblages along a land-use gradient in the upper River Njoro watershed of Lake Nakuru drainage basin, Kenya. *Lakes & Reservoirs: Research and Management, 12*(2), 107–117.

Klopp, J. M. (2012). Deforestation and democratization: Patronage, politics and forests in Kenya. *Journal of Eastern African Studies, 6*(2), 351–370.

Lakes, T., Müller, D., & Krüger, C. (2009). Cropland change in southern Romania: A comparison of logistic regressions and artificial neural networks. *Landscape Ecology, 24*(9), 1195–1206.

Lambin, E. F., Geist, H. J., & Lepers, E. (2003). Dynamics of land use and land cover changes in tropical regions. *Annual Reviews of Environmental Resources, 28*, 205–241.

Lambin, E. F., Rounsevell, M. D. A., & Geist, H. J. (2000). Are agricultural land use models able to predict changes in land-use intensity? *Agriculture, Ecosystems & Environment, 82*(1-3), 321–331.

Li, S., Verburg, P. H., Lv, S., Wu, J., & Li, X. (2012). Spatial analysis of the driving factors of grassland degradation under the conditions of climate change and intensive use in inner Mongolia, China. *Regional Environmental Change, 12*(3), 461–474.

Li, X., Zhou, W., & Ouyang, Z. (2013). Forty years of urban expansion in Beijing: What is the relative importance of physical, socio-economic and neighborhood factors? *Applied Geography, 38*, 1–10.

López, S., & Sierra, R. (2010). Agricultural change in the Pastaza River Basin: A spatially explicit model of native Amazonian cultivation. *Applied Geography, 30*(3), 355–369.

Martinez, J. Á., Suárez-seoane, S., & Calabuig, E. L. (2011). Modelling the risk of land cover change from environmental and socio-economic drivers in heterogeneous and changing landscapes: The role of uncertainty. *Landscape and Urban Planning, 101*(2), 108–119.

McCall, G. J. H. (1967). Geology of the Nakuru-Thomson's Falls-Lake Hannington area: degree sheet No. 35, S.W. Quarter and 43 N.W. Quarter, Report No. 78. Government Printer, Nairobi.

Menard, S. (2002). *Applied logistic regression analysis, quantitative applications in the social sciences, No. 106*. London: Sage.

Mertens, B., Hua, L., Belcher, B., Ruiz-Pérez, M., Maoyi, F., & Xiaosheng, Y. (2008). Spatial patterns and processes of bamboo expansion in southern China. *Applied Geography, 28*(1), 16–31.

Millington, J. D. A., Perry, G. L. W., & Romero-Calcerrada, R. (2007). Regression techniques for examining land use/

cover change: A case study of a Mediterranean landscape. *Ecosystems, 10*(4), 562–578.
Monteiro, A. T., Fava, F., Hiltbrunner, E., Marianna, G. D., & Bocchi, S. (2011). Assessment of land cover changes and spatial drivers behind loss of permanent meadows in the lowlands of Italian Alps. *Landscape and Urban Planning, 100*(3), 287–294.
Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introduction to linear regression analysis*. New Jersey: Wiley.
Müller, D., & Mburu, J. (2009). Forecasting hotspots of forest clearing in Kakamega forest, western Kenya. *Forest Ecology and Management, 257*(3), 968–977.
Müller, R., Müller, D., Schierhorn, F., & Gerold, G. (2011). Spatio-temporal modelling of the expansion of mechanized agriculture in the Bolivian lowland forests. *Applied Geography, 31*(2), 631–640.
Müller, R., Müller, D., Schierhorn, F., Gerold, G., & Pacheco, P. (2012). Proximate causes of deforestation in the Bolivian lowlands: An analysis of spatial dynamics. *Regional Environmental Change, 12*(3), 445–459.
Munroe, D. K., Southworth, J., & Tucker, C. M. (2004). Modelling spatially and temporally complex land-cover change: The case of western Honduras. *The Professional Geographer, 56*(4), 544–559.
Muriuki, G., McAlpine, C., Seabrook, L., & Baxter, G. (2011). The role of squatters in retention of native vegetation: A case study of the Chyulu Hills, Kenya. *Applied Geography, 31*(2), 577–589.
Mwetu, K. K., Mutua, B. M., Kundu, P. M., Fürst, J., & Loiskandl, W. (2009). Effects of deforestation and climate variability on river discharge in the Njoro river catchment, Kenya. *Proceedings of the Sumawa Mau forest complex conference*, Egerton University, 27–29th April 2009.
Odada, E. O., Ochola, W. O., & Olago, D. O. (2009). Drivers of ecosystem change and their impact on human well-being in Lake Victoria basin. *African Journal of Ecology, 47*(1), 46–54.
Odada, E. O., Raini, J., & Ndetei, R. (2006). *Lake Nakuru: Experience and lessons learned brief* [online]. http://www.worldlakes.org/uploads/18_Lake_Nakuru_27February2006.pdf. Accessed 10 January 2014.
Overmars, K. P., de Koning, G. H. J., & Veldkamp, A. (2003). Spatial autocorrelation in multi-scale land use models. *Ecological Modelling, 164*(2–3), 257–270.
Overmars, K. P., & Verburg, P. H. (2005). Analysis of land use drivers at the watershed and household level: Linking two paradigms at the Philippine forest fringe. *International Journal of Geographical Information Science, 19*(2), 125–152.
Pontius, R. G., & Schneider, L. C. (2001). Land cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment, 85*(1–3), 239–248.
Potter, C., Genovese, V., Gross, P., Boriah, S., Steinbach, M., & Kumar, V. (2007). Revealing land cover change in California with satellite data. *EOS, Transactions, American Geophysical Union, 88*(26), 269.
Prishchepov, A. V., Müller, D., Dubinin, M., Baumann, M., & Radeloff, V. C. (2013). Determinants of agricultural land abandonment in post-Soviet European Russia. *Land Use Policy, 30*(1), 873–884.
R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0. http://www.R-project.org/.
Raini, J. A. (2009). Impact of land use changes on water resources and biodiversity of Lake Nakuru catchment basin, Kenya. *African Journal of Ecology, 47*(1), 39–45.
Rutherford, G. N., Bebi, P., Edwards, P. J., & Zimmermann, N. E. (2008). Assessing land use statistics to model land cover change in a mountainous landscape in the European Alps. *Ecological Modelling, 212*(3–4), 460–471.
Schmitt-Harsh, M. (2013). Landscape change in Guatemala: Driving forces of forest and coffee agro-forest expansion and contraction from 1990 to 2010. *Applied Geography, 40*, 40–50.
Schneider, L. C., & Pontius, R. G. (2001). Modelling land use change in the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems & Environment, 85*(1–3), 83–94.
Serneels, S., & Lambin, E. F. (2001). Proximate causes of land-use change in Narok District, Kenya: A spatial statistical model. *Agriculture, Ecosystems & Environment, 85*(1–3), 65–81.
Serra, P., Pons, X., & Saurí, D. (2008). Land cover and land use change in a Mediterranean landscape: A spatial analysis of driving forces integrating biophysical and human factors. *Applied Geography, 28*(3), 189–209.
Shivoga, W. A., Muchiri, M., Kibichii, S., Odanga, J., Miller, S. N., & Baldyga, T. J. (2007). Influences of land use/cover on water quality in the upper land middle reaches of river Njoro, Kenya. *Lakes & Reservoirs: Research and Management, 12*(2), 97–105.
van Dessel, W., van Rompaey, A., & Szilassi, P. (2011). Sensitivity analysis of logistic regression parameterization for land use and land cover probability estimation. *International Journal of Geographical Information Science, 25*(3), 489–508.
Wanjogu, S. N., Kibe, J. M., Wagate, P. N., & Mwangi, J. A. (2010). *Soil characteristics and land degradation in Lake Nakuru catchment*. Nairobi: Kenya Soil Survey.
Were, K. O., Dick, Ø. B., & Singh, B. R. (2013). Remotely sensing the spatial and temporal land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya. *Applied Geography, 41*, 75–86.
Wu, B., Huang, B., & Fung, T. (2009). Projection of land use change patterns using kernel logistic regression. *Photogrammetric Engineering & Remote Sensing, 75*(8), 971–979.
Wyman, M. S., & Stein, T. V. (2010). Modelling social and land use/land cover change data to assess drivers of smallholder deforestation in Belize. *Applied Geography, 30*(3), 329–342.

# Paper III

**Were, K.O.,** Singh, B.R., Dick, Ø.B. (2014). Effects of land cover changes on soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya (Chapter 6). In: Lal, R., Singh, B.R., Mwaseba, D.L., Kraybill, D., Hansen, D.O., Eik, L.O. (eds.), *Sustainable intensification to advance food security and enhance climate resilience in Africa*. Springer International Publishing, Switzerland, doi: 10.1007/978-3-319-09360-4_6 (in press).

# EFFECTS OF LAND COVER CHANGES ON SOIL ORGANIC CARBON AND TOTAL NITROGEN STOCKS IN THE EASTERN MAU FOREST RESERVE, KENYA

K.O. Were[1,2], B.R. Singh[3], Ø.B. Dick[1]

**Abstract**

This study analysed the variations of soil organic carbon (SOC) and total nitrogen (TN) stocks under natural forests (NF), plantation forests (PF), bamboo forests (BF), and croplands that had been converted from such forests (i.e., NF2C, PF2C and BF2C) in the Eastern Mau Forest Reserve using field, laboratory, spatial, and statistical techniques. The results displayed significant differences in SOC and TN stocks between NF and NF2C ($p$<0.0001), and between PF and PF2C ($p$<0.0001). For instance, the surface soils (0-15cm) of NF had the highest SOC and TN stocks (71.6 and 7.1 Mg ha$^{-1}$, respectively), while NF2C had the lowest (35.4 and 3.5 Mg ha$^{-1}$). Similarly, the subsurface soils (15-30cm) of NF had the highest stocks (55.7 and 5.6 Mg ha$^{-1}$), while NF2C had the lowest (32.5 and 3.2 Mg ha$^{-1}$). This reflects a decline in both SOC and TN stocks by about 51% in the surface and about 42% in the subsurface soils after NF conversion. There were also significant differences in SOC and TN stocks ($p$<0.05) between the surface and subsurface soils of different land cover types. The stocks decreased as soil depth increased. This trend suggests that (i) forest-to-cropland conversions are undermining the ecosystem's capacity for carbon sequestration, and (ii) subsurface soils have potential for carbon sequestration. SOC and TN losses in the croplands may be mitigated by adopting best management practices (BMPs), especially agro-forestry. These findings are useful for designing sustainable land management (SLM) and carbon sequestration projects.

**Keywords:** Land cover changes • soil organic carbon • total nitrogen • soil carbon sequestration • Eastern Mau • Kenya

## 1.0 Introduction

Soil organic carbon (SOC) comprises organic compounds (i.e., plant, animal and microbial residues at all stages of decay) that are highly enriched in carbon (Lal 2008; Post and Kwon 2000; Solomon et al. 2000). SOC is a major determinant of the physical, chemical, and biological properties that are necessary for soil's proper functioning. For example, SOC ensures soil quality by supplying nutrients, enhancing cation exchange capacity (CEC), supporting biodiversity, and improving soil aggregation and water-holding capacity (Bationo et al. 2007). The quantity of SOC varies spatially and temporally because of multiple climatic, edaphic, biotic (flora, fauna and humans), topographical, and lithological

[1] Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway

[2] Kenya Agricultural Research Institute, Kenya Soil Survey, P.O. Box 14733-00800, Nairobi, Kenya

[3] Department of Environmental Sciences, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway

factors, which influence the balance between the gains and losses of soil carbon. However, the greatest carbon fluxes between the atmosphere and the Earth surface are attributed to anthropogenic factors, including land use and land cover changes (IPCC 2013). Consequently, land cover change is a core theme of climate change research, which emphasizes the understanding of SOC responses to land cover dynamics. This is because the world's soils contain about 1500 Pg C to 1m depth, while the atmosphere contains about 750 Pg C, and the vegetation contains about 610 Pg C (Smith 2004, 2008; Lal 2004). Therefore, even slight changes in the SOC pool can significantly affect the global carbon cycle, climate, and soil properties (Powlson et al. 2011). This indicates that soil carbon sequestration is a potential strategy to mitigate climate change through reduction of $CO_2$ emissions as required by the United Nation Framework Convention for Climate Change.

Many studies have reported that converting natural vegetation such as forests or grasslands to arable land impinges on soil carbon storage and fluxes (Demessie et al. 2013; Jafarian and Kavian 2013; Muñoz-Rojas et al. 2012; Biro et al. 2011; Don et al. 2011; Jiao et al. 2009; Awiti et al. 2008; Wang et al. 2008; Yimer et al. 2007; Evrendilek et al. 2004; Jing-Cheng et al. 2004; Powers 2004; Osher et al. 2003; Murty et al. 2002; Islam and Weil 2000; Solomon et al. 2000; Brown and Lugo 1990). Such conversions invariably result in SOC losses and $CO_2$ emissions because of the attendant changes in quality and quantity of biomass carbon inputs, accelerated decomposition of soil organic matter (SOM), leaching of dissolved organic carbon (DOC), and loss of particulates through mechanical clearing, water, and wind (Powlson et al. 2011; Detwiler 1986). However, the ultimate direction, magnitude, and rate of changes in SOC after land cover conversions depend on the initial carbon content of the soil, method of land clearance, terrain, soil type, climate, time since conversion, changes in the microbial community and nitrogen cycling, chemical properties of the litter, and land management practices (Vågen et al. 2005; Murty et al. 2002; Brown and Lugo 1990).

In Kenya, human actions have brought dramatic changes to many crucial ecosystems; for example, the Eastern Mau Forest Reserve. It has experienced wanton destruction and degradation since the 1990s thanks to illegal logging, encroachments, and charcoal burning (Fig. 1), as well as ill-advised political decisions, particularly the excision of ~61,023ha for human settlement in 2001 (Government of Kenya 2009; UNEP 2009). The Eastern Mau Forest Reserve is an important study site because it constitutes part of the largest closed-canopy indigenous montane forest in Eastern Africa, and is also one of Kenya's five key water catchment areas, which offers various ecosystem services, such as carbon sequestration, micro-climate regulation, ground water recharge, water storage, flood mitigation, etc. (UNEP 2009). Previous studies of land cover change in the area employed satellite remote sensing and GIS techniques (Were et al. 2013; Baldyga et al. 2007) to map the hotspots; that is, places where forests and shrublands had been converted to croplands. Although studies on the impacts on soil properties at such hotspots have been conducted globally, little attention has been paid to Eastern Africa. In particular, there is insufficient knowledge of how SOC has responded to deforestation and forest degradation in the Eastern Mau Forest Reserve. This study was designed to address these issues. The results will improve our understanding of SOC dynamics, the ecosystem's productivity and its role in climate change, as well as our capacity to monitor and predict carbon fluxes. This is essential to formulate realistic and effective policies for sustainable land management (SLM) and climate change mitigation.

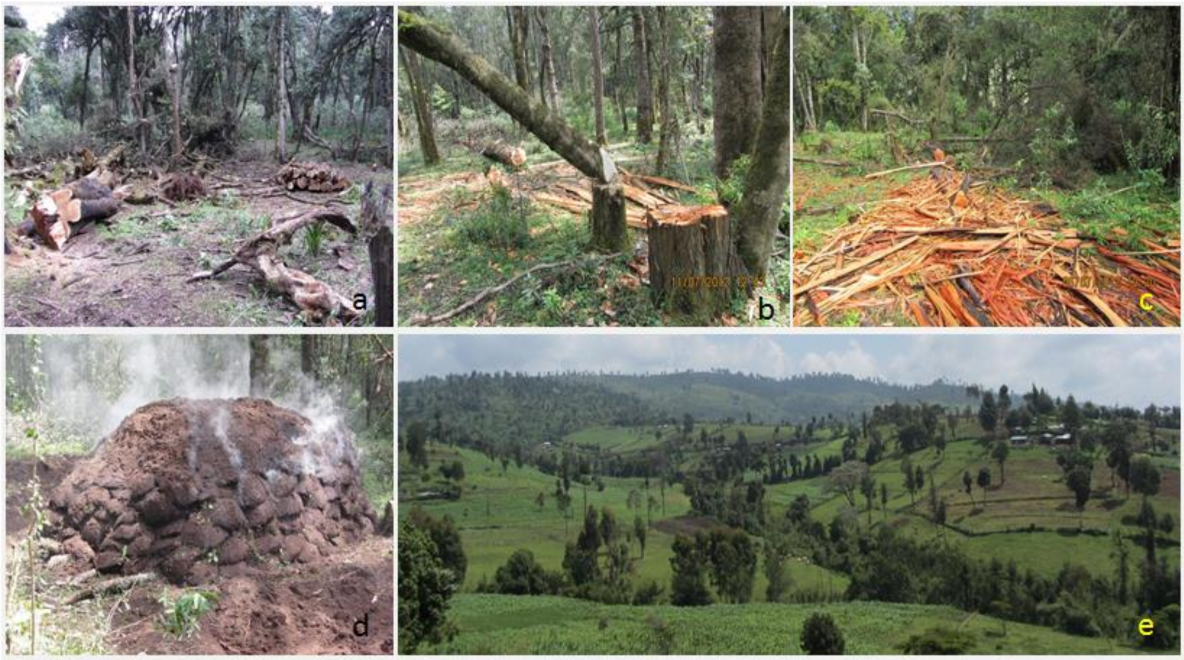



**Fig. 1:** Human activities in the Eastern Mau Forest Reserve: (**a**, **b**, and **c**) illegal felling of trees; (**d**) charcoal burning; and (**e**) agricultural expansion and human settlement (Source: Author).

In this study, we analysed the effects of forest to cropland conversions on SOC, total nitrogen (TN), and bulk density (BD) in the Eastern Mau Forest Reserve. We then recommended the best management practices (BMPs) for SOC sequestration based on our results. We included TN in the study because of the intricate linkage between soil C and N cycles. The study's guiding hypothesis was that SOC, TN, and BD varied significantly between (i) forests and cropland establishments, and (ii) surface (0-15cm) and subsurface soils (15-30cm) of the land cover types.

## 2.0 Materials and methods

### 2.1 Study area

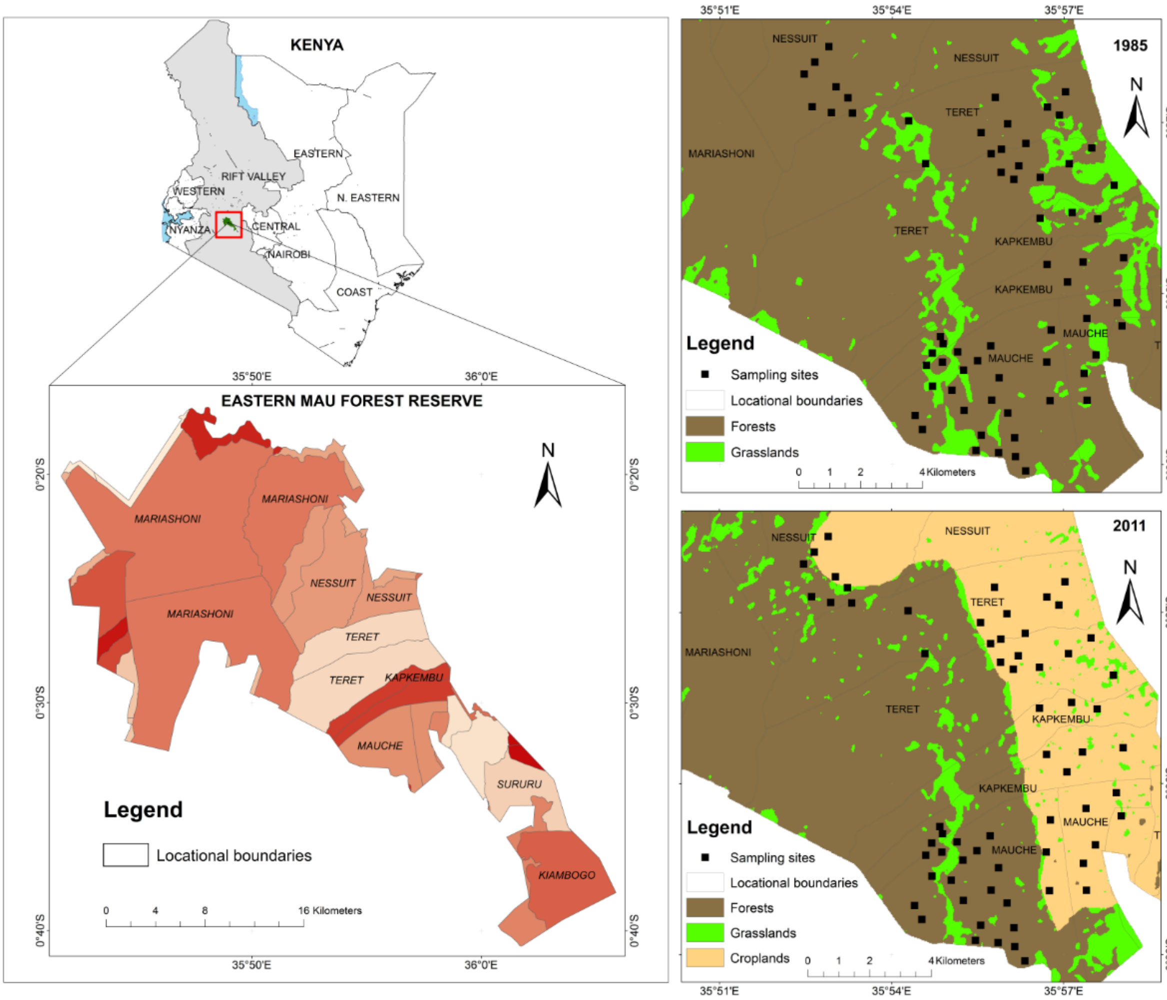



**Fig. 2:** Geographical location of the study area and sampling points. The sampling points are superimposed on the land cover maps of 1985 and 2011 (on the right) that were derived from digital classification of Landsat TM imagery

The study area covered Nessuiet, Teret, Kapkembu, and Mauche locations in the Eastern Mau Forest Reserve defined by the latitudes 0º 15´- 0º 40´S and longitudes 35º 40´- 36º 10´E (Fig. 2), and the altitudes ranging from 2210 to 3070m above sea level. The climate is cool and humid, with the mean annual rainfall varying between 935 and 1287 mm, and the mean annual temperature ranging from 9.8 to 17.5 °C (Jaetzold et al. 2010). The rainfall pattern is tri-modal with peaks in April, August, and November. The Njoro, Naishi, and Larmudiac Rivers drain the eastern slopes into Lake Nakuru, while the Nessuiet flows into Lake Bogoria, and the Rongai River into the Baringo. The area's physiography and lithology are characterized by major scarps and uplands comprising pyroclastic rocks, such as Pumice tuffs, of tertiary-quaternary volcanic age. These soft, light brown rocks have insets of yellow pumice and angular trachyte, which decompose into deep to very deep, dark reddish brown clayey soil aggregates (McCall 1967). The soils, classified as *Mollic Andosols*, are friable and smeary with humic topsoils (Jaetzold et al. 2010). The vegetation comprises indigenous trees, such as red stinkwood (*Prunus Africana*), bamboo (*Arundinaria alpina*), red cedar (*Juniperus procera*), African wild olive (*Olea europaea* ssp. *Africana*), East African olive (*Olea capensis* ssp. *hochstetteri*), broad-leaved yellowwood (*Podocarpus latifolius*), brittlewood (*Nuxia congesta*), clematis (*Clematis hirsuta*), schefflera (*Schefflera volkensii*), and forest dombeya (*Dombeya torrida*), exotic trees, such as pine (*Pinus patula*) and cypress (*Cupressus lusitanica*), as well as grasses like kikuyu grass (*Pennisetum clandestinum*). The major crops grown are maize (*Zea mays*), beans (*Phaseolus vulgaris*), wheat (*Triticum aestivum*), and potatoes (*Solanum tuberosum*).

## 2.2 Data sources, processing and analyses

Figure 3 summarizes the data and methods applied in this study, while the following sub-sections give the details:

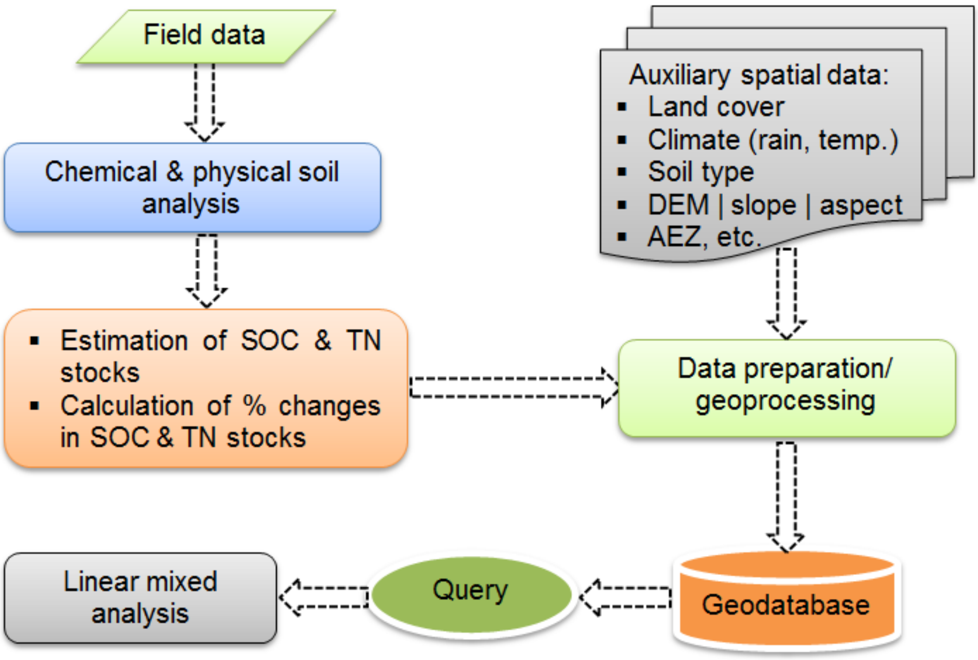



**Fig. 3:** Schematic representation of the data and methods used in the study

### 2.2.1 Field and laboratory methods

#### 2.2.1.1 Sampling design and soil sampling

Fieldwork was conducted between June and August 2012. The sites were selected to minimize the variations in climate, soil type, and slope. Four to thirteen sampling plots

(30×30m) were laid out in a completely randomized design within the natural forests (NF), plantation forests (PF), bamboo forests (BF), and croplands that had been established on natural forests (NF2C), plantation forests (PF2C) and bamboo forests (BF2C) (Fig. 4). In each plot, an auger was used to collect soil samples from the centre and four corners of the plot at two depths, one at 0-15cm and the other at 15-30cm. The samples taken from corresponding depths were thoroughly mixed and bulked into one composite sample of about 500g. To determine BD, a core sampler (5 cm in diameter and 5cm in height) was used to collect one undisturbed sample per depth from each plot centre. Geographical position, elevation, vegetation, and land management practices at each plot were also recorded. A total of 120 soil samples were collected and transported to the National Agricultural Research Laboratories for chemical and physical analyses.


**Fig. 4**: Illustrations of the land cover types: (**a**) mixed natural forest; (**b**) pine and cypress plantation forest; (**c**) bamboo forest; and (**d**) croplands (Source: Author).

#### 2.2.1.2 Physical and chemical soil analysis

At the laboratory, the soil samples were air-dried, ground and passed through a 2 mm mesh. SOC concentration was determined using Walkley-Black wet oxidation method (Nelson and Sommers 1982), while TN concentration was determined using Kjeldahl digestion method (Bremner and Mulvaney 1982). BD was estimated using the core method after oven-drying a specific volume of soil at 105 °C for 48 hours (Blake 1965). Particle size distribution was analysed using the hydrometer method after dispersing soil and eliminating organic matter

(Day 1965). Potassium (K) was measured using a flame-photometer, calcium (Ca) and magnesium (Mg) using an atomic absorption spectrophotometer, and pH (1:2.5 soil-water) using a pH meter (Okalebo et al. 2002). Phosphorous (P) was analysed using the Mehlich method (Okalebo et al. 2002). The data on soil properties are found in Table 1.

#### 2.2.1.3 Estimation of soil organic carbon and total nitrogen stocks

Stocks of SOC (Mg C $ha^{-1}$) for each depth were calculated using Eq. 1:

$$\mathrm{SOC_{st}} = \frac{\mathrm{SOC}}{100} \times \mathrm{BD} \times \mathrm{D} \times 100 \qquad (1)$$

where: $SOC_{st}$ is the soil organic carbon stock (Mg C $ha^{-1}$); SOC is the soil organic carbon concentration (%), which is then converted to g C $g^{-1}$ soil; BD is the bulk density (g $cm^{-3}$); D is the depth (cm); 100 is the multiplication factor to convert the SOC per unit area from g C $cm^{-2}$ to Mg C $ha^{-1}$. Coarse particles were negligible due to the softness of the volcanic rocks; hence, Eq. 1 does not account for them. Similarly, TN mass per unit area ($TN_{st}$; Mg N $ha^{-1}$) for each depth was computed by substituting TN for SOC in Eq. 1. The percentage changes (Δ) in $SOC_{st}$ (or $TN_{st}$) following NF, PF or BF conversions were then estimated using Eq. 2:

$$\Delta\ (\%) = \frac{\mathrm{SOC}_{st}\ (\text{or } \mathrm{TN}_{st})\ \text{under NF2C (PF2C or BF2C)} - \mathrm{SOC}_{st}\ (\text{or } \mathrm{TN}_{st})\ \text{under NF (PF or BF)}}{\mathrm{SOC}_{st}\ (\text{or } \mathrm{TN}_{st})\ \text{under NF (PF or BF)}} \times 100 \qquad (2)$$

SOC and TN stocks in the surface (0-15cm) and subsurface soils (15-30cm) were summed up to obtain the total stocks in the soil from the surface to a depth 30cm.

### 2.2.2 Remote sensing and GIS methods

Land cover maps (Fig. 2) produced through classification of terrain-corrected Landsat 5 TM images acquired in 1985 and 2011 were taken from Were et al. (2013). The images were atmospherically corrected, geometrically co-registered, and subsets made for classification using partitioning, hybrid classification, and spatial reclassification techniques. The subset for each date was further subdivided into spectrally distinct segments for separate classification. Depending on the degree of spatial heterogeneity, 15-30 spectral clusters were defined for each segment using an unsupervised classification procedure involving the ISODATA algorithm. The resultant clusters were assigned land cover labels (grassland, forest, or cropland) based on the analyst's knowledge of the area, as well as the ancillary and field data. In case of errors, supplementary spectral signatures were extracted, merged with the signatures of ISODATA clusters, and classified using the maximum likelihood algorithm. The land cover classes in each classified segment were recoded and, subsequently, the segments for each date were mosaicked. Finally, a majority filter was applied to reduce noise on the resultant seamless land cover maps. The quality of these maps was assessed using ancillary, temporally-invariant, and ground data. The overall accuracy of the 1985 and 2011 land cover maps were 95% and 89%, respectively.

Existing databases provided the auxiliary spatial data used to describe the topographical, climatic, agro-ecological, and pedological attributes of the area. Climate data (mean annual temperature and rainfall) were obtained from www.worldclim.org, soil data (soil type) from the Kenya Soil Survey, data on agro-ecological zonation from www.ilri.org/gis, and the digital elevation model (DEM) from http://srtm.csi.cgiar.org. Slope and aspect were extracted from the DEM. All these data were transformed to the Universal Transverse Mercator coordinate system (UTM WGS84 Zone 36S). The area of interest was clipped from each thematic layer, and all layers in vector format were rasterized. The datasets were then resampled to 100m and a geodatabase was built. The field and laboratory data were also integrated into the geodatabase as points using the geographical coordinates that were recorded at each sampling plot. The attribute values from each raster dataset (e.g., slope, rainfall, soil type) were extracted to these points. This facilitated querying of the geodatabase to only select those points that met the criteria for statistical analyses. All geoprocessing and analyses were performed using ArcGIS® 10.1 and ERDAS IMAGINE® 2011.

### 2.2.3 Statistical methods

Soil attributes of the point data in the geodatabase were summarized by land cover types and soil depths. Descriptive and correlation statistics were used to explore the distributions and relationships among various soil characteristics. Subsequently, linear mixed models were fitted to test the effects of land cover, soil depth, and sampling plot on SOC, $SOC_{st}$, TN, $TN_{st}$, and BD for each category of forest to cropland conversion: NF vs. NF2C; PF vs. PF2C; and, BF vs. BF2C. Equation 3 shows the form of the statistical model (Montgomery 2006):

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + \epsilon_{ijkl} \qquad (3)$$

where: $\mu$ is the overall mean, $\tau_i$ is the fixed effect of the *i*th land cover treatment, $\beta_j$ is the fixed effect of the *j*th soil depth treatment, $\gamma_k$ is the random effect of sampling plot, and $\epsilon_{ijkl}$ is the normally and independently distributed random error with zero mean and constant variance. Pairwise comparisons between the types of land cover and soil depths were based on *post hoc* t-tests at a 5% significance level. The *p*-values were adjusted by single-step method. Homoscedasticity was checked using residual plots and normality using normal probability plots. All analyses were carried out using package "nlme" and "lme4" in R version 3.0.1 (R Core Team 2013) and Microsoft Excel® 2010.

## 3.0 Results

### 3.1 Basic soil properties under different land cover types and soil depths

**Table 2:** Pearson's correlation coefficients of the soil properties

| Soil properties | Soil depth (0 - 15 cm) | | | | | | Soil depth (15 – 30 cm) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | TN | SOC | BD | Clay | Silt | Sand | TN | SOC | BD | Clay | Silt | Sand |
| TN | 1.00 | | | | | | 1.00 | | | | | |
| SOC | 0.99 | 1.00 | | | | | 0.99 | 1.00 | | | | |
| BD | -0.19 | -0.19 | 1.00 | | | | -0.06 | -0.05 | 1.00 | | | |
| Clay | 0.66 | 0.66 | -0.14 | 1.00 | | | 0.23 | 0.21 | -0.31 | 1.00 | | |
| Silt | -0.69 | -0.69 | 0.26 | -0.61 | 1.00 | | -0.50 | -0.49 | 0.32 | -0.53 | 1.00 | |
| Sand | -0.09 | -0.10 | -0.10 | -0.58 | -0.29 | 1.00 | 0.35 | 0.35 | -0.08 | -0.32 | -0.64 | 1.00 |

The means and standard deviations of select physical and chemical soil properties of different land cover types are presented in Table 1. In the surface soils (0-15cm), the highest BD was in PF and the lowest in NF2C, while in the subsurface soils (15-30cm), the highest BD was in PF and the lowest in BF. At all sites, BD never exceeded 1.0 g $cm^{-3}$ and was higher in the subsurface soils. Conversely, SOC and TN concentrations ranged from moderate to high and diminished as soil depth increased. In the surface soils, the highest SOC and TN concentrations were in NF (6.1% and 0.6%, respectively) and the lowest in NF2C (3.1% and 0.3%), while in the subsurface soils, the highest SOC and TN concentrations were in NF (4.2% and 0.4%) and the lowest in PF2C (2.3% and 0.2%). The proportions of soil separates ranged from 32 to 40% for sand, 31 to 45% for silt, and 21 to 32% for clay. Sand content was generally higher in the surface soils, while silt content was higher in the subsurface soils. The soils were moderately acidic with the pH levels varying between 5.0 and 6.1. NF2C had the highest pH values, while PF had the lowest pH values in both soil depths. The lowest available phosphorus and potassium were both in the surface and subsurface soils of PF, while the highest was in the surface soils of BF2C. Correlation patterns in the matrix show that SOC concentration was positively correlated with TN concentration and clay content, but negatively with BD, silt, and sand content, both in the surface and subsurface soils (Table 2). The correlations of TN concentration with other soil properties showed similar trends to SOC concentration.

### 3.2 Estimated SOC and TN stocks under different land cover types and soil depths

According to Table 3, in the surface soils (0-15cm), the highest SOC and TN stocks were in NF (71.6 and 7.1 Mg $ha^{-1}$, respectively) and the lowest in NF2C (35.4 and 3.5 Mg $ha^{-1}$). In the subsurface soils (15-30cm), the highest stocks were still in NF (55.7 and 5.6 Mg $ha^{-1}$), but the lowest were in PF2C (32.3 and 3.2 Mg $ha^{-1}$). Both SOC and TN stocks decreased as depth increased at all sites as shown in Figs. 5 and 6. The highest proportions of change in the stocks followed conversions from NF, while the lowest followed conversions from BF. In particular, cultivation of NF reduced both SOC and TN stocks by about 51% in the surface soils, and about 42% in the subsurface soils (Table 3; Figs. 7 and 8). Further, cultivation of PF reduced SOC and TN stocks by about 28% each in the surface soils, and about 36% each in the subsurface soils. However, cultivation of BF presented mixed results. In the surface soils, SOC stocks increased by 1%, while TN stocks decreased by 0.6%. In the subsurface soils, SOC and TN stocks were about 13% lower than at the surface. The same patterns were observed even when the entire topsoil (0-30cm) was considered. The highest stocks of SOC amounting to 127 Mg $ha^{-1}$ were in NF, which cropland conversion reduced by 46.8%. The lowest stocks of 101.5 Mg $ha^{-1}$ were in BF, which cropland conversion reduced by 4.4%. Similarly, the highest stocks of TN were in NF (12.7 Mg $ha^{-1}$), which cropland conversion reduced by 47%. The lowest stocks were in BF (10.3 Mg $ha^{-1}$), which reduced by 5.4% when converted to croplands.

### 3.3 Effects of land cover and soil depth on SOC, $SOC_{st}$, TN, $TN_{st}$ and BD

Table 4a shows a statistically significant land-cover effect on all the soil properties ($p$<0.0001) except for BD ($p$=0.6648) in the NF vs. NF2C category. SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), and $TN_{st}$ ($p$<0.001) in NF differed from NF2C. However, BD were equal ($p$=0.9450). There was also a highly significant soil-depth effect on SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0002), TN ($p$<0.0001), $TN_{st}$ ($p$=0.0002), and BD ($p$<0.0001). SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), $TN_{st}$ ($p$<0.001), and BD ($p$<0.0001) between soil depth 0-15cm and 15-30cm were significantly different.

**Table 1:** Some soil properties of the different land cover types (Means and standard deviations)

| | | | Soil properties | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Land cover | *n* | Depth | SOC (%) | TN (%) | P (ppm) | K (me %) | pH | BD (g cm$^{-3}$) | Sand (%) | Silt (%) | Clay (%) |
| NF | 15 | 0 - 15 | 6.1±1.3 | 0.6±0.1 | 26.3±6.5 | 0.9±0.30 | 5.6±0.3 | 0.8±0.2 | 37.1±6.5 | 31.5±4.4 | 31.5±7.2 |
| | 15 | 15 - 30 | 4.2±0.8 | 0.4±0.1 | 22.9±8.0 | 0.8±0.47 | 5.5±0.3 | 0.9±0.1 | 33.3±5.1 | 39.6±3.6 | 27.1±3.0 |
| NF2C | 15 | 0 - 15 | 3.1±0.8 | 0.3±0.1 | 29.9±8.6 | 1.2±0.31 | 6.1±0.5 | 0.8±0.1 | 38.1±6.7 | 39.5±5.3 | 22.4±4.7 |
| | 15 | 15 - 30 | 2.4±0.5 | 0.2±0.1 | 26.1±7.7 | 1.0±0.26 | 6.0±0.5 | 0.9±0.1 | 31.7±8.3 | 44.8±6.3 | 23.5±4.3 |
| PF | 13 | 0 - 15 | 3.9±0.5 | 0.4±0.1 | 16.4±4.7 | 0.5±0.17 | 5.3±0.3 | 0.9±0.1 | 32.9±4.1 | 38.6±5.0 | 28.5±4.6 |
| | 13 | 15 - 30 | 3.5±0.5 | 0.3±0.1 | 14.9±2.8 | 0.4±0.11 | 5.0±0.2 | 1.0±0.1 | 34.0±3.6 | 41.4±5.3 | 24.6±3.8 |
| PF2C | 13 | 0 - 15 | 3.2±0.8 | 0.3±0.1 | 27.6±8.0 | 1.0±0.37 | 5.9±0.2 | 0.8±0.1 | 39.7±4.8 | 38.9±5.6 | 21.4±5.4 |
| | 13 | 15 - 30 | 2.3±0.4 | 0.2±0.1 | 25.8±8.9 | 0.9±0.29 | 6.0±0.3 | 0.9±0.1 | 34.2±5.3 | 43.9±5.4 | 22.0±5.7 |
| BF | 4 | 0 - 15 | 5.2±2.3 | 0.5±0.2 | 28.3±7.0 | 1.3±0.26 | 5.6±0.2 | 0.8±0.1 | 35.5±3.4 | 30.0±5.2 | 34.5±6.0 |
| | 4 | 15 - 30 | 3.3±0.8 | 0.3±0.1 | 25.0±10.8 | 1.3±0.36 | 5.6±0.6 | 0.8±0.1 | 36.0±5.4 | 28.5±5.0 | 35.5±4.4 |
| BF2C | 4 | 0 - 15 | 5.2±2.2 | 0.5±0.2 | 32.3±10.6 | 1.5±0.44 | 6.0±0.5 | 0.8±0.1 | 38.5±8.5 | 31.5±5.0 | 30.0±9.8 |
| | 4 | 15 - 30 | 2.7±0.7 | 0.3±0.1 | 26.3±11.2 | 1.1±0.29 | 5.6±0.6 | 0.8±0.1 | 34.0±5.2 | 38.5±4.1 | 27.5±1.9 |

Note: NF = Natural forest; NF2C = Natural forest converted to cropland; PF= Plantation forest; PF2C = Plantation forest converted to cropland; BF = Bamboo forest; and, BF2C = Bamboo forest converted to cropland

**Table 3:** Soil organic carbon and total nitrogen stocks under different land cover types and soil depths (Means and standard deviations)

| | Soil depth (0 - 15 cm) | | | | Soil depth (15 - 30 cm) | | | | Soil depth (0 - 30 cm) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Land cover | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change | $SOC_{st}$ (Mg ha$^{-1}$) | % change | $TN_{st}$ (Mg ha$^{-1}$) | % change |
| NF | 71.6±17.9 | | 7.1±1.8 | | 55.7±11.3 | | 5.58±1.2 | | 127.3±26.5 | | 12.7±2.6 | |
| NF2C | 35.4±10.8 | -50.6 | 3.5±1.1 | -50.4 | 32.5±7.4 | -41.8 | 3.20±0.7 | -42.7 | 67.8±15.5 | -46.8 | 6.7±1.6 | -47.0 |
| PF | 54.6±3.9 | | 5.4±0.4 | | 49.9±7.0 | | 4.94±0.7 | | 104.3±8.9 | | 10.3±1.0 | |
| PF2C | 39.3±11.2 | -28.0 | 3.9±1.1 | -27.6 | 32.3±6.5 | -35.3 | 3.16±0.6 | -36.0 | 71.5±14.2 | -31.48 | 7.1±1.5 | -31.6 |
| BF | 61.6±27.3 | | 6.3±2.8 | | 39.9±9.0 | | 4.05±0.9 | | 101.5±32.3 | | 10.3±3.3 | |
| BF2C | 62.2±20.1 | 1.0 | 6.2±2.0 | -0.6 | 34.8±10.2 | -12.7 | 3.52±1.0 | -13.1 | 97.0±19.8 | -4.48 | 9.7±1.9 | -5.4 |

Note: NF = Natural forest; NF2C = Natural forest converted to cropland; PF= Plantation forest; PF2C = Plantation forest converted to cropland; BF = Bamboo forest; and, BF2C = Bamboo forest converted to cropland. See also the errata

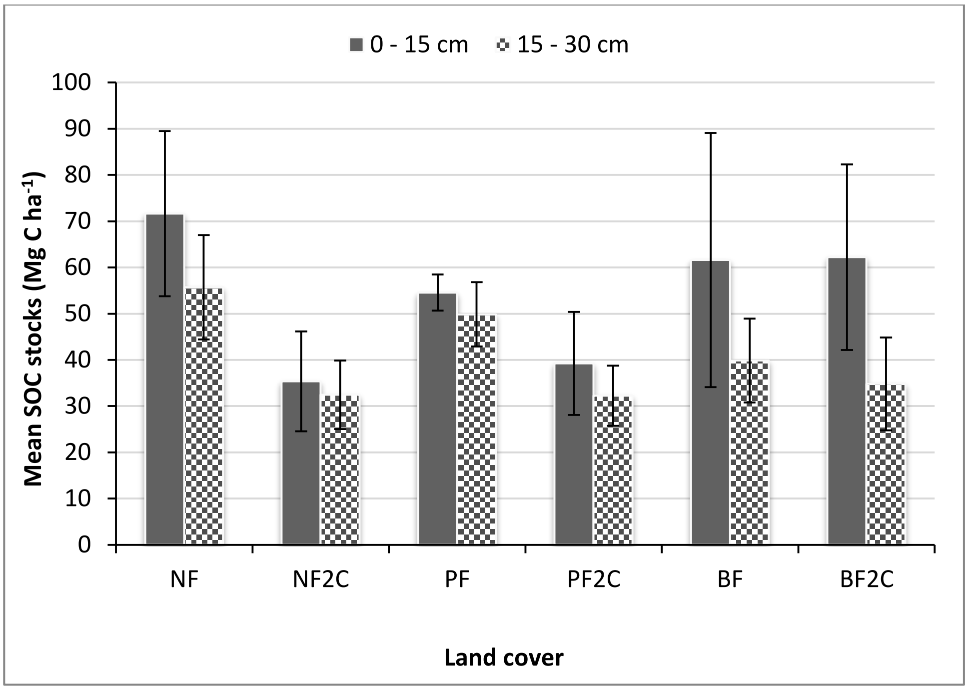



**Fig. 5:** Soil organic carbon stocks under different land cover types and soil depths. NF = natural forest; NF2C = natural forest converted to cropland; PF= plantation forest; PF2C = plantation forest converted to cropland; BF = bamboo forest; and BF2C = bamboo forest converted to cropland

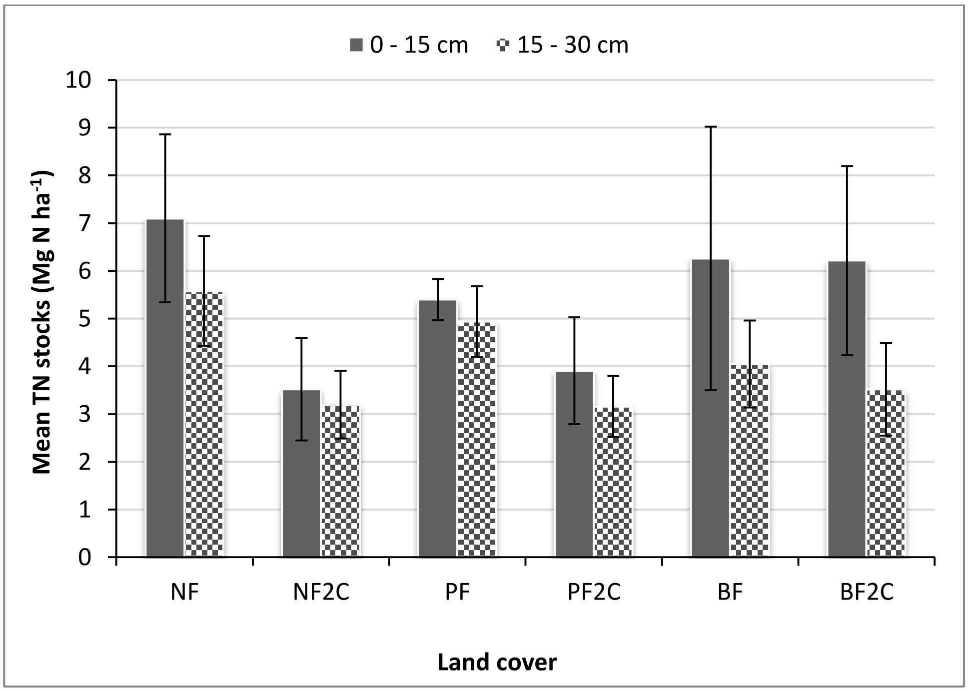



**Fig. 6:** Total nitrogen stocks under different land cover types and soil depths. NF = natural forest; NF2C = natural forest converted to cropland; PF= plantation forest; PF2C = plantation forest converted to cropland; BF = bamboo forest; and BF2C = bamboo forest converted to cropland

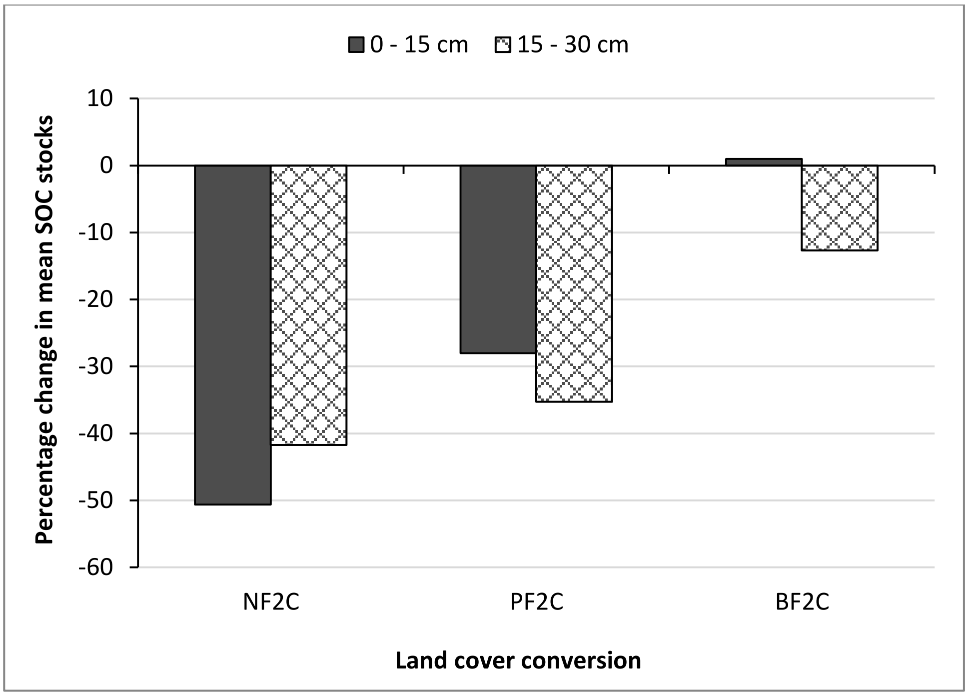



**Fig. 7:** Percentage change in SOC stocks following forest conversions. NF2C = natural forest converted to cropland; PF2C = plantation forest converted to cropland; and BF2C = bamboo forest converted to cropland

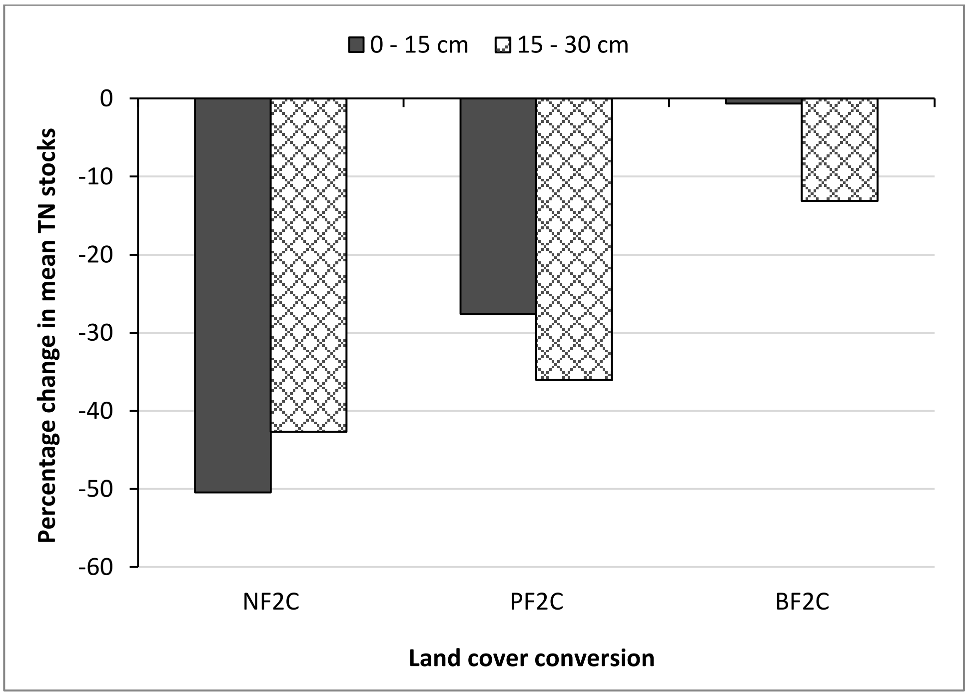



**Fig. 8**: Percentage change in TN stocks following forest conversion. NF2C = natural forest converted to cropland; PF2C = plantation forest converted to cropland; and BF2C = bamboo forest converted to cropland

Similarly, land cover had a highly significant effect on SOC ($p$=0.0001), $SOC_{st}$ ($p$=0.0001), TN ($p$=0.0002), $TN_{st}$ ($p$<0.0001), and BD ($p$=0.0191) in the PF vs. PF2C category (Table 4b). There were significant differences in SOC ($p$<0.0001), $SOC_{st}$ ($p$<0.001), TN ($p$<0.0001), $TN_{st}$ ($p$<0.001), and BD ($p$=0.0320) between PF and PF2C. Soil depth also had a highly significant effect on SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0038), TN ($p$<0.0001), $TN_{st}$ ($p$=0.0026), and BD ($p$=0.0153). SOC ($p$<0.0001), $SOC_{st}$ ($p$=0.0041), TN ($p$<0.0001), $TN_{st}$ ($p$=0.0026), and BD ($p$=0.0255) in soil depth 0-15cm differed from soil depth 15-30cm.

In contrast, land cover had no significant effect on SOC ($p$=0.7894), $SOC_{st}$ ($p$=0.8217), TN ($p$=0.7460), $TN_{st}$ ($p$=0.7749), and BD ($p$=0.6992) in the BF vs. BF2C category (Table 4c). SOC ($p$=0.9825), $SOC_{st}$ ($p$=0.990), TN ($p$=0.9699), $TN_{st}$ ($p$=0.9794), and BD ($p$=0.9540) in BF were similar to BF2C. However, soil depth had a significant effect on all the soil properties (SOC ($p$=0.0284), $SOC_{st}$ ($p$=0.0206), TN ($p$=0.0279), and $TN_{st}$ ($p$=0.0204)) except for BD ($p$=0.4545). Pairwise comparisons revealed differences in SOC ($p$=0.0157), $SOC_{st}$ ($p$=0.008), TN ($p$=0.0158), and $TN_{st}$ ($p$=0.008), but similarities in BD ($p$=0.7610), between soil depths 0-15cm and 15-30cm.

**Table 4:** Statistical summary of land-cover and soil-depth effects on SOC, $SOC_{st}$, TN, $TN_{st}$, and BD

| | SOC | | $SOC_{st}$ | | TN | | $TN_{st}$ | | BD | |
|---|---|---|---|---|---|---|---|---|---|---|
| **(a) NF vs. NF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 63.22 | <0.0001 | 56.46 | <0.0001 | 66.09 | <0.0001 | 57.44 | <0.0001 | 0.19 | 0.6648 |
| Soil depth | 72.51 | <0.0001 | 18.02 | 0.0002 | 76.39 | <0.0001 | 17.65 | 0.0002 | 32.01 | <0.0001 |
| Land cover × Soil depth | 16.47 | 0.0004 | 8.61 | 0.0066 | 16.17 | 0.0004 | 7.42 | 0.0110 | - | - |
| **(b) PF vs. PF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 22.17 | 0.0001 | 49.89 | <0.0001 | 20.04 | 0.0002 | 44.92 | <0.0001 | 6.32 | 0.0191 |
| Soil depth | 37.75 | <0.0001 | 10.18 | 0.0038 | 40.00 | <0.0001 | 11.15 | 0.0026 | 6.78 | 0.0153 |
| **(c) BF vs. BF2C** | | | | | | | | | | |
| Source of variation | F | P | F | P | F | P | F | P | F | P |
| Land cover | 0.08 | 0.7894 | 0.06 | 0.8217 | 0.12 | 0.7460 | 0.09 | 0.7749 | 0.16 | 0.6992 |
| Soil depth | 7.57 | 0.0284 | 8.87 | 0.0206 | 7.65 | 0.0279 | 8.90 | 0.0204 | 0.62 | 0.4545 |

Note: The degree of freedom was 1 in all cases; F = *F*-value; P = *p*-value; NF = Natural forest; NF2C = natural forest converted to cropland; PF= plantation forest; PF2C = plantation forest converted to cropland; BF = bamboo forest; and BF2C = bamboo forest converted to cropland

## 4.0 Discussion

The empirical results suggest that forest-to-cropland conversions have reduced SOC and TN concentrations and stocks in the Eastern Mau Forest Reserve. The mean SOC and TN concentrations and stocks between forests and cropland establishments differed significantly as hypothesized (Table 4), and sometimes the difference was by as much as half (Table 3). This is consistent with the findings of previous studies in the tropics (Enanga et al. 2011; Walker and Desanker 2004; Lemenih et al. 2005; Bewketa and Stroosnidjer 2003; Solomon et al. 2000; Detwiler 1986). In these studies, the rates and magnitude of decrease in SOC and TN stocks varied with soil type and time since conversion to cropland. The declining trend mostly persists until new steady states of carbon and nitrogen are reached after years of continuous cultivation (Eaton et al. 2008; Lemenih et al. 2005; Evrendilek et al. 2004).

The decrease in SOC and TN stocks *per se* can be explained by the subsequent disruption of the balance between inputs and outputs of carbon and nitrogen in the soil system after forest conversion. Forest ecosystems usually have a higher net primary productivity (NPP) than agro-ecosystems; thus, their inputs of detritus to the soils are also higher (Eclesia et al. 2012; Smith 2008). In Eastern Mau, the NPP of forests and inputs of carbon and nitrogen to their soils is even higher because of the extremely fertile *Andosols* and high rainfall amounts. Despite the lower NPP of agro-ecosystems, the bulk of their biomass is usually removed from the crop fields after harvest for use as food or fuel. Only a small amount of readily decomposable residues remain on the fields to accumulate SOM. Removal of crop biomass after harvest also aggravates the erosion processes, which were initiated by forest clearance, in the predominant uplands leading to SOC and TN losses. Additionally, frequent tillage and other perturbations disintegrate soil aggregates, redistribute crop residues, and alter soil aeration, moisture, and temperature. This accelerates microbial decomposition and oxidation of the soil's organic matter to $CO_2$, which is ultimately emitted to the atmosphere (Wiesmeier et al. 2012; Batlle-Aguilar et al., 2011; Lal 2004; Powers 2004; Murty et al. 2002; Follett 2001). The reduction of SOC and TN stocks after forest conversions occurred regardless of the application of inorganic fertilizers in most croplands. This implies that supplementing fertilization by agro-forestry techniques, such as planting fast-growing, highly productive, deep-rooted, and nitrogen-fixing tree species within NF2C and PF2C, may be the optimal option to restore and enhance SOC and TN stocks.

In contrast, the mean SOC and TN concentrations and stocks between BF and BF2C were similar (Table 4c). This can be attributed to the establishment of croplands within BF less than ten years ago. The sample sizes for BF and BF2C were also small ($n$=4); hence, the data may not have fully represented the variations within these land cover treatments. Future studies should increase the sample sizes in these two land cover groups to reduce the relatively large standard deviations from the mean SOC and TN stocks.

Further, the results revealed that BD in the croplands had not significantly changed (Table 4) except for PF2C. This was unexpected because BD tends to increase as tillage breaks down soil aggregates. But as Walker and Desanker (2004) argued, the tillage of most croplands by hand may have only caused minimal disturbances to substantially increase the

BD. This may have obviated the confounding influence of BD changes on estimating the changes in SOC and TN stocks after deforestation, as well as on analysing land-cover effect on these stocks. There was no confounding either in the PF vs. PF2C category because comparable results were obtained even when BD was included as a covariate in the statistical analyses.

Finally, SOC and TN concentrations and stocks decreased significantly as soil depth increased under all land cover treatments, which is in accordance with previous studies (Demessie et al. 2013; Li et al. 2013; Zhang et al. 2013; Fang et al. 2012; Girmay and Singh 2012; Zhang et al. 2012; Han et al. 2010; Wang et al. 2010; Chen et al. 2009; Birch-Thomsen et al. 2007; Yimer et al. 2007; Brown and Lugo 1990). The cause of this decline is that organic material inputs to forest soils (litter fall, exudates, leachates, dead roots, etc.) and agricultural soils (crop residues, manures, fertilizers, etc.) mostly reside in the upper layers, with only small amounts penetrating much deeper. High precipitation in the area may also instigate leaching of dissolved organic carbon and nitrogen compounds from the subsurface soils (15-30cm) to deeper soils that were not sampled. The lower concentrations and stocks of SOC in the subsurface soils also help to explain the corresponding higher BD values. Other factors that may account for higher BD in the subsurface soils include reduced aggregation, root penetration, and soil micro-organism populations, as well as the compacting weight of surface soils (USDA 2008).

## 5.0 Conclusions and recommendations

This study assessed the effects of forest to cropland conversions on SOC, TN, and BD in the Eastern Mau Forest Reserve. Based on the results, we conclude that (i) conversion of forests, particularly NF and PF, to croplands has led to a significant decline in the concentrations and stocks of SOC and TN, but no significant BD changes, and (ii) the surface soils contain significantly more concentrations and stocks of SOC and TN, while BD is significantly higher in the subsurface soils. This indicates that (i) transformation from natural to human-dominated landscape is increases the risk of soil degradation and restricts the ecosystem’s capacity to store carbon and nitrogen, and (ii) the subsurface soils have potential for carbon and nitrogen storage. Thus, intervention measures to enhance carbon and nitrogen storage should focus not only on surface soils, but also on subsurface soils. BMPs may reduce carbon and nitrogen losses in the croplands, especially agro-forestry practices that introduce fast-growing, highly productive, deep-rooted, and nitrogen-fixing trees. Long-term carbon and nitrogen storage in the forest soils depends on proper management and protection of the forests from further deforestation and degradation. Appropriate land use and land use change policies are needed to protect the soils.

The findings of this study improve our knowledge of the impacts of human activities on soil properties in the area. They also provide a basis to design sustainable land management and carbon sequestration strategies. In view of the ongoing soil degradation and requirements for sequestration of atmospheric $CO_2$, future research should also employ remote sensing and GIS approaches to model and map the spatial patterns of carbon and nitrogen stocks. These approaches can afford holistic information and deeper understanding of carbon and nitrogen storage and fluxes throughout East Africa.

**Acknowledgements**

This study was funded by the Research Council of Norway through the Norwegian University of Life Sciences. Special thanks go to Mr. P. Owenga for technical support during fieldwork, Mr. E. Thairu for excellent driving under difficult field conditions, and the two reviewers for improving the manuscript.

## References


Awiti AO, Walsh MG, Kinyamario J (2008) Dynamics of topsoil carbon and nitrogen along a tropical forest–cropland chronosequence: Evidence from stable isotope analysis and spectroscopy. Agriculture, Ecosystems and Environment 127: 265–272

Baldyga T, Miller SN, Driese KL, Gichaba CM (2007) Assessing land cover change in Kenya's Mau forest region using remotely-sensed data. African Journal of Ecology 46: 46-54

Bationo A, Kihara J, Vanlauwe B, Waswa B, Kimetu J (2007) Soil organic carbon dynamics, functions and management in West African agro-ecosystems. Agricultural Systems 94: 13-25

Batlle-Aguilar J, Brovelli A, Porporato A, Barry DA (2011) Modelling soil carbon and nitrogen cycles during land use change- A review. Agron Sustain Dev 31: 251-274

Bewketa W, Stroosnijder L (2003) Effects of agro-ecological land use succession on soil properties in Chemoga watershed, Blue Nile basin, Ethiopia. Geoderma 111: 85–98

Birch-Thomsen T, Elberling B, Fog B, Magid J (2007) Temporal and spatial trends in soil organic carbon stocks following maize cultivation in semi-arid Tanzania, East Africa. Nutr Cycl Agroecosyst 79: 291–302

Biro K, Pradhan B, Buchroithner M, Makeschin F (2011) Land use/ land cover change analysis and its impact on soil properties in the northern part of Gadarif region, Sudan. Land Degradation and Development 24: 90–102

Blake GR (1965) Bulk density. In: Black CA (ed) Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling. American society of agronomy, Inc., Madison, Wisconsin, USA

Bremner JM, Mulvaney CS (1982) Nitrogen - total. In: Page AL (ed) Methods of soil analysis, Part 2. Chemical and microbiological properties, 2nd edn. American society of agronomy, Inc., Madison, Wisconsin, USA

Brown S, Lugo AE (1990) Effects of forest clearing and succession on the carbon and nitrogen content of soils in Puerto Rico and US Virgin Islands. Plant and Soil 124: 53-64

Chen H, Marhan S, Billen N, Stahr K (2009) Soil organic carbon and total nitrogen stocks as affected by different land uses in Baden-Württemberg (southwest Germany). J Plant Nutr Soil Sci 172: 32–42

Day PR (1965) Particle fractionation and particle size analysis. In: Black CA (ed) Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling. American society of agronomy, Inc., Madison, Wisconsin, USA

Demessie A, Singh BR, Lal R (2013) Soil carbon and nitrogen stocks under chronosequence of farm and traditional agro-forestry uses in Gambo district, southern Ethiopia. Nutr Cycl Agroecosys 95: 365-375

Detwiler RP (1986) Land use change and the global carbon cycle: The role of tropical soils. Biogeochemistry 2: 67-93

Don A, Schumacher J, Freibauer A (2011) Impact of tropical land use change on soil organic carbon stocks - A meta-analysis, Global Change Biology 17: 1658–1670

Eaton JM, McGoff NM, Byrne KA, Leahy P, Kiely G (2008) Land cover change and soil organic carbon stocks in the Republic of Ireland 1851–2000. Climatic Change 91:317–334

Eclesia RP, Jobbagy EG, Jackson RB, Biganzoli F, Piñeiro G (2012) Shifts in soil organic carbon for plantation and pasture establishment in native forests and grasslands of South America. Global Change Biology 18: 3237–3251

Enanga EM, Shivoga WA, Maina-Gichaba C, Creed IF (2011) Observing changes in riparian buffer strip soil properties related to land use activities in the River Njoro watershed, Kenya. Water Air Soil Pollut 218:587–601

Evrendilek F, Celik I, Kilic S (2004) Changes in soil organic carbon and other physical soil properties along adjacent Mediterranean forest, grassland, and cropland ecosystems in Turkey. Journal of Arid Environments 59: 743–752

Fang X, Xue Z, Li B, An S (2012) Soil organic carbon distribution in relation to land use and its storage in a small watershed of the Loess Plateau, China. Catena, 88, 6–13

Follett RF (2001) Soil management concepts and carbon sequestration in cropland soils. Soil and Tillage Research 61: 77-92

Girmay G, Singh BR (2012) Changes in soil organic carbon stocks and soil quality: Land use system effects in northern Ethiopia. Acta Agriculturae Scandinavica, Section B &#x2013, Soil & Plant Science 62 (6): 519-530

Government of Kenya (2009) Report of the prime minister's task force on the conservation of the Mau forest complex. [Online]. Available: http://www.maurestoration.go.ke/. [Accessed 2012, December 20]

Han F, Hu H, Zheng J, Du F, Zhang X (2010) Estimating soil organic carbon storage and distribution in a catchment of Loess Plateau, China. Geoderma 154: 261–266

IPCC (2013) Summary for Policymakers. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Islam K, Weil RR (2000) Land use effects on soil quality in a tropical forest ecosystem of Bangladesh. Agriculture, Ecosystems and Environment 79: 9–16

Jaetzold R, Schmidt H, Hornetz B, Shisanya C (2010) Farm management handbook of Kenya, Vol. II. Natural conditions and farm management information, 2nd edn, Part B Central Kenya, Subpart B1a Southern Rift Valley Province. Ministry of Agriculture, Kenya and German Agency for Technical Cooperation (GTZ), Nairobi

Jafarian Z, Kavian A (2013) Effects of land use change on soil organic carbon and nitrogen. Communications in Soil Science and Plant Analysis 44 (1-4): 339-346

Jiao Y, Xu Z, Zhao J (2009) Effects of grassland conversion to cropland and forest on soil organic carbon and dissolved organic carbon in the farming-pastoral ecotone of Inner Mongolia. Acta Ecologica Sinica 29: 150–154

Jing-cheng Y, Jian-hui H, Qing-min P, Jian-wei T, Xing-guo H (2004) Long-term impacts of land use change on dynamics of tropical soil carbon and nitrogen pools. Journal of Environmental Sciences 16(2): 256-261

Lal R (2008) Carbon sequestration. Phil. Trans. R. Soc. B 363: 815–830

Lal R (2004) Soil carbon sequestration to mitigate climate change. Geoderma 123: 1-22

Lemenih M, Karltun E, Olsson M (2005) Assessing soil chemical and physical property responses to deforestation and subsequent cultivation in smallholders farming system in Ethiopia. Agriculture, Ecosystems and Environment 105: 373–386

Li M, Zhang X, Pang G, Han F (2013) The estimation of soil organic carbon distribution and storage in a small catchment area of the Loess Plateau. Catena 101: 11–16

McCall GJH (1967) Geology of the Nakuru-Thomson's falls-Lake Hannington area: degree sheet No. 35, S.W. Quarter and 43 N.W. Quarter, Report No. 78. Government Printer, Nairobi

Montgomery DC (2006) Design and analysis of experiments. John Wiley & sons, Inc., New Jersey

Muñoz-Rojas M, Jordán A, Zavala LM, de la Rosa D, Abd-Elmabod SK, Anaya-Romero M (2012) Impact of land use and land cover changes on organic carbon stocks in Mediterranean soils (1956–2007). Land Degradation and Development, doi: 10.1002/ldr.2194

Murty D, Kirschbaum, MF McMurtrie, RE McGilvray, H (2002) Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. Global Change Biology 8: 105-123

Nelson DW, Sommers LE (1982) Total carbon, organic carbon and organic matter. In: Page, AL (ed) Methods of soil analysis, Part 2, Chemical and microbiological properties, 2nd edn. American society of agronomy, Inc., Madison, Wisconsin, USA

Okalebo JR, Gathna KW, Woomer PL (2002) Laboratory methods for soil and plant analysis: A working manual, 2nd edition. Tropical Soil Biology and Fertility Programme, Nairobi

Osher LJ, Matson PA, Amundson R (2003) Effect of land use change on soil carbon in Hawaii. Biogeochemistry 65: 213–232

Post WM, Kwon KC (2000) Soil carbon sequestration and land use change: Processes and potentials. Global Change Biology 6: 317-327

Powers JS (2004) Changes in soil carbon and nitrogen after contrasting land-use transitions in north-eastern Costa Rica. Ecosystems 7: 134–146

Powlson DS, Gregory PJ, Whalley WR, Quinton JN, Hopkins DW, Whitmore AP, Hirsch PR, Goulding KWT (2011) Soil management in relation to sustainable agriculture and ecosystem services. Food Policy 36: S72–S87

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://wwwR-projectorg/

Smith P (2008) Land use change and soil organic carbon dynamics. Nutr Cycl Agroecosyst 81: 169–178

Smith P (2004) Soils as carbon sinks: the global context. Soil Use and Management 20: 212-218

Solomon D, Lehmann J, Zech W (2000) Land use effects on soil organic matter properties of chromic luvisols in semi-arid northern Tanzania: Carbon, nitrogen, lignin and carbohydrates. Agriculture, Ecosystems and Environment 78: 203–213

UNEP (2009) Kenya: Atlas of our changing environment. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP). [Online]. Available: http://www.unep.org/dewa/africa/kenyaatlas/. [Accessed 2013, August 28]

USDA (2008) Soil quality indicators: Bulk density. [Online]. Available: http://soils.usda.gov/sqi/assessment/assessment.html. [Accessed 2013, August 28]

Vågen TG, Lal R, Singh BR (2005) Soil carbon sequestration in sub-Saharan Africa: A review. Land Degradation and Development 16: 53-71

Walker SM, Desanker PV (2004) The impact of land use on soil carbon in Miombo Woodlands of Malawi. Forest Ecology and Management 203: 345–360

Wang Y, Fu B, Lü Y, Song C, Luan Y (2010) Local-scale spatial variability of soil organic carbon and its stock in the hilly area of the Loess Plateau, China. Quaternary Research 73: 70–76

Wang Z, Han X, Li L (2008) Effects of grassland conversion to croplands on soil organic carbon in the temperate Inner Mongolia. Journal of Environmental Management 86: 529–534

Were KO, Dick ØB, Singh BR (2013) Remotely sensing the spatial and temporal land cover changes in Eastern Mau Forest Reserve and Lake Nakuru drainage basin, Kenya. Applied Geography 41: 75-86

Wiesmeier M, Spörlein P, Geuß U, Hangen E, Haug S, Reischl A, Schilling B, von Lützow M, Kögel-Knabner I (2012) Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. Global Change Biology 18: 2233–2245

Yimer F, Ledin S, Abdelkadir A (2007) Changes in soil organic carbon and total nitrogen contents in three adjacent land use types in the Bale Mountains, south-eastern highlands of Ethiopia. Forest Ecology and Management 242: 337–342

Zhang C, Liu G, Xue S, Sun C (2013) Soil organic carbon and total nitrogen storage as affected by land use in a small watershed of the Loess Plateau, China. European Journal of Soil Biology 54: 16-24

Zhang L, Xie Z, Zhao R, Wang Y (2012) The impact of land use change on soil organic carbon and labile organic carbon stocks in the Longzhong region of Loess Plateau. Journal of Arid Land 4(3): 241−250

# Paper IV

**Were, K.O.,** Singh, B.R., Dick, Ø.B. (2014). Spatially-distributed modelling and mapping of soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya. Submitted to *Catena* (under review).

# Spatially-distributed modelling and mapping of soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya

K.O. Were[a,b], B.R. Singh[c], Ø.B. Dick[a]

**Abstract**

Detailed knowledge about the spatial distribution of soil organic carbon (SOC) and total nitrogen (TN) stocks is fundamental for sustainable land management and climate change mitigation. This study aimed at: (i) modelling and mapping the spatial patterns, and (ii) estimating SOC and TN stocks to 30cm depth in the Eastern Mau Forest Reserve using field sampling, remote sensing, geographical information systems (GIS), and statistical modelling approaches. This is a critical ecosystem that offers essential goods and services, but the sustainability is threatened by deforestation and degradation. Results revealed that elevation, silt content, TN concentration and Landsat 8 Operational Land Imager band 11 explained 72% of the variability in SOC stocks, while the same factors (excluding silt content) explained 71% of the variability in TN stocks. The results further showed that the observed patterns of SOC and TN stocks were controlled more by TN and SOC concentrations, respectively, than the other factors. Forests stored the highest amounts of SOC and TN (3.78 Tg C and 0.38 Tg N) followed by croplands (2.46 Tg C and 0.25 Tg N) and grasslands (0.57 Tg C and 0.06 Tg N). Overall, the Eastern Mau Forest Reserve stored approximately 6.81 Tg C and 0.69 Tg N. The highest estimates of SOC and TN stocks (*hotspots*) occurred on the western and north-western parts where forests dominated, while the lowest estimates (*coldspots*) occurred on the eastern side where croplands had been established. Therefore, the *hotspots* calls for policies that promote conservation, while the *coldspots* those that support accumulation of SOC and TN stocks.

**Keywords:** Soil organic carbon • total nitrogen • digital soil mapping • remote sensing • GIS • Eastern Mau • Kenya

## 1.0 Introduction

Soil organic carbon (SOC) and total nitrogen (TN) are key determinants of biogeochemical cycling, soil quality, and various physical, chemical, and biological soil properties (Obade and Lal, 2013; Wang et al., 2013). They vary spatially and temporally in response to different climatic, edaphic, biotic, topographical, and lithological factors. Such dynamics also affect their contribution to the atmospheric greenhouse gases, particularly $CO_2$ and $N_2O$. Therefore, mapping the spatial patterns of SOC and TN stocks over time is important

[a] Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway

[b] Kenya Agricultural Research Institute, Kenya Soil Survey, P.O. Box 14733-00800, Nairobi, Kenya

[c] Department of Environmental Sciences, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway

to understand climate change, as well as land degradation at different scales. Unfortunately, the traditional soil mapping techniques are expensive, time-consuming, and yield coarse qualitative information (Mora-Vallejo et al., 2008; Mehrjardi et al., 2014). As a result, there is increasing effort to develop and apply new techniques for spatial prediction of soil properties, especially SOC stocks. The spotlight is on SOC stocks because the world's soils contain about 1500 Pg C to 1m depth, which is twice the amount of C in the atmospheric, and three times in the biotic pool (Lal, 2004; Smith, 2004, 2008). Hence, even slight changes in SOC pool can significantly affect the global C cycle and climate.

The existing techniques for spatial prediction of target soil variables have been classified as: (i) measure and multiply (MM), and (ii) soil-landscape modelling (SLM) techniques (Mishra et al., 2010; Cambule et al., 2014). In MM approach, the area is stratified and the point estimates of the target soil property in a stratum are averaged and multiplied by the area of the stratum. Whereas in SLM approach, the variability of target soil property is explained by its relationships with soil-forming factors such as topography, climate, land use, vegetation, or soil type. Field observations of the target soil variables and environmental data are used to calibrate models, which are then applied to generate prediction surfaces of the target soil variable for a given area (Mishra et al., 2010; Li et al., 2013a; Cambule et al., 2014). The SLM approaches have been boosted by availability of inexpensive and spatially explicit environmental data from remote sensors and existing geodatabases. Although, MM approach is simple, it ignores the complex interactions of the environmental factors with the target soil variables, which lead to spatial variability. Thus, MM approach is more likely to yield predictions with lesser accuracy than SLM approach. McKenzie and Ryan (1999), McBratney et al. (2003), and Scull et al. (2003) have provided thorough reviews of spatial prediction of soil properties using environmental data, also known as digital soil mapping.

Literature abounds with examples of the SLM techniques that have been applied so far to model and map the spatial patterns of SOC and TN stocks. The techniques range from multiple linear regression (Lesch and Corwin, 2008; Meersmans et al., 2008) to partial least square regression (Selige et al., 2006; Amare et al., 2013), generalized linear models (Yang et al., 2008), classification and regression trees (Kheir et al., 2010; Martin et al., 2011; Razakamanarivo et al., 2011), kriging (Wu et al., 2009; Zhang et al., 2010; Liu et al., 2011; Li et al., 2013b; Cambule et al., 2014), regression-kriging (Hengl et al., 2004, 2007; Lamsal et al., 2006; Mora-Vallejo et al., 2008; Sumfleth and Duttmann, 2008; Li, 2010; Vasques et al., 2010a, 2010b; Dorji et al., 2014; Martin et al., 2014), geographically weighted regression (Mishra et al., 2010; Zhang et al., 2011; Mishra and Riley, 2012; Kumar et al., 2013; Wang et al., 2013), geographically weighted regression-kriging (Kumar et al., 2012), neural networks (Malone et al., 2009; Jaber and Al-Qinna, 2011; Li et al., 2013a), random forests (Grimm et al., 2008; Vågen and Winowiecki, 2013a, 2013b), rule-based models (Lacoste et al., 2014), and linear mixed models (Doetterl et al., 2013; Karunaratne et al., 2014). Of these methods, multiple linear regression (MLR) is the most popular because of its simplicity, computational efficiency, and straightforward interpretation (Li et al., 2013a). However, its assumptions of spatial stationarity in the effects of environmental variables and spatial independence in the target soil properties are rarely met leading to misspecification of prediction models. Hybrid methods, particularly regression-kriging (MLRK), which combines ordinary kriging with MLR are also gaining currency in digital soil mapping because of their detailed results and lower prediction errors compared to pure geostatistical, or statistical methods (Hengl et al., 2004). Geographically weighted regression (GWR) is the most recent technique, which has drawn the attention of environmental scientists. GWR was designed to deal with the spatially varying relationships between the target and environmental variables (spatial non-stationarity); thus, the estimated parameters also vary spatially (Wang et al., 2013). Although some comparative studies have shown that it outperforms MLRK in the spatial prediction of SOC stocks (Mishra et al., 2010), its application is still limited. Few studies have also satisfactorily attempted to couple GWR with kriging (geographically weighted regression-kriging; GWRK) to predict the spatial distribution of environmental phenomena. For example, urban heat island in Wrocław, Poland (Syzomanoski and Kryza, 2012) and SOC stocks in Pennsylvania State, USA (Kumar and Lal, 2011; Kumar et al., 2012).

The objective of this study was to estimate and map the spatial distribution of SOC and TN stocks to 30cm depth in the Eastern Mau Forest Reserve using an integrated field sampling, remote sensing, geographical information systems (GIS), and statistical modelling approach. The 30cm depth was consistent with the Intergovernmental Panel on Climate Change (IPCC) guidelines (IPCC, 2006). The Eastern Mau Forest Reserve was appropriate for this study because it had undergone wanton deforestation and degradation since the mid-1990s thanks to ill-advised forest excisions and illegal logging, encroachments, and charcoal burning (Government of Kenya 2009; UNEP 2009). Despite this, no complete studies had been undertaken to quantify the storage and spatial patterns of SOC and TN. This study aimed to fill this gap; hence, contribute knowledge to formulate site-specific and effective measures for ecosystem restoration, sustainable land management and climate change mitigation.

## 2.0 Materials and methods

### 2.1 Study area

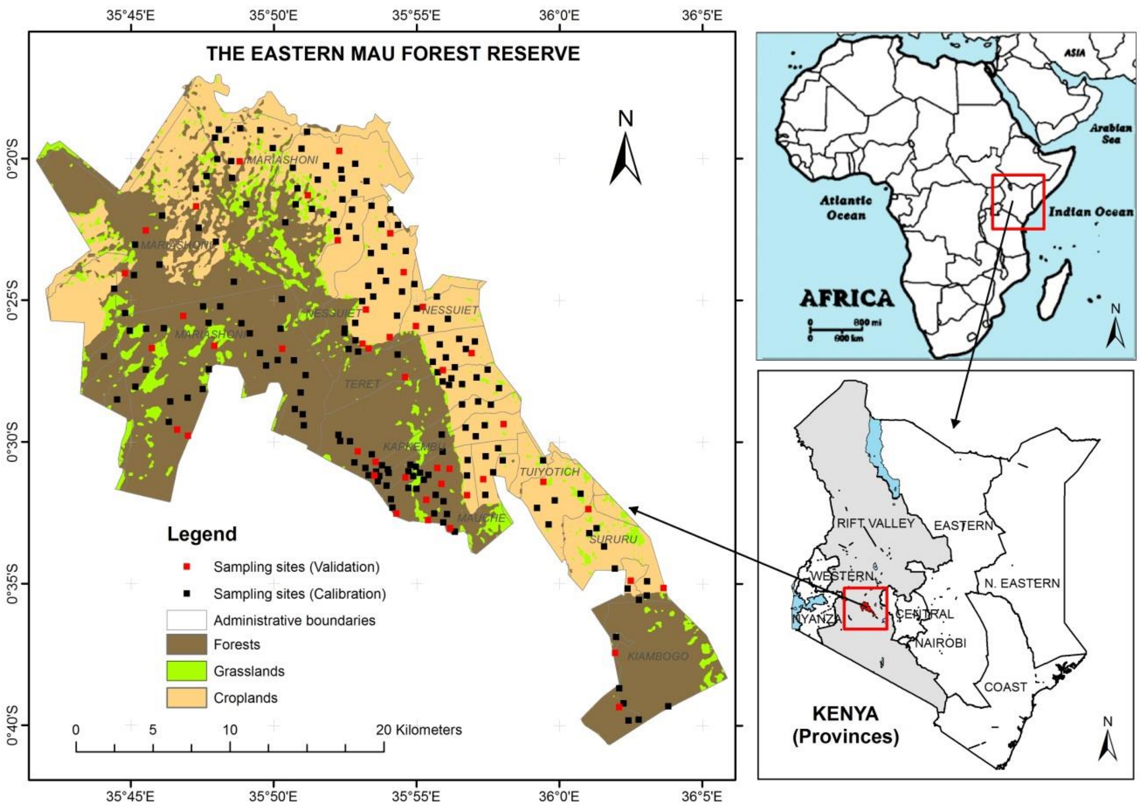



**Fig. 1:** Geographical location of the study area

The Eastern Mau Forest Reserve, which is bounded by the latitudes 0º 15´- 0º 40´S and longitudes 35º 40´- 36º 10´E (Fig. 1), is part of East Africa’s largest closed-canopy indigenous montane forest, and Kenya’s key water catchment area. It covers approximately 650 km$^2$, with the altitudes ranging from 2210 to 3070m above sea level. The climate is cool and humid, with the mean annual rainfall varying between 935 and 1287 mm, and the mean annual temperature ranging from 9.8 to 17.5 °C (Jaetzold et al., 2010). The rainfall pattern is tri-modal with peaks in April, August, and November. The Njoro, Naishi, and Larmudiac Rivers drain the eastern slopes into Lake Nakuru, while the Nessuiet flow northwards into Lake Bogoria, and the Rongai River into Lake Baringo. The area’s physiography and lithology are characterized by major scarps and uplands comprising pyroclastic rocks, such as pumice tuffs, of tertiary-quaternary volcanic age. These soft, light brown rocks have insets of yellow pumice and angular trachyte, which decompose into deep to very deep, dark reddish brown clayey soil aggregates (McCall, 1967). The soils, classified as *Mollic Andosols*, are friable and smeary with humic topsoils (Jaetzold et al., 2010). The major land cover types are forests, grasslands, and croplands. The floristic composition of forests and grasslands include: indigenous tree species, such as *Prunus africana*, *Arundinaria alpina*, *Juniperus procera*, *Olea europaea* ssp. *africana*, *Olea capensis* ssp. *hochstetteri*, *Podocarpus latifolius*, *Nuxia congesta*, *Clematis hirsuta*, *Schefflera volkensii*, and *Dombeya torrida*, exotic tree species, such as *Pinus patula* and *Cupressus lusitanica*, and grass species, such as *Pennisetum clandestinum*. The major crops grown are maize (*Zea mays*), beans (*Phaseolus vulgaris*), wheat (*Triticum aestivum*), and potatoes (*Solanum tuberosum*).

## 2.2 Data sources and pre-processing

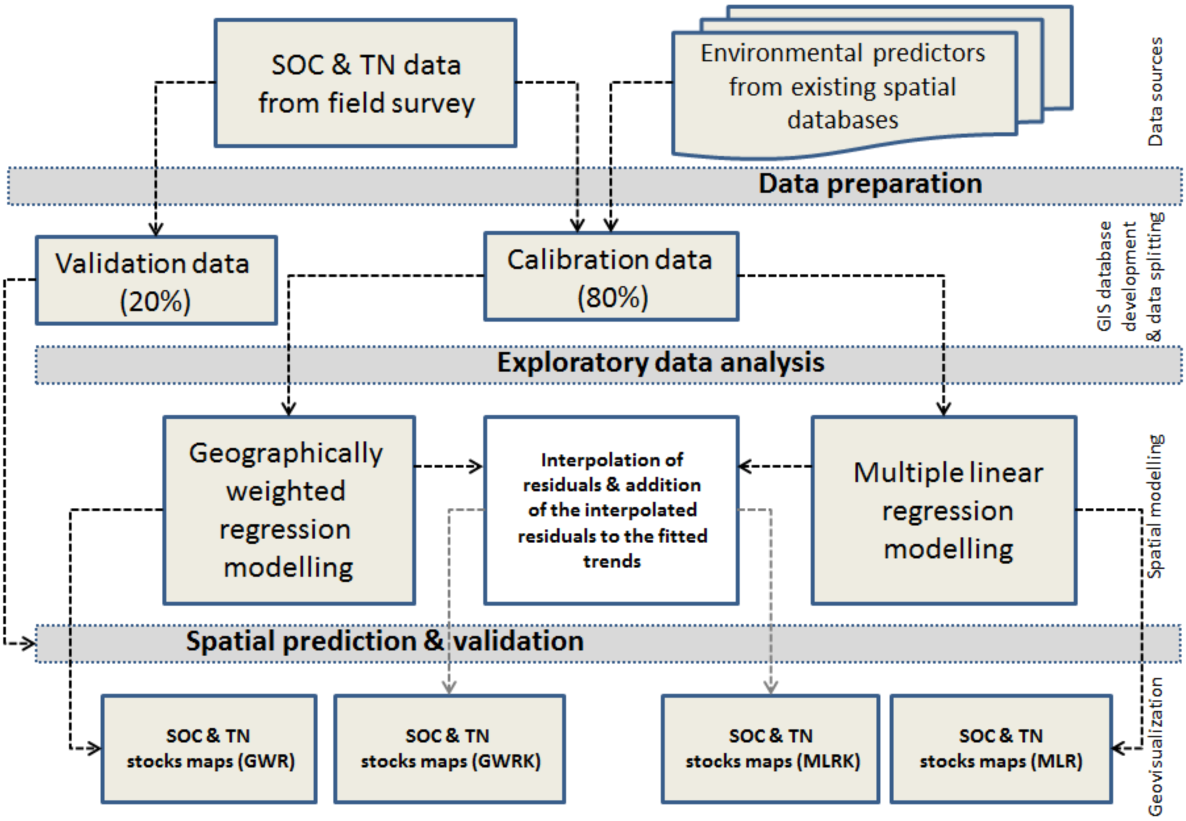



**Fig. 2:** Illustration of the data sources and modelling framework

Figure 2 summarizes the data sources and spatial modelling framework of the study. The overall methodology involved seven major steps: (i) soil sampling and analysis, (ii) pre-processing the environmental predictors and target soil variables, (iii) calibrating the regression-based models, (iv) applying the models, (v) interpolating the regression-based residuals and adding them to the fitted trend surfaces, (vi) validating, and (vii) producing thematic maps of SOC and TN stocks.

### 2.2.1 Soil sampling and analysis

The soil sampling campaign was conducted between June and August 2012. Before the campaign, sampling points were generated in a completely randomized design using agro-ecological zone map as the base in a GIS. A map showing the distribution of these sampling points guided the field visit. In each sampling point, a 30 × 30m plot was laid, and an auger used to collect samples from the centre and four corners of the plot, at 0-15cm and 15-30cm depths. The samples taken from corresponding depths were mixed thoroughly and bulked into one composite sample of about 500g. To determine bulk density (BD), a core sampler (5 cm in diameter and 5cm in height) was used to collect one undisturbed sample at the centre of each plot and each depth. The geographical coordinates, elevation, vegetation, and land management practices were also recorded at each plot. A total of 320 soil samples were collected from 160 sampling plots to analyse the chemical and physical properties, and a similar number to determine BD at the National Agricultural Research Laboratories. Supplementary soil data that had been collected similarly from 60 other sampling plots to evaluate the impact of land cover changes on SOC and TN stocks (Were et al., 2014) were also used. Overall, soil data from 220 sampling plots (Fig. 1) were used to model the spatial distribution of SOC and TN stocks.

The soil samples were air-dried, ground and sieved through a 2mm mesh. The Walkley-Black wet oxidation method (Nelson and Sommers, 1982), Kjeldahl digestion method (Bremner and Mulvaney, 1982), and core method (Blake, 1965) then determined SOC concentrations, TN concentrations, and BD, respectively. These properties were used to calculate SOC and TN stocks (target variables) at each depth. Additional soil properties were also analysed. The hydrometer method (Day, 1965) determined particle size distribution, while the Mehlich method (Okalebo et al., 2002) estimated phosphorous (P) content. A flame-photometer measured the content of potassium (K), an atomic absorption spectrophotometer the contents of calcium (Ca) and magnesium (Mg), and a pH meter measured pH (1:2.5 soil-water) (Okalebo et al., 2002).

Eq. (1) was used to calculate SOC stocks (Mg C $ha^{-1}$) for each depth:

$$SOC_{st} = \frac{SOC}{100} \times BD \times D \times 100 \qquad (1)$$

where: $SOC_{st}$ is the soil organic carbon stock (Mg C $ha^{-1}$), SOC is the soil organic carbon concentration (%, which is then converted to g C $g^{-1}$ soil), BD is the bulk density (g $cm^{-3}$), D is the depth (cm), and 100 is the multiplication factor to convert the SOC per unit area from g C $cm^2$ to Mg C $ha^{-1}$. Stone contents were negligible due to the softness of the volcanic rocks; hence, Eq. (1) does not account for them. Similarly, TN stocks ($TN_{st}$; Mg N $ha^{-1}$) for each depth were computed by substituting TN for SOC in Eq. (1). The SOC and TN stocks in the surface (0-15cm) and subsurface soils (15-30cm) were summed up to obtain the total stocks to 30cm depth.

#### 2.2.2 Remote sensing and GIS analysis

Existing spatial databases provided the twenty candidate environmental predictors that had been selected *a priori* based on *scorpan* conceptual model (McBratney et al., 2003). This conceptual model captures the key soil-forming factors; namely, soil properties (s), climate (c), organisms (o), topography (r), parent material (p), age (a), and space (n). Table 1 provides the sources of climate (mean annual temperature and rainfall), land cover, elevation (digital elevation model; DEM), and Landsat 8 Operational Land Imager (OLI) data. Slope, curvature, aspect, and compound topographic index (CTI) were extracted from the DEM. Eq. (2) extracted the CTI grid:

$$CTI = \ln\left(\frac{A_s}{\tan\beta}\right) \qquad (2)$$

where: $A_s$ is the upslope area and $\beta$ is the slope (McKenzie and Ryan, 1999).

The Normalized Difference Vegetation Index (NDVI) (Eq. (3)) was derived after the digital numbers of OLI band 4 (red; R) and 5 (near infra-red; NIR) were converted to top-of-atmosphere reflectance ($\rho$) (http://landsat.usgs.gov/Landsat8_Using_Product.php).

$$NDVI = \frac{\rho NIR - \rho R}{\rho NIR + \rho R} \qquad (3)$$

Moreover, principal component analysis was performed to reduce the dimensionality, while maximizing the variability of OLI band 2, 3, 4, 5, 6, and 7. The first principal component (PC1), which explained 98% of the variability, was chosen for spatial modelling. All the raster grids were transformed to Universal Transverse Mercator coordinate system (UTM WGS84 Zone 36S) prior to extracting the area of interest from each. The 1km climatic grids were resampled to 30m to synchronize them with the rest. Soil data from the laboratory, including sand content, silt content, clay content, TN concentrations, C concentrations, pH, Mg, Ca, P, and K were also integrated into the geodatabase both as points in vector format and raster surfaces after interpolation by ordinary kriging. Ordinary kriging is a spatial interpolation technique, which has been widely used to optimize the prediction of soil properties at unsampled locations in pedological studies (Chaplot et al., 2010; Pachomphon et al., 2010; Kumar and Lal, 2011; Tesfahunegn et al., 2011; Marchetti et al., 2012; Elbasiouny et al., 2014). Finally, the attribute values of all the other raster grids (e.g., slope, rainfall, temperature) were extracted to the points to allow the analysis of relationships between the target variables and predictors.

**Table 1:** Properties of the environmental predictors for spatial modelling

| Variables | Data format | Date | Source | Scale | Soil-forming factor |
|---|---|---|---|---|---|
| *Target variables* | | | | | |
| 1. SOC stocks | Points | 2012 | Field work | | |
| 2. TN stocks | Points | 2012 | Field work | | |
| *Predictor variables* | | | | | |
| 1. SOC concentration | Raster | 2012 | Interpolated field data | 30m | S |
| 2. TN concentration | Raster | 2012 | Interpolated field data | 30m | S |
| 3. Magnesium | Raster | 2012 | Interpolated field data | 30m | S |
| 4. Potassium | Raster | 2012 | Interpolated field data | 30m | S |
| 5. Calcium | Raster | 2012 | Interpolated field data | 30m | S |
| 6. Clay content | Raster | 2012 | Interpolated field data | 30m | S |
| 7. Silt content | Raster | 2012 | Interpolated field data | 30m | S |
| 8. Sand content | Raster | 2012 | Interpolated field data | 30m | S |
| 9. pH | Raster | 2012 | Interpolated field data | 30m | S |
| 10. Elevation | Raster | - | ASTER GDEM http://gdem.ersdac.jspacesystems.or.jp/ | 30m | R |
| 11. Slope | Raster | - | ASTER GDEM | 30m | R |
| 12. Aspect | Raster | - | ASTER GDEM | 30m | R |
| 13. Curvature | Raster | - | ASTER GDEM | 30m | R |
| 14. CTI | Raster | - | ASTER GDEM | 30m | S |
| 15. Temperature | Raster | 1950-2000 | www.worldclim.org | 1km | C |
| 16. Rainfall | Raster | 1950-2000 | www.worldclim.org | 1km | C |
| 17. Surface reflectance & thermal emission | Raster | 30.05.2013 | Landsat 8 OLI (bands 2,3,4,5,6,7,10 & 11) http://earthexplorer.usgs.gov/ | 30m | C, S |
| 18. NDVI | Raster | 30.05.2013 | Landsat 8 OLI (bands 4 & 5) | 30m | O |
| 19. PC bands | Raster | 30.05.2013 | Landsat 8 OLI (bands 2,3,4,5,6 & 7) | 30m | S |
| 20. Land cover | Raster | 17.01.2011 | Landsat 5 TM; Were et al. (2013) | 30m | O |

Note: SOC=soil organic carbon; TN=total nitrogen; CTI=compound topographic index; NDVI=normalized difference vegetation index; PC= principal component; S=soil properties; C=climate; O=organisms; and, R=topography

## 2.3 Spatial modelling

### 2.3.1 Exploratory data analysis

Descriptive statistics of the target variables were first estimated. This was followed by pairwise Pearson's product-moment correlation analysis to detect collinearity between the predictors, as well as their correlation with the target variables. Predictors entered the model if their correlation with the target variable was, or exceeded 0.2. Also, two highly correlated predictors ($r \geq 0.8$) were retained in a model only if their variance inflation factors (VIFs) did not exceed 10; otherwise, one was removed.

### 2.3.2. Model development

The processed point data from the geodatabase ($n$=220) was randomly split into two: (i) training data ($n$=176) to calibrate the models of SOC and TN stocks, and create prediction surfaces, and (ii) test data ($n$=44) to validate the surfaces. Multiple linear regression (MLR), multiple linear regression-kriging (MLRK), geographically weighted regression (GWR), and geographically weighted regression-kriging (GWRK) techniques were used to calibrate the models.

#### 2.3.2.1 MLR and MLRK

Eq. (4) gives the form of MLR model used to define the relationship between the target variables and predictors at the sampled locations (Montgomery et al., 2006):

$$\mathrm{y_i} = \beta_0 + \sum_k \beta_k \mathrm{x_{ik}} + \varepsilon_i \tag{4}$$

where: $y_i$ is the value of the target variable at $i$th location, $\beta$ are the regression coefficients, $x_i$ is the value of the predictor variable at $i$th location, $k$ is the number of predictors, and $\varepsilon_i$ is the error term.

The ordinary least square estimator fitted the full MLR models, and then all possible regressions variable selection method ranked the best subset models based on Mallow's $C_p$. The final reduced model was selected from the three best subset models after scrutiny for physical correctness. T-tests determined the significance of the model parameters, while analysis of variance $F$ tests determined the significance of the regression at a level of 5%. Residual plots, normal probability plots, measures of influence and leverage (e.g., Cook's

$D$), VIFs, and coefficients of determination ($R^2$) checked the adequacy of the models. Adequate models were applied to create prediction surfaces of the target variables.

To develop MLRK models and prediction surfaces, the deterministic component of the target variable modelled by MLR (Eq. (4)), and the spatially correlated stochastic component modelled by kriging the residuals were summed up. Eq. (5) summarizes the MLRK model (Hengl et al., 2004; Vasques et al., 2010a):

$$y(u_i v_i) = m_{mlr}(u_i v_i) + \varepsilon'_{ok}(u_i v_i) + \varepsilon''(u_i v_i) \tag{5}$$

where: $y(u_i v_i)$ is the target variable at location $(u_i v_i)$, $(u_i v_i)$ are the coordinates of the $i$th location, $m_{mlr}$ $(u_i v_i)$ is the deterministic component, $\varepsilon'_{ok}$ $(u_i v_i)$ is the spatially correlated random component, and $\varepsilon''(u_i v_i)$ is the spatially independent residuals error (noise).

#### 2.3.2.2 GWR and GWRK

Similar predictors were used to build the GWR models to allow comparison with the MLR models. Unlike MLR that assumes spatial stationarity and locational independence, GWR takes into account the spatial location of samples. This allows the estimated parameters to vary locally, which better represent the spatially varying relationships between the target and predictor variables (Zhang et al., 2011). Eq. (6) expresses the form of GWR model (Fotheringham et al., 2002):

$$y_i = \beta_0(u_i v_i) + \sum_k \beta_k(u_i v_i) x_{ik} + \varepsilon_i \tag{6}$$

where: $y_i$ is the value of the target variable at $i$th location, $(u_i v_i)$ are the coordinates of the $i$th location, $\beta(u_i v_i)$ are the regression coefficients, $x_i$ is the value of the predictor variable at $i$th location, $k$ is the number of predictors, and $\varepsilon_i$ is the error term.

The GWR parameters were estimated using adaptive (*bi*-square) spatial kernel functions, where the bandwidth of the samples included for estimation varied with sample density (Fotheringham et al., 2002; Wang et al., 2013). The Akaike Information Criterion (AICc) determined the optimal bandwidth. The estimated parameters were applied to create spatially distributed maps of the target variables.

To develop GWRK models, the deterministic component of the target variable explained by GWR (Eq. (6)), and the spatially correlated stochastic component represented by kriged residuals were added. Eq. (7) provides the form of GWRK model (Kumar et al., 2012):

$$\mathrm{y}(\mathrm{u_i v_i}) = \mathrm{m_{gwr}}(\mathrm{u_i v_i}) + \varepsilon'_{\mathrm{ok}}(\mathrm{u_i v_i}) + \varepsilon''(\mathrm{u_i v_i}) \quad (7)$$

where: $y(u_i v_i)$ is the target variable at location $(u_i v_i)$, $m_{gwr}(u_i v_i)$ is the deterministic component, $\varepsilon'_{ok}(u_i v_i)$ is the spatially correlated random component, and $\varepsilon''(u_i v_i)$ is the spatially independent residuals error.

Additionally, Moran's *I* measured spatial autocorrelation in the residuals. The Moran's *I* values range from -1 to +1, with 0 indicating absence of spatial autocorrelation, and positive values indicating positive autocorrelation and vice versa.

### 2.3.3 Model evaluation

A ten-fold validation procedure was employed to evaluate the prediction surfaces produced by the fitted models. In this procedure, the original dataset ($n$=220) was randomly split into training ($n$=176) and testing ($n$=44) datasets ten times. The training data were used to calibrate models and generate prediction surfaces, while the testing data were used to validate them. Root mean squared error (RMSE) and mean error (ME) were computed from the differences between the predicted and measured values to determine the precision and bias of the predictions, respectively (Eq. (8 and 9)):

$$\mathrm{RMSE} = \sqrt{\textstyle\sum_1^n (y_i} - \hat{y}_i)^2 / \mathrm{n} \quad (8)$$

$$\mathrm{ME} = \textstyle\sum_1^n (y_i - \hat{y}_i) / \mathrm{n} \quad (9)$$

where: $\hat{y}_i$ is the estimated value, $y_i$ is the measured value, and *n* is the number of measured values in the testing data. The ME should be close to zero, while RMSE should be as small as possible. Average ME and RMSE values of the ten-fold validation are reported in this paper. Visual inspection of the spatial patterns of the target variables supplemented statistical validation.

The method with the lowest prediction error indices provided the final estimates of SOC and TN stocks for the study area. To estimate the stocks under different land cover types, the prediction maps were overlaid with the land cover maps. ArcGIS® 10.1, ERDAS IMAGINE® 2013, GWR4, Microsoft Excel® 2010, and R version 3.0.1 (R Core Team, 2013), with add-in packages "sp" (Pebesma et al., 2013) and "automap" (Hiemstra, 2013) performed all the data management, analyses, and geovisualization tasks.

## 3.0 Results

### 3.1 Exploratory data analysis

Table 2 presents the numerical summaries of SOC and TN stocks at the sampled locations. Soil organic carbon stocks range from 42.0 to 193.4 Mg C ha$^{-1}$, with a mean of 102.7 Mg C ha$^{-1}$. The standard deviation is 24.6 Mg C ha$^{-1}$ and coefficient of variation is 23.9%, which suggests moderate variability. The skewness is 0.39 indicating an approximately normal distribution of the data, whereas kurtosis is 0.97 implying less peaked values in the distribution of the data. Similarly, TN stocks vary from 4.2 to 19.1 Mg N ha$^{-1}$, with a mean of 10.3 Mg C ha$^{-1}$. The standard deviation is 2.4 Mg C ha$^{-1}$ and coefficient of variation is 23.8%, while skewness and kurtosis are 0.28 and 0.76, respectively. Again, this shows moderate variability and minimal departure from normality. Hence, spatial modelling of both SOC and TN stocks was performed using the raw, non-transformed data. Pearson's correlation analysis shows that some of the predictors were highly correlated ($r \geq 0.80$), and that only 13 met the threshold correlation ($r \geq 0.20$) with the target variables (Table 3). Thus, the predictors used to develop full models reduced from 20 to 13; namely, elevation, aspect, rainfall, temperature, TN, Mg, silt, clay, land cover, PC1, NDVI, and OLI band 10 and 11 (Table 3). The predictors that were highly correlated include: temperature and elevation, temperature and land cover, elevation and land cover, elevation and OLI band 11, land cover and OLI band 11, and PC1 and OLI band 11. Therefore, VIFs of the predictors in the reduced models were also checked for multi-collinearity.

**Table 2:** Descriptive statistics of SOC and TN stocks (0-30cm)

| Variable | $n$ | Mean | Median | SD | CV (%) | Min. | Max. | Range | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| $SOC_{st}$ | 220 | 102.7 | 103.2 | 24.6 | 23.9 | 42.0 | 193.4 | 151.4 | 0.39 | 0.97 |
| $TN_{st}$ | 220 | 10.3 | 10.3 | 2.4 | 23.8 | 4.2 | 19.1 | 14.9 | 0.28 | 0.76 |

SD=standard deviation; CV=coefficient of variation; $n$=number of observations

**Table 3:** Pearson's correlation coefficient between the predictors and target variables selected for spatial modelling

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. SOC stock | 1.00 | | | | | | | | | | | | | | |
| 2. TN stock | 0.99 | 1.00 | | | | | | | | | | | | | |
| 3. TN content | **0.84** | 0.85 | 1.00 | | | | | | | | | | | | |
| 4. SOC content | 0.85 | **0.84** | 0.99 | 1.00 | | | | | | | | | | | |
| 5. Silt | **-0.41** | **-0.42** | -0.56 | -0.55 | 1.00 | | | | | | | | | | |
| 6. Magnesium | **0.35** | **0.35** | 0.44 | 0.44 | -0.36 | 1.00 | | | | | | | | | |
| 7. Clay | **0.28** | **0.29** | 0.40 | 0.39 | -0.61 | 0.08 | 1.00 | | | | | | | | |
| 8. Temperature | **-0.50** | **-0.50** | -0.63 | -0.63 | 0.28 | -0.04 | -0.34 | 1.00 | | | | | | | |
| 9. Rainfall | **0.44** | **0.45** | 0.56 | 0.55 | -0.23 | 0.25 | 0.10 | -0.61 | 1.00 | | | | | | |
| 10. Elevation | **0.51** | **0.51** | 0.65 | 0.65 | -0.30 | 0.06 | 0.35 | -0.99 | 0.65 | 1.00 | | | | | |
| 11. Aspect | **0.22** | **0.23** | 0.18 | 0.17 | -0.08 | 0.01 | 0.02 | -0.16 | 0.16 | 0.16 | 1.00 | | | | |
| 12. NDVI | **0.30** | **0.30** | 0.39 | 0.39 | -0.25 | 0.07 | 0.24 | -0.50 | 0.25 | 0.50 | 0.11 | 1.00 | | | |
| 13. Land cover | **-0.48** | **-0.48** | -0.54 | -0.53 | 0.31 | 0.00 | -0.41 | 0.83 | -0.46 | -0.84 | -0.16 | -0.56 | 1.00 | | |
| 14. PC1 | **-0.48** | **-0.48** | -0.52 | -0.52 | 0.15 | -0.03 | -0.23 | 0.71 | -0.50 | -0.73 | -0.28 | -0.32 | 0.74 | 1.00 | |
| 15. Landsat 8 OLI band 11 | **-0.58** | **-0.58** | -0.65 | -0.65 | 0.35 | -0.09 | -0.37 | 0.81 | -0.56 | -0.84 | -0.29 | -0.63 | 0.89 | 0.82 | 1.00 |

Note: SOC=soil organic carbon; TN=total nitrogen; NDVI=normalized difference vegetation index; PC= principal component. Bold form shows that the correlation coefficient between the predictors and target variables exceeded the threshold value ($r > 0.2$).

### 3.2 Spatial models

#### 3.2.1 MLR

The subset SOC and TN stock models selected by all possible regressions method had lower Mallow's $C_p$ values (3.8 and 10.3, respectively) than the number of model parameters. Table 4 provides the summaries of the models. Elevation, silt content, TN concentration, and OLI band 11 have significant effects on SOC stocks explaining 72% of its variability (adjusted $R^2$=0.72), whereas OLI band 11, elevation, and SOC concentration have significant effects on TN stocks explaining 71% of its variability (adjusted $R^2$=0.71). Total nitrogen concentrations have the largest magnitude of effect on SOC stocks, while SOC concentrations have the largest magnitude of effect on TN stocks. OLI band 11 has the smallest magnitude of effect on both TN and SOC stocks. Visual analysis of the residual and normal probability plots indicated equality of variance and normality in the distribution of error terms, as well as linearity in the model parameters. The few outliers that were evident on these plots were not sufficiently influential to warrant their removal from the data because Cook's *D* indices were less than 1. Despite the high correlation between elevation and OLI band 11 ($R^2$=0.84), the associated VIFs do not exceed 10 in the models. Moran's indices are very low, yet statistically significant; that is, 0.11 ($p$=0.0141) and 0.08 ($p$=0.0550) for SOC and TN stocks models, respectively. This shows very weak tendency for clustering of similar residuals. The high nugget-to-sill ratios (NSRs) of 78.6% for the residuals of SOC stock model, and 73.7% for the residuals of TN stock model to a range of 4km (Table 6; Fig. 3) also demonstrate this weak spatial structure. However, the spatial dependency of SOC and TN stocks data are moderate (NSRs of 58.1% and 45.6%, respectively) to a range of 4.8km. Total sills for the residuals are 182 Mg C ha$^{-1}$ and 1.9 Mg N ha$^{-1}$, which are close to the variance ($\sigma^2$) estimates of the respective MLR models (170.6 Mg C ha$^{-1}$ and 1.7 Mg N ha$^{-1}$).

**Table 4:** Parameter estimates of the MLR models

| | SOC stocks model | | | | | TN stocks model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | Estimate | SE | *t* value | Pr(>\|t\|) | VIF | Estimate | SE | *t* value | Pr(>\|t\|) | VIF |
| Intercept | 143.502 | 50.757 | 2.827 | 0.0053** | - | 16.741 | 5.131 | 3.263 | 0.0013** | - |
| Silt | 0.443 | 0.202 | 2.191 | 0.0298* | 1.531 | - | - | - | - | - |
| Band 11 | -0.003 | 0.001 | -2.360 | 0.0194* | 3.511 | -0.000 | 0.000 | -2.475 | 0.0143* | 3.489 |
| Elevation | -0.022 | 0.009 | -2.503 | 0.0133* | 3.613 | -0.002 | 0.001 | -2.305 | 0.0223* | 3.558 |
| TN | 178.200 | 12.269 | 14.524 | 0.0000*** | 2.471 | - | - | - | - | - |
| SOC | - | - | - | - | - | 1.597 | 0.106 | 15.103 | 0.0000*** | 1.807 |
| Adjusted $R^2$ | 0.72 | | | | | 0.71 | | | | |
| RMSE | 13.07 | | | | | 1.33 | | | | |
| Moran's I | 0.11 | | | | | 0.08 | | | | |

Significance codes: 0 '***' 0.001 '**' 0.01 '*'

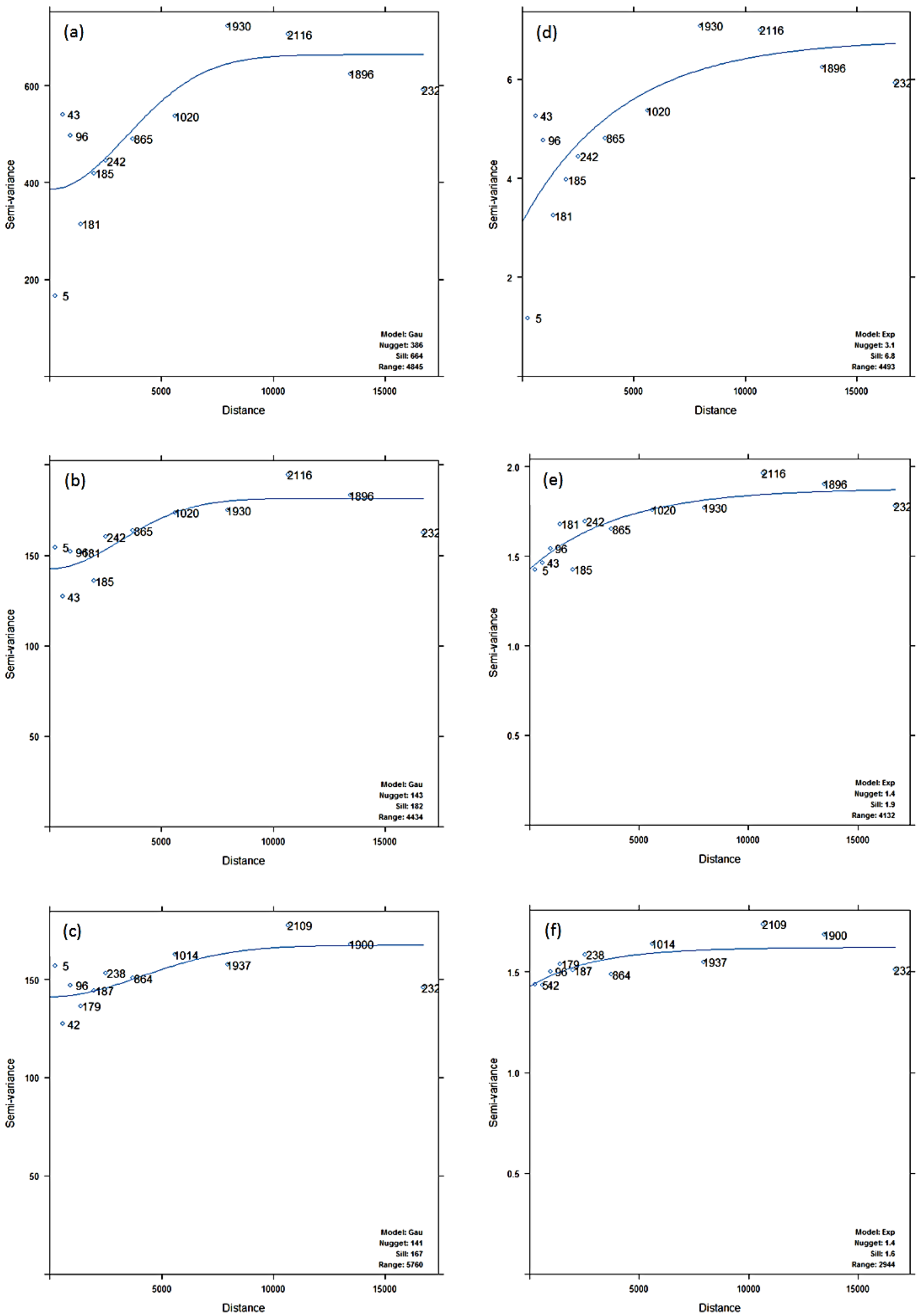


**Fig. 3:** (**a**) Experimental variograms (points) and fitted models (lines) of SOC stocks (**b**) $MLR_{soc}$ residuals (**c**) $GWR_{soc}$ residuals (**d**) TN stocks (**e**) $MLR_{tn}$ residuals, and (**f**) $GWR_{tn}$ residuals. $MLR_{soc}$ and $MLR_{tn}$ refer to the multiple linear regression models for SOC and TN stocks, respectively. While $GWR_{soc}$ and $GWR_{tn}$ refer to the geographically weighted regression models for SOC and TN stocks, respectively.

### 3.2.2 GWR

Table 5 shows the summaries of parameter estimates of the GWR models for SOC and TN stocks. Once more, TN concentrations have the highest magnitude of effect on SOC stocks, while SOC concentrations have the highest magnitude of effect on TN stocks. OLI band 11 has the lowest magnitude of effect on both TN and SOC stocks. Unlike the MLR models, the GWR models show that the magnitude of the effects of predictors varies with sampling location. This means that the interactions between the target variables and environmental factors are spatially non-stationary. Hence, the summaries of GWR estimates are given in ranges instead of mean values. Although the magnitudes vary spatially, the directions of the effects are constant. The Moran's indices are lower than for MLR models; that is, 0.06 ($p$=0.1798) and 0.02 ($p$=0.1620) for SOC and TN stocks models, respectively. This indicates that the GWR residuals are approximately uncorrelated, and that the models are better specified than the MLR models. The high NSRs of 84.4% for the GWR residuals of SOC stocks, and 87.5% for the GWR residuals of TN stocks to a range of 6km and 3km, respectively (Table 6; Fig. 3), also reveal this weak spatial dependency. Total sills for the residuals are 167 Mg C ha$^{-1}$ and 1.6 Mg N ha$^{-1}$, which are lower than for the MLR residuals. However, the range is shorter for the GWR residuals than for the MLR residuals only in the case of TN stocks.

**Table 5:** Parameter estimates of the GWR models

| | SOC stocks model | | | | | TN stocks model | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Parameter | Mean | SD | Min. | Max. | Range | Mean | SD | Min. | Max. | Range |
| Intercept | 129.525 | 31.412 | 50.031 | 199.881 | 149.851 | 14.790 | 4.771 | 8.037 | 26.423 | 18.387 |
| Silt | 0.436 | 0.115 | 0.203 | 0.614 | 0.411 | - | - | - | - | - |
| Band 11 | -0.003 | 0.001 | -0.004 | -0.001 | 0.004 | -0.000 | 0.000 | -0.000 | -0.000 | 0.000 |
| Elevation | -0.021 | 0.007 | -0.041 | -0.004 | 0.037 | -0.002 | 0.001 | -0.005 | -0.001 | 0.004 |
| TN | 177.230 | 23.072 | 142.790 | 238.558 | 95.768 | - | - | - | - | - |
| SOC | - | - | - | - | - | 1.576 | 0.197 | 1.087 | 1.899 | 0.812 |
| Global adjusted $R^2$ | 0.73 | | | | | 0.72 | | | | |
| Global RMSE | 12.86 | | | | | 1.29 | | | | |
| Moran's I | 0.06 | | | | | 0.02 | | | | |

**Table 6:** Parameters of the fitted variogram models for SOC and TN stocks, and the residuals of the respective GWR and MLR models

| Variable | Model | Nugget Mg ha$^{-1}$ | Partial sill Mg ha$^{-1}$ | Total sill Mg ha$^{-1}$ | Range (m) | Nugget-to-sill ratio (%) | Spatial dependence |
|---|---|---|---|---|---|---|---|
| SOC stocks | Gaussian | 386 | 278 | 664 | 4845 | 58.1 | Moderate |
| $MLR_{soc}$ residuals | Gaussian | 143 | 39 | 182 | 4434 | 78.6 | Weak |
| $GWR_{soc}$ residuals | Gaussian | 141 | 26 | 167 | 5760 | 84.4 | Weak |
| TN stocks | Exponential | 3.1 | 3.7 | 6.8 | 4493 | 45.6 | Moderate |
| $MLR_{tn}$ residuals | Exponential | 1.4 | 0.5 | 1.9 | 4132 | 73.7 | Weak |
| $GWR_{tn}$ residuals | Exponential | 1.4 | 0.2 | 1.6 | 2944 | 87.5 | Weak |

### 3.3 Spatial distribution and estimates of SOC and TN stocks

Figures 4 and 5 display the different prediction surfaces of SOC and TN stocks produced by MLR, MLRK, GWR, and GWRK models. The maps reveal similar spatial patterns of SOC and TN stocks meaning that SOC and TN stocks respond similarly to the environmental factors. There is a general decrease of SOC and TN stocks from the west to east. The highest estimates of SOC and TN stocks occur on the western and north-western parts, which according to the environmental data, have higher forest cover, elevations, and SOC and TN concentrations, but lower silt contents and surface temperatures. These *hotspots* are parts of the Logoman, Nessuiet, Kiptunga, and Baraget forests that have not undergone deforestation. The lowest estimates, on the other hand, occur on the eastern side where croplands have been established, including Teret, Nessuiet, Kapkembu, Tuiyotich, and Sururu locations. These *coldspots* are areas with higher crop cover, silt contents, and surface temperatures, but lower elevations, and SOC and TN concentrations. In the northern and south-eastern parts where crop cover is also high, the SOC and TN stocks are moderate to high. The GWR and GWRK prediction surfaces give more realistic pictures of the moderate to high SOC and TN stocks at Sururu forest on the south-eastern most part, which is more degraded than the forests on the western and north-western parts.

The models generated minimum and maximum values that approximate the measured values (cf. Table 3). The MLR and MLRK estimates of TN stocks range from 5.8 to 15.1 Mg N ha$^{-1}$, whereas the GWR and GWRK estimates vary from 5.3 to 15.8 Mg N ha$^{-1}$. Similarly, the MLR and MLRK estimates of SOC stocks range from 56.7 to 146 Mg C ha$^{-1}$, while the GWR and GWRK estimates vary from 55.6 to 146 Mg C ha$^{-1}$.

**Table 7:** Soil organic carbon and nitrogen stocks under different land cover types

| Land cover | Area | SOC stocks | | | | TN stocks | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Min. | Max. | Mean | Total | Min. | Max. | Mean | Total |
| | Ha | Mg ha$^{-1}$ | | | Tg | Mg ha$^{-1}$ | | | Tg |
| Forests | 32228.4 | 75.5 | 142.9 | 110.4 | 3.78 | 7.5 | 15.3 | 11.1 | 0.38 |
| Grasslands | 5509.4 | 66.7 | 129.8 | 103.5 | 0.57 | 6.7 | 12.6 | 10.4 | 0.06 |
| Croplands | 25828.1 | 62.9 | 126.9 | 95.2 | 2.46 | 6.5 | 12.2 | 9.6 | 0.25 |
| **Total** | **65565.9** | | | | **6.81** | | | | **0.69** |

Table 7 gives the magnitude of SOC and TN stocks under different land cover categories based on GWR method, which has lower prediction error indices compared to other methods. Forests stores the highest amounts of SOC and TN (3.78 Tg C and 0.38 Tg N) followed by croplands (2.46 Tg C and 0.25 Tg N), and grasslands (0.57 Tg C and 0.06 Tg N) (1 Tg = $10^{12}$ g = 1 million tons). This is because forests cover the largest area (32,228 ha), while grasslands cover the smallest area (5,509 ha). Overall, the Eastern Mau Forest Reserve stores about 6.81 Tg and 0.69 Tg of SOC and TN, respectively.

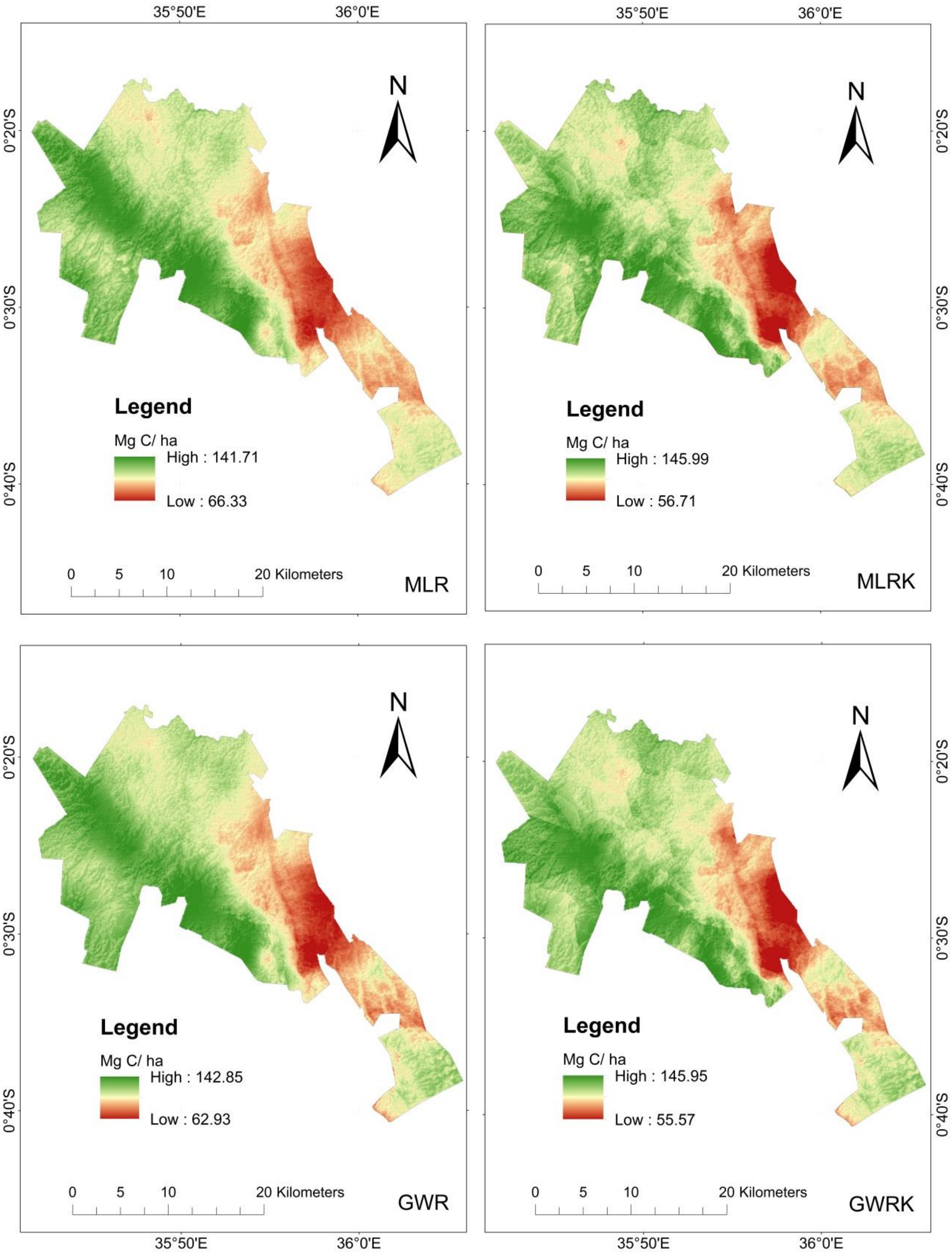



**Fig. 4:** Maps showing the spatial patterns of the predicted SOC stocks using MLR, MLRK, GWR, and GWRK

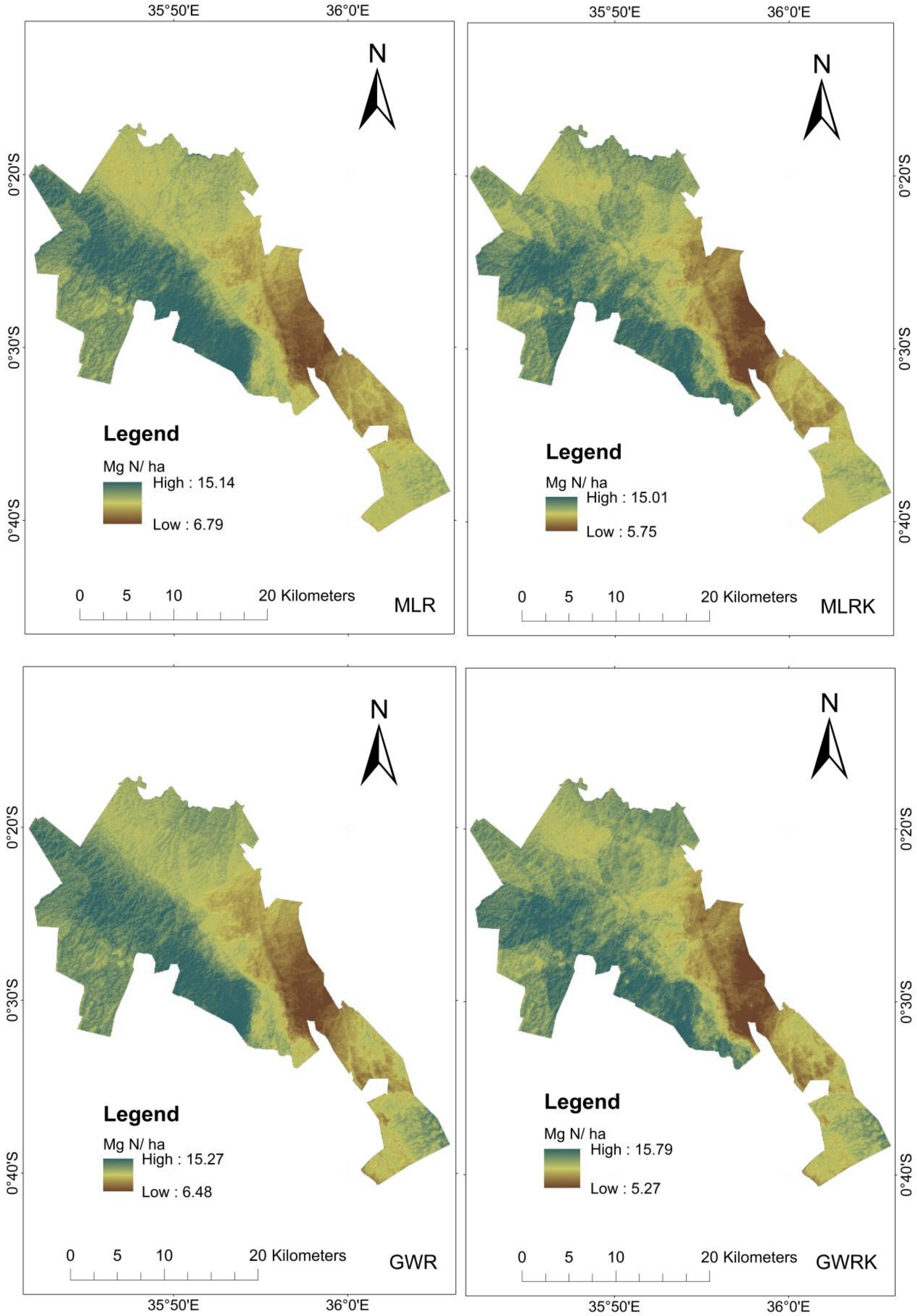



**Fig. 5:** Maps showing the spatial patterns of the predicted TN stocks using MLR, MLRK, GWR, and GWRK

## 3.4 Model evaluation

Table 8 presents the results of ten-fold validation procedure used to quantify the errors attached to the prediction maps of SOC and TN stocks. The average MEs for all prediction models are close to 0, which indicates a small tendency for over- or underestimation. The average RMSEs of the validation range from 16.7 to 19.9 Mg C ha$^{-1}$, and from 1.5 to 1.9 Mg N ha$^{-1}$, which are slightly higher than the RMSEs of the fitted models (13.07 Mg C ha$^{-1}$ and 1.33 Mg N ha$^{-1}$ for MLR models, and 12.86 Mg C ha$^{-1}$ and 1.29 Mg N ha$^{-1}$ for GWR models). This suggests that the models do not predict new data as precise as they fit the original ones. However, the differences in RMSEs are slight, and it should also be noted that the proportions of observations used for calibration and validation were different. The GWR models show better performance in predicting SOC and TN stocks at new locations than MLR models given their lower average RMSEs and MEs. The average RMSEs are also slightly lower than the standard deviations of the measured values (cf. Table 3), which means that incorporation of the predictors and spatial correlation gives better estimations than what would be achieved by just using the measured values for predictions. However, addition of the stochastic part (kriged residuals) to the GWR and MLR outputs does not reduce the prediction errors. The RMSEs of the GWRK and GWR models are similar, and so are the RMSEs of the MLRK and MLR models.

**Table 8:** Summary statistics of the spatial prediction errors

| | $SOC_{st}$ (Mg C ha$^{-1}$) | | $TN_{st}$ (Mg N ha$^{-1}$) | |
|---|---|---|---|---|
| | ME | RMSE | ME | RMSE |
| GWRK | -0.48 | 16.74 | 0.04 | 1.53 |
| GWR | -0.86 | 16.66 | -0.03 | 1.51 |
| MLRK | 0.39 | 19.42 | -0.31 | 1.93 |
| MLR | 0.30 | 19.89 | -0.33 | 1.93 |

ME= mean error; RMSE=root mean squared error

## 4.0 Discussion

## 4.1 Spatial models

The significant effects of elevation, silt content, TN concentration, and OLI band 11 in the SOC models, and elevation, SOC concentration, and OLI band 11 in the TN models implies that topographical, edaphic and climatic factors control the spatial patterns of SOC and TN stocks in the Eastern Mau Forest Reserve. Pachomphon et al. (2010) reported similar combination of controlling factors in Laos, and Li and Shao (2014) in north-western China. The magnitudes of the effects of these predictors indicates that soil properties, particularly

TN and SOC concentrations, are more important than the other factors in determining the observed variability of SOC and TN stocks, respectively. This was expected because of the high statistical correlation between them, and the theoretical tight coupling between C and N cycles. Nitrogen supply increases the net uptake of C in terrestrial ecosystems, which in turn leads to higher inputs of C and N to the soils (Zaehle et al., 2011). The high coefficient of determination (i.e., $R^2$>0.70) obtained for the fitted MLR and GWR models further confirms the explanatory power of these soil properties. In a similar study at four contrasting East African landscapes, Vågen and Winowiecki (2013a) also concluded that intrinsic soil properties determined more the SOC dynamics than other environmental factors alone. The significant effects of OLI band 11 (proxy for surface temperature) and elevation on SOC and TN stocks suggests that: (i) some solutions to the problem of up-scaling soil survey data to landscape level in the region exist in the freely available remotely sensed and DEM data, and (ii) computationally intensive remote sensing- and DEM-derived parameters (e.g., CTI, NDVI) do not always improve the spatial prediction of soil properties. Generally, poor prediction performance ($R^2$<0.50) has been the norm in the region. For instance, Mora-Vallejo et al. (2008) developed MLR and MLRK models using topographical and geomorphological variables that explained less than 25% of SOC variability in south-eastern Kenya. But recently, Vågen et al. (2013b) reported good prediction performance of SOC and pH (i.e., $R^2$>0.70) using Landsat 7 ETM+ imagery and random forest models in Ethiopia.

In terms of spatial structure, the NSRs of raw SOC and TN stocks data revealed moderate spatial dependency (Table 6), which compare with the findings of Sumfleth and Duttmann (2008). This suggests that in the short-range, random and structural processes are equally influential in explaining the spatial variability of SOC and TN stocks. The structural processes that determine the variability of SOC and TN stocks in the Eastern Mau Forest Reserve are the natural soil-forming factors, including topography, soil properties, and climate, while the random processes that explain the remaining variability are human activities, such as illegal logging, encroachments, and charcoal burning, as well as land management practices. In contrast, the residuals obtained from the GWR and MLR models exhibited weaker patterning as evidenced by the low Moran's indices and high NSRs. This means that the global trend models partly explained the variability and spatial correlation of SOC and TN stocks leaving a small, less structured, short-range variation unexplained (Vasques et al., 2010a, 2010b). The unexplained short-range spatial variation reflects the inherent data errors and spatial sources of variations at distances smaller than the shortest sampling interval. Theoretically, this can be resolved by increasing the sampling intensity, but practically, this may be difficult to implement due to resource constraints. The NSRs also hint at the proportion of variation that can be explained by the spatial models. As expected, the NSRs for the MLR models of SOC and TN stocks (79% and 74%) were close to the proportions of variation that the models explained (adjusted $R^2$=72% and 71%).

### 4.2 Spatial distribution and estimates of SOC and TN stocks

The prediction maps revealed spatial patterns of SOC and TN stocks that were similar, and a reflection of the environmental predictors. The given characteristics of the *hotspots* of SOC and TN stocks on the western and north-western parts, in addition to the highly fertile *Andosols* of the area, favours accumulation of SOC and TN stocks. For instance, the high rainfall and low temperatures associated with higher altitudes increase net primary productivity of the forests and decrease SOC turnover. The lower silt content relative to clay content in the forest soils is also an indication of the presence of organo-complexes, or allophane, imogolite, and ferrihydrite clay minerals, which stabilize organic matter and plant nutrients (Lemenih et al., 2005; Chaplot et al., 2010). The smaller pore spaces of clay particles also promote aggregation and physical protection of SOC. In contrast, the characteristics of the *coldspots* of SOC and TN stocks on the eastern side are unfavourable for accumulation of SOC and TN stocks. For example, the higher crop cover is attributed to the conversion of forests to croplands, which began in the mid-1990s. In these croplands, biomass removal after harvesting, erosive processes, and frequent tillage, which breaks up the soil aggregates and alters aeration, can explain the lower SOC and TN stocks (Murty et al., 2002; Smith 2008; Eclesia et al., 2012; Wiesmeier et al., 2012). Thus, the *coldspots* of SOC and TN stocks also highlight the human-induced soil degradation, and sources of C and N emissions. The altitudinal gradient in SOC stocks mentioned above corresponds with previous studies in the Tana River basin, Kenya (Tamooh et al., 2012), while the highest SOC and TN stocks under forests coincide with other studies in the region (Bewketa and Stroosnidjer, 2003; Lemenih et al., 2005; Girmay and Singh, 2012; Demessie et al., 2013). The *hotspots* and *coldspots* on the prediction maps are the locations to target the best management practices for climate change mitigation and sustainable land management. For example, the western and north-western parts need practices that promote retention, whereas the eastern part requires those that enhance accumulation of SOC and TN stocks.

The SOC and TN stocks in the Eastern Mau Forest Reserve to 30cm depth were estimated at 6.81 Tg and 0.69 Tg, respectively. This accounts for 0.36% of the total SOC stock to 30cm depth reported for Kenya (Batjes, 2004). Batjes (2004) also reported that *Andosols* of the humid and semi-humid regions in Kenya stored an average of 9.1 Kg C $m^{-2}$ (91 Mg C $ha^{-1}$) to 30cm depth, which slightly differs from the present findings (i.e., 10.3 Kg C $m^{-2}$ or 102.7 Mg C $ha^{-1}$). This difference can be attributed to the properties of data in the two studies. Batjes (2004) used coarse resolution legacy data from SOil and TERrain (SOTER) database and Africa Land Cover Characteristics (ALCC) database, while the present study used newly collected, fine resolution data to estimate SOC stocks.

### 4.3 Model evaluation

The GWR-based models were better than MLR-based models in predicting new data (Table 8); so, GWR models were chosen to quantify the total stocks of SOC and TN in the area and different land cover types (Table 8). Mishra et al. (2010), Zhang et al. (2011), and Syzomanoski and Kryza (2012) also obtained similar results. Basically, the MLR approach assumed that the environmental factors, which affected the variability of SOC and TN stocks, were spatially stationary. Hence, it represented their relationships using a global statistic. However, such global values can lead to large errors and be very misleading since most of the variability in SOC and TN stocks result from the local interaction of processes (Kumar et al., 2012). In contrast, the GWR approach applied regressions locally, which accounted for both the spatial trends and local variations resulting in superior estimation of SOC and TN stocks. Despite this, the GWR approach also has its own weaknesses; for example, the variation of regression coefficients locally did not lead to selection of different predictors at different locations (Zhang et al., 2011; Kumar et al., 2013). This means that some predictors may have been redundant at some locations. Unlike other studies that reported better performance with regression-kriging methods (Mishra et al., 2010; Kumar et al., 2012; Syzomanoski and Kryza, 2012; Zhang et al., 2012), addition of stochastic component (kriged residuals) to the MLR and GWR outputs did not yield lower prediction errors here. The small proportion of spatially correlated random component in the residuals as indicated by the low Moran's indices and NSRs (Table 6) explains this. Mora-Vallejo et al. (2008) and Li et al. (2013a) also found that MLRK did not outperform MLR.

### 4.4 Limitations of the study

We acknowledge some limitations of the study. Firstly, the soil properties used as predictors were themselves products of interpolation by ordinary kriging. Thus, interpolation errors may have been propagated to the subsequent prediction of SOC and TN stocks. These predictors would have enhanced the prediction accuracy more had they been sampled more intensely than the target variables. Similarly, estimation of SOC and TN stocks under different land cover types was based on land cover maps that had been produced through classification of Landsat 5 TM satellite imagery. Thus, the inherent classification errors may have influenced the estimates of SOC and TN stocks in the different land cover classes. In addition, the auxiliary spatial data (e.g., DEM, Landsat imagery, and climate) were sourced from different databases; hence, they were of different quality. Poor coverage of samples in the south-eastern most and middle parts dominated by thick impenetrable bamboo forests may have also affected prediction accuracy in these areas. Lastly, some soil-forming factors (e.g., parent material and age) were omitted owing to lack of suitable data. Their inclusion, if significant, may improve the predictive power of future models. The foregoing factors introduced uncertainties, the quantification of which was beyond the scope of the study. Future work will assess the implications of error propagation through sensitivity analysis of model parameters estimated using multi-source auxiliary spatial data of different accuracy levels.

## 5.0 Conclusions and recommendations

This study has demonstrated an integrated approach of field sampling, GIS, remote sensing, and statistical analysis to quantify and map the spatial patterns of SOC and TN stocks to 30cm depth in the Eastern Mau Forest Reserve. Based on the results, the conclusions drawn are: (i) forests have the largest SOC and TN pools followed by croplands and grasslands. Overall, the Eastern Mau Forest Reserve stores about 6.81 Tg of C and 0.69 Tg of N, (ii) the *hotspots* of SOC and TN stocks are the native systems on the western and north-western parts, including Logoman, Nessuiet, Kiptunga, and Baraget forests, while the *coldspots* are the human-dominated landscapes on the eastern part covering Teret, Nessuiet, Kapkembu, Tuiyotich, and Sururu locations. Thus, conversion of forests to croplands is a causal factor of soil degradation in this area, and (iii) a mixture of climatic, edaphic, and topographic factors control the observed spatial patterns of SOC and TN stocks; however, soil properties, particularly TN and SOC concentrations are the most important determinants. Despite the limitations, this study provides the first spatially exhaustive soil information for Eastern Mau forest reserve at a finer scale. The resultant outputs will assist to monitor SOC and TN stocks, as well as to formulate spatially targeted climate change mitigation and sustainable land management policies. Also, the approach used offers a cost-effective framework to derive knowledge of soil processes and multi-purpose soil information in other data-poor environments in Eastern Africa.


## Acknowledgements

This study was funded by the Research Council of Norway through the Norwegian University of Life Sciences. Special thanks go to Mr. P. Owenga for technical support during fieldwork and Mr. E. Thairu for excellent driving in difficult field conditions.

## References


Amare, T., Hergarten, C., Hurni, H., Wolfgramm, B., Yitaferu, B., Selassie, Y.G., 2013. Prediction of soil organic carbon for Ethiopian highlands using soil spectroscopy. *ISRN Soil Science* 2013, 720589 (11 pp), http://dx.doi.org/10.1155/2013/720589.

Batjes, N.H., 2004. Soil carbon stocks and projected changes according to land use and management: a case study for Kenya. *Soil Use and Management* 20, 350-356.

Bewketa, W., Stroosnijder, L., 2003. Effects of agro-ecological land use succession on soil properties in Chemoga watershed, Blue Nile basin, Ethiopia. *Geoderma* 111, 85–98.

Blake, G.R., 1965. Bulk density. In: *Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling*, Black, C.A. (ed.). American society of agronomy, Inc., Madison, Wisconsin, USA.

Bremner, J.M., Mulvaney, C.S., 1982. Nitrogen - total. In: *Methods of soil analysis, Part 2. Chemical and microbiological properties*, 2nd edition, Page, A.L. (ed.). American society of agronomy, Inc., Madison, Wisconsin, USA.

Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., Smaling, E.M.A., 2014. Soil organic carbon stocks in the Limpopo National Park, Mozambique: Amount, spatial distribution and uncertainty. *Geoderma* 213, 46-56.

Chaplot, V., Bouahom, B., Valentin, C., 2010. Soil organic carbon stocks in Laos: spatial variations and controlling factors. *Global Change Biology* 16, 1380-1393.

Day, P.R., 1965. Particle fractionation and particle size analysis. In: *Methods of soil analysis, Part 1. Physical and mineralogical properties, including statistics of measurement and sampling,* Black, C.A. (ed.). American society of agronomy, Inc., Madison, Wisconsin, USA.

Demessie, A., Singh, B.R., Lal, R., 2013. Soil carbon and nitrogen stocks under chronosequence of farm and traditional agro-forestry uses in Gambo district, southern Ethiopia. *Nutr Cycl Agroecosys* 95, 365-375.

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* 204-205, 31-42.

Dorji, T., Odeh, I.O.A., Field, D.J., Baillie, I.C., 2014. Digital soil mapping of soil organic carbon stocks under different land use and land cover types in montane ecosystems, Eastern Himalayas. *Forest Ecology and Management* 318, 91-102.

Eclesia, R.P., Jobbagy, E.G., Jackson, R.B., Biganzoli, F., Piñeiro, G., 2012. Shifts in soil organic carbon for plantation and pasture establishment in native forests and grasslands of South America. *Global Change Biology*, 18, 3237-3251.

Elbasiouny, H., Abowaly, M., Abu_Alkheir, A., Gad, A. 2014. Spatial variation of soil carbon and nitrogen pools by using ordinary kriging method in an area of north Nile delta, Egypt. *Catena* 113, 70-78.

Fotheringham, A.S., Brunsdon, C., Charlton, M.E., 2002. *Geographically weighted regression: The analysis of spatially varying relationships*. John Wiley & Sons Inc., England.

Girmay, G., Singh, B.R., 2012. Changes in soil organic carbon stocks and soil quality: Land use system effects in northern Ethiopia. Acta Agriculturae Scandinavica, Section B &#x2013, *Soil & Plant Science* 62, 6, 519-530.

Government of Kenya, 2009. Report of the prime minister's task force on the conservation of the Mau forest complex. [Online]. Available: http://www.kws.org/export/sites/kws/info/maurestoration/maupublications/Mau_Forest_Complex_Report.pdf [Accessed 2014, January 19].

Grimm, R., Behrens, T., Märker, 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.

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

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. *Computers & Geosciences* 33, 1301-1315.

Hiemstra, P., 2013. Classes and methods for spatial data in R. [Online]. Available: http://cran.r-project.org/web/packages/automap/automap.pdf. [Accessed 2013, December 15].

IPCC, 2006. IPCC Guidelines for national greenhouse gas inventories, prepared by the national greenhouse gas inventories programme, Eggleston, H.S., Buendia, L., Miwa, K., Ngara T., Tanabe, K. (Eds). Published: IGES, Japan.

Jaber, S.M., Al-Qinna, M.I. (2011). Soil organic carbon modelling and mapping in a semi-arid environment using thematic mapper data. *Photogrammetric Engineering & Remote Sensing* 77(7), 709-719.

Jaetzold, R., Schmidt, H., Hornetz, B., Shisanya, C., 2010. *Farm management handbook of Kenya, Vol. II. Natural conditions and farm management information, 2nd edition, Part B Central Kenya, Subpart B1a Southern Rift Valley Province*. Ministry of Agriculture, Kenya and German Agency for Technical Cooperation (GTZ), Nairobi.

Karunaratne, S.B., Bishop, T.F.A., Baldock, J.A., Odeh, 1.O.A., 2014. Catchment scale mapping of measureable soil organic carbon fractions. *Geoderma* 219-220, 14-23.

Kheir, R.B., Greve, M.H., BØcher, P.K., Greve, M.B., Larsen, R., McCloy, K., 2010. Predictive mapping of soil organic carbon in wet cultivated lands using classification tree-based models: the case study of Denmark. *Journal of Environmental Management* 91, 1150-1160.

Kumar, S., Lal, R., 2011. Mapping the organic carbon stocks of surface soils using local spatial interpolator. *Journal of Environmental Monitoring* 13, 3128-3135.

Kumar, S., Lal, R., Liu, D., 2012. A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma* 189–190, 627-634.

Kumar, S., Lal, R., Liu, D., 2013. Estimating the spatial distribution of organic carbon density for the soils of Ohio, USA. *Journal of Geographical Sciences* 23(2), 280-296.

Lacoste, M., Minasny, B., McBratney, A., Michot, D., 2014. High resolution 3D mapping of soil organic carbon in a heterogeneous agricultural landscape. *Geoderma* 213, 296-311.

Lal, R., 2004. Soil carbon sequestration to mitigate climate change. *Geoderma* 123, 1-22.

Lamsal, S., Grunwald, S., Bruland, G.L., Bliss, C.M., Comerford, N.B., 2006. Regional hybrid geospatial modeling of soil nitrate–nitrogen in the Santa Fe River watershed. *Geoderma* 135, 233-247.

Lemenih, M., Karltun, E., Olsson, M., 2005. Assessing soil chemical and physical property responses to deforestation and subsequent cultivation in smallholders farming system in Ethiopia. *Agriculture, Ecosystems and Environment* 105, 373–386.

Lesch, S.M., Corwin, D.L., 2008. Prediction of spatial soil property information from ancillary sensor data using ordinary linear regression: Model derivations, residual assumptions and model validation tests. *Geoderma* 148, 130-140.

Li, Y., 2010. Can the spatial prediction of soil organic matter contents at various sampling scales be improved by using regression kriging with auxiliary information? *Geoderma* 159, 63-75.

Li, Q., Yue, T., Wang, C., Zhang, W., Yu, Y., Li, B., Yang, J., Bai, G., 2013a. Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach. *Catena* 104, 210-218.

Li, M., Zhang, X., Pang, G., Han, F., 2013b. The estimation of soil organic carbon distribution and storage in a small catchment area of the Loess Plateau. *Catena* 101, 11-16.

Li, D., Shao, M., 2014. Soil organic carbon and influencing factors in different landscapes in an arid region of north-western China. *Catena* 116, 95-104.

Liu, Z., Shao, M., Wang, Y., 2011. Effect of environmental factors on regional soil organic carbon stocks across the Loess Plateau region, China. *Agriculture, Ecosystems and Environment* 142, 184–194.

Malone, B.P., McBratney, A.B., Minasny, B., Laslett, G.M., 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma* 154, 138-152.

Marchetti, A., Piccini, C., Francaviglia, R., Mabit, L., 2012. Spatial distribution of soil organic matter using geostatistics: a key indicator to assess soil degradation status in central Italy. *Pedosphere* 22, 2, 230-242.

Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., Arrouays, D., 2011. Spatial distribution of soil organic carbon stocks in France. *Biogeosciences* 8, 1053-1065.

Martin, M.P., Orton, T.G., Lacarce, E., Meersmans, J., Saby, N.PA., Paroissien, J.B., Jolivet, C., Boulonne, L., Arrouays, D., 2014. Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale. *Geoderma*, http://dx.doi.org/10.1016/j.geoderma.2014.01.005.

McBratney, A.B., Santos, M.L.M., Minasny, B., 2003. On digital soil mapping. *Geoderma* 117, 3-52.

McCall, G.J.H., 1967. Geology of the Nakuru-Thomson's falls-Lake Hannington area: degree sheet No. 35, S.W. Quarter and 43 N.W. Quarter, Report No. 78. Government Printer, Nairobi.

McKenzie, N.J., Ryan, P.J., 1999. Spatial prediction of soil properties using environmental correlation. *Geoderma* 89, 67-94.

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, 1-13.

Mehrjardi, R.T., Minasny, B., Sarmadian, F., Malone, B.P., 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma* 213, 115-28.

Mishra, U., Lal, R., Liu, D., van Meirvenne, M., 2010. Predicting the spatial variation of the soil organic carbon pool at a regional scale. *Soil Science Society of America Journal* 74, 906–914.

Mishra, U., Riley, W.J., 2012. Alaskan soil carbon stocks: Spatial variability and dependence on environmental factors. *Biogeosciences* 9, 3637-3645.

Montgomery, D.C, Peck, E.A., Vining, G.G. (2006). *Introduction to linear regression analysis*. John Wiley & Sons, Inc., New Jersey.

Mora-Vallejo, A., Claessens, L., Stoorvogel, J., Heuvelink, G.B.M., 2008. Small-scale digital soil mapping in southeastern Kenya. *Catena* 76, 44-53.

Murty, D., Kirschbaum, M.F., McMurtrie, R.E., McGilvray, H., 2002. Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology* 8, 105-123.

Nelson, D.W., Sommers, L.E., 1982. Total carbon, organic carbon and organic matter. In: *Methods of soil analysis, Part 2, Chemical and microbiological properties*, Second edition, Page, A.L. (Ed.). American society of agronomy, Inc., Madison, Wisconsin, USA.

Obade, V.P., Lal, R., 2013. Assessing land cover and soil quality by remote sensing and geographical information systems (GIS). *Catena* 104, 77-92.

Okalebo, J.R., Gathna, K.W., Woomer, P.L., 2002. Laboratory methods for soil and plant analysis: A working manual, 2nd edition. Tropical Soil Biology and Fertility Programme, Nairobi.

Pachomphon, K., Dlamini, P., Chaplot, V., 2010. Estimating carbon stocks at regional level using soil information and easily accessible auxiliary variables. *Geoderma* 155, 372-380.

Pebesma, E., Bivand, R.S., Rowlingson, B., Gomez-Rubio, V., 2013. Classes and methods for spatial data in R. Available: http://cran.r-project.org/web/packages/sp/sp.pdf. [Accessed 2013, December 15].

R Core Team, 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://wwwR-projectorg/.

Razakamanarivo, R.H., Grinand, C., Razafindrakoto, M.A., Bernoux, M., Albrecht, A., 2011. Mapping organic carbon stocks in eucalyptus plantations of the central highlands of Madagascar: A multiple regression approach. *Geoderma* 162, 335-346.

Scull, P., Franklin, J., Chadwick, O.A., McArthur, D., 2003. Predictive soil mapping: a review. *Progress in Physical Geography* 27, 2, 171–197.

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.

Smith, P., 2004. Soils as carbon sinks: the global context. *Soil Use and Management* 20, 212-218.

Smith, P., 2008. Land use change and soil organic carbon dynamics. *Nutr Cycl Agroecosyst*, 81, 169-178.

Sumfleth, K., Duttmann, R., 2008. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators. *Ecological Indicators*, 485-501.

Szymanowski, M., Kryza, M., 2012. Local regression models for spatial interpolation of urban heat island-an example from Wrocław, SW Poland. *Theor Appl Climatol* 108, 53-71.

Tamooh, F., van den Meersche, K., Meysman, F., Marwick, T.R., Borges, A.V., Merckx, R., Dehairs, F., Schmidt, S., Nyunja, J., Bouillon, S., 2012. Distribution and origin of suspended matter and organic carbon pools in the Tana River basin, Kenya. *Biogeosciences* 9, 2905-2920.

Tesfahunegn, G.B., Tamene, L., Vlek, P.L.G., 2011. Catchment scale spatial variability of soil properties and implications on site-specific soil management in northern Ethiopia. *Soil & Tillage Research*, 117, 124-139.

UNEP, 2009. Kenya: Atlas of our changing environment. Division of Early Warning and Assessment (DEWA), United Nations Environment Programme (UNEP). [Online]. Available: http://www.unep.org/dewa/africa/kenyaatlas/. [Accessed 2013, August 28].

Vågen, T.G., Winowiecki, L.A., 2013a. Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters* 8, 015011 (9pp) doi:10.1088/1748-9326/8/1/015011.

Vågen, T.G., Winowiecki, L.A., Abegaz, A., Hagdu, K.M., 2013b. Landsat-based approaches for mapping of land degradation prevalence and soil functional properties in Ethiopia. *Remote Sensing of Environment* 134, 266-275.

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

Vasques, G.M., Grunwald, S., Sickman, J.O., Comerford, N.B., 2010b. Up-scaling of dynamic soil organic carbon pools in a north-central Florida watershed. *Soil Science Society of America Journal* 74, doi:10.2136/sssaj2009.0242.

Wang, K., Zhang, C., Li., W., 2013. Predictive mapping of soil total nitrogen at a regional scale: A comparison between geographically weighted regression and co-kriging. *Applied Geography* 42, 73-85.

Were, K.O., Dick. Ø.B., Singh, B.R., 2013. Remotely sensing the spatial and temporal land cover changes in Eastern Mau forest reserve and Lake Nakuru drainage basin, Kenya. *Applied Geography* 41, 75-86.

Were, K.O., Singh, B.R., Dick, Ø.B. (2014). Effects of land cover changes on soil organic carbon and total nitrogen stocks in the Eastern Mau Forest Reserve, Kenya (Chapter 6). In: Lal, R., Singh, B.R., Mwaseba, D.L., Kraybill, D., Hansen, D.O., Eik, L.O. (Eds.), *Sustainable intensification to advance food security and enhance climate resilience in Africa*. Springer International Publishing, Switzerland, DOI: 10.1007/978-3-319-09360-4_6 (in press).

Wiesmeier, M., Spörlein, P., Geuß, U., Hangen, E., Haug, S., Reischl, A., Schilling, B., von Lützow, M., Kögel-Knabner, I., 2012. Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biology*, 18, 2233-2245.

Wu, C., Wu, J., Luo, Y., Zhang, L., DeGloria, S.D., 2009. Spatial prediction of soil organic matter content using co-kriging with remotely sensed data. *Soil Science Society of America Journal* 73, 1202-1208.

Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J. and Zhu, B., 2008. Storage, patterns and controls of soil organic carbon in the Tibetan grasslands, *Global Change Biology,* 14: 1592-1599.

Zaehle, S., Ciais, P., Friend, A.D., Prieur, V., 2011. Carbon benefits of anthropogenic reactive nitrogen offset by nitrous oxide emissions. *Nature Geoscience* 4, 601-605.

Zhang, Z., Yu, C., Shi, X., Warner, E., Ren, H., Sun, W., Tan, M., Wang, H., 2010. Application of categorical information in the spatial prediction of soil organic carbon in the red soil area of China. *Soil Science and Plant Nutrition* 56, 307-318.

Zhang, C., Tang, Y., Xu, X., Kiely, G., 2011. Towards spatial geochemical modelling: Use of geographically weighted regression for mapping soil organic carbon contents in Ireland. *Applied Geochemistry* 26, 1239-1248.

Zhang, S., Huang, Y., Shen, C., Ye, H., Du, Y., 2012. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information. *Geoderma* 171-172, 35-43.