Masters
Program
in **Geospatial**
Technologies

IMPACT OF LAND COVER CHANGES ON CARBON STOCK TRENDS IN KENYA USING FREE OPEN DATA

Nicodemus Ontweka Nyamari

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS
UNIVERSITAT JAUME·I
Westfälische Wilhelms-Universität Münster

IMPACT OF LAND COVER CHANGES ON CARBON STOCK TRENDS IN KENYA USING FREE OPEN DATA

Dissertation supervised by:

Pedro Cabral, PhD

Co-supervised by:

Joel Dinis Silva, PhD

Jorge Mateu, PhD

February 2019

ACKNOWLEDGEMENTS

Firstly, I would like to express gratitude to Almighty God for His providence of life and good health to compile this dissertation. I am also profoundly grateful to the consortium of the Erasmus Mundus Master's program in Geospatial Technologies for funding and giving me this opportunity to pursue the course. I also express my gratitude to my supervisor for their invaluable comments, assistance, and guidance throughout this research. Finally, utmost gratitude goes to my family, friends, program fellows and mentors, I will always appreciate your contribution to my achievement.

IMPACT OF LAND COVER CHANGES ON CARBON STOCK TRENDS IN KENYA USING FREE OPEN DATA

ABSTRACT

Terrestrial carbon stock estimates information has significant importance in planning decisions for amicable mitigation of global warming and climate change related disasters. However, conventional estimation methods are usually expensive and time demanding particularly on national or regional scales. Therefore, this study sought to estimate and analyze carbon stock changes in Kenya as a consequence of land cover change (LCC) using open data and software to provide affordable and timely solutions. Using Random Forest (RF) decision trees, the land cover for 2028 was modelled from 2004 and 2016 land cover under Business as Usual (BAU) and an alternative, Reducing of Emissions from Forest Degradation and Deforestation (REDD+) scenarios. The modelled land cover maps were thereafter input in InVEST carbon model for estimation and valuation of carbon stock between 2004 and 2028. The results show a 16% decline in carbon stock between 2004 and 2028 with a likelihood of losing up to 21 billion US$ under BAU scenario at a national level. On a regional scale, the results revealed a gradual decline in carbon stock in the Coastal and Central regions of the study area while other regions exhibited mixed results. However, the trend can be reversed by implementation of REDD+ scenario with a possible increase of 1.6% between 2016 and 2028, translating to a gain of approximately 1 billion US$. This study contributes to the understanding of spatiotemporal carbon stock changes under different scenarios for effective spatial planning, land use policy development and keeping balances during natural resource utilization.

KEYWORDS

Ecosystems services

InVEST carbon model

Land cover changes modelling

Random Forest Decision Trees

REDD+

ACRONYMS

BAU	Business as Usual
DEM	Digital Elevation Model
ENVISAT	ESA Environmental Satellite
ES	Ecosystems Services
ESA	European Space Agency
FPR	False Positive Rate
GeoDB	GeoDatabase
GDP	Gross Domestic Product
GHC	Greenhouse Gases
ILRI	International Livestock Research Institute
InVEST	Integrated Valuation of Environmental Services and Tradeoffs
IPCC	International Panel of Climate Change
ITCZ	Inter-Tropical Convergence Zone
LC	Land Cover
LCC	Land Cover Change
MERIS	Medium Resolution Imaging Spectrometer Instrument
NEMA	National Environment Management Authority
NCPD	National Council for Population and Development
REDD+	Reducing of Emissions from Forest Degradation and Deforestation
ROC	Relative Operating Characteristic
TPR	True Positive Rate
US$	United States Dollar
USGS	United State Geological Survey
UTM	Universal Transverse Mercator
WRI	World Resource Institute

INDEX OF THE TEXT

Page

ACKNOWLEDGEMENTS ii

ABSTRACT iii

KEYWORDS iv

ACRONYMS v

INDEX OF THE TEXT vi

INDEX OF TABLES vii

INDEX OF FIGURES viii

1 INTRODUCTION 1

2 STUDY AREA 4

3 DATA AND METHODS 6
- 3.1 Data . . . 6
- 3.2 Methods . . . 6
  - 3.2.1 Data Preprocessing . . . 6
  - 3.2.2 Land cover change modelling . . . 8
  - 3.2.3 Model validation . . . 10
  - 3.2.4 Land cover 2028 scenarios development . . . 10
  - 3.2.5 Carbon stock estimation and change analysis . . . 11

4 RESULTS 13
- 4.1 Land change modelling and validation . . . 13
- 4.2 LCC under BAU and REDD+ scenarios . . . 15
- 4.3 Impact of land cover change on carbon stock . . . 16

5 DISCUSSION 19

6 CONCLUSION 21

References 26

A APPENDICES 27
- A.1 Land cover change modelling python script . . . 27
- A.2 Exploratory data analysis . . . 27

INDEX OF TABLES

Page

3.1 Datasets 6
3.2 Simplified land cover classes; adopted from Niquisse, Cabral, Rodrigues, and Augusto (2017) 8
3.3 Carbon pool assignment adopted from Zhang, Zhan, Zhang, Yao, and Liu (2017) 12

4.1 Land cover model confusion matrix 13
4.2 Area of coverage for each land cover class under all the scenarios 15
4.3 Carbon stock change between 2004 and 2028 under all scenarios 16
4.4 Carbon stock change rate by region between 2004 and 2028 18

A.1 Description of modelling independent variables 27

INDEX OF FIGURES

Page

2.1 Study area map 4

3.1 Flowchart of the overall methodology 7

4.1 Actual (a) and simulated (b) land cover maps for 2016 13
4.2 Relative operating characteristic curves for each land cover class 14
4.3 Historic land cover maps 15
4.4 Land cover maps for 2028 under BAU and REDD+ scenarios 16
4.5 Carbon stock change between 2014 and 2028 17
4.6 Carbon stock changes between 2004 and 2028 17
4.7 Carbon stock change rate (%) by region between 2004 and 2028 18

A.1 Boxplots of land cover proximity distances and biophysical variables . . 27
A.2 Land cover modelling independent variables correlation heat map 28
A.3 RF measure of level of importance of each independent variable in prediction 28

1. INTRODUCTION

Global warming and climate change have elicited greatest of all time environmental crisis thereby predicating anticipated serial and irreversible disasters (Parry et al., 2007). Therefore, it is necessary to aptly combat them sooner than later. These phenomena are mostly attributed to the gradual increase of greenhouse gases (GHG) in the atmosphere (Nakicenovic et al., 2000). According to Gibbs, Brown, Niles, and Foley (2007), the total emission of GHG to the atmosphere is about 70%. Carbon dioxide forms the main contributor to pollutant greenhouse gases. Carbon dioxide is the form by which carbon occurs in the atmosphere.

According to Stocker (2014), the concentration of carbon dioxide in the atmosphere has significantly increased by approximately 40% over the past two and half centuries. The increase of atmospheric carbon dioxide and other GHG gases has been fostered by primarily by burning of fossils fuels, secondly, by land cover changes and land degradation particularly from the conversion of forest or biomass contain vegetation to other land uses.

Forest and biomass vegetative plants play a pivotal role in the carbon sequestering; a process of carbon dioxide removal from the atmosphere and act as carbon stock (carbon storage) reservoirs. These forest carbon reservoirs amount to about 80% of the aboveground carbon and 40% of belowground carbon stocks (Zhang, Zhan, Zhang, Yao, & Liu, 2017). The carbon stock forms an integral part of ecosystems, functioning as a climate regulating ecosystem service (Gómez-Baggethun & Barton, 2013). However, as noted by Pellikka et al. (2018), the reduction of forests due to rapid deforestation and conversion to other land uses mainly to agriculture and urban especially in developing countries are threatening their existence. Hence, expecting to increase in the concentration of carbon dioxide even further.

For instance, as reported by Pellikka et al. (2018), in sub-Saharan Africa, cropland increased by 57% at a rate of 2.3% per year as from 1975 to 2000. Secondly, in the horn of Africa where Kenya is included, cropland increased by 28% at the rate of 1.4% per year between 1990 to 2010. In the East African region; forests, woodlands, and shrub land declined by a mean average of 13.4% between 2002 and 2008 (Pfeifer et al., 2013). The increase of agricultural and other human activities have a direct impact on forests and thus the reduction of ecosystem services and loss of the biodiversity (Pellikka et al., 2018).

A number of carbon conservation based initiatives have been established with a sole purpose of mitigation of atmospheric carbon dioxide concentration increase and maintenance the carbon stock balance (Trexler & Kosloff, 1998). For instance, the United Nations Programme on Reducing Emission from Deforestation and Forest Degradation (REDD+) initiative for the reduction of emissions in developing countries through encouragement of afforestation, reforestation, agroforestry, sustainable intensification of resource utilization and improvement of agricultural practices in contrast to the expansion of agricultural land (UN-REDD, 2019).

For the success of such programs, there is a need to provide reliable and cost-effective measurement tools, constant monitoring, and informed information management. As observed by Niquisse, Cabral, Rodrigues, and Augusto (2017) and Wangai, Burkhard, and Müller (2016), the mapping of the ecosystem service (ES) including carbon stock as an indicator of climate regulation service forms one of the tools for effective management. The mapping of carbon stock and other ES provides a picture for understanding the balance tradeoffs in resource utilization.

Mapping of carbon stock is highly anchored on the provision of data by estimation of carbon through various methods such as biome averages, ground-based biomass measurements, allometric equations, remote sensing methods and a hybrid of one or more methods (Gibbs et al., 2007). These methods are however clouded by some inherent deficiencies such as destructive in nature since the trees have to be cut, time-consuming and/or expensive to maintained particularly at large scale such as national or regional level (Vashum & Jayakumar, 2012).

Alternatively, carbon stocks can be estimated indirectly as a consequence of from the land cover change (Niquisse et al., 2017), (Zhang et al., 2017) and (Kindu, Schneider, Döllerer, Teketay, & Knoke, 2018) and (Leh, Matlock, Cummings, & Nalley, 2013). The estimation can be implemented in InVEST (Integrated Valuation of Environmental Services and Tradeoffs) tool (Sharp et al., 2014). The InVEST tool uses a simplified carbon cycle model that estimates different carbon pools; above-ground, below-ground, dead and soil organic carbon matter based on each land cover classes. Therefore, it provides a simplified means of estimating carbon storage particularly over a large coverage area for trend analysis.

According to the review on ecosystem services studies in Africa carried out by (Wangai et al., 2016), there are little research works on ecosystem services quantification and mapping due to lack of adequate data and heterogeneity of the influencing factors. The review also reported that most of the studies were either context-based (Willcock et al., 2016), or specific landscape-based (Pellikka et al., 2018) or study area coverage is limited to the local level (Kindu et al., 2018). The limited number of studied in Africa was also echoed by (Zhang et al., 2017). In their study, (Zhang et al., 2017), highlighted a need for carbon stock assessment based on multiple land cover classes rather than single land cover class change evaluation and they also noted their study was one of its kind within the East African region. However, none of the studies has attempted to demonstrate how the existing policies such as REDD+ can be actualized through spatial analysis and mapping more particularly on a large geographic scale. The lack of such studies forms an important obstacle in informed decisions making process for the full realization of the policy missions.

Therefore, the main aim of this study was to analyze the trends of carbon stock as an indicator of climate regulation ecosystem service between 2004 and 2028 under different (BAU and REDD+) scenarios based on the land cover changes in Kenya.

The specific objectives of the study are:

- To model the land cover changes in the study area between 2004 and 2028 for each land cover class i.e. cropland, forest, non-forest vegetation such as shrubs and grassland, built-up areas, bare areas, and water;
- To estimate the carbon stock changes in the study area between 2004 and 2028 using InVEST carbon model under BAU and alternative REDD+ scenarios;
- To analyze carbon stock change trends under BAU and REDD+ scenarios.

Therefore, this study expects to provide information for the understanding of effects land cover dynamics on carbon stock under different scenarios for effective spatial plan developments, land use policy development, actualization of REDD+ initiative, and keeping balances during natural resource utilization.

2. STUDY AREA

The study area's full name is the Republic of Kenya. It is located in the Eastern region of Africa with an approximate area of $580,000Km^2$. The country is subdivided into eight regions (with 47 proactive county administrations) as shown in Figure 2.1. It is geographically located along the equator, 5^o N and 5^o S of latitude, between 34^o E and 42^o E of longitude with a vast diversity of landforms ranging from escarpments, plains, mountains, highlands and coastal strip to the southeast.

Kenya is a lower middle-income country with an estimated Gross Domestic Product (GDP) of US$ 79.3 billion as per 2017. It has approximately population of 50 million people and a growth of approximately 3% per year (World Bank, 2019). According to NCPD (2018), about 30% of the population resides in the urban areas, the increase of urban population occurred by almost double in the last two decades. Despite the rural-urban migration, there is over-exploitation of natural resources occurring in the rural areas. Kenya climate is mainly influenced by the Inter-Tropical Convergence

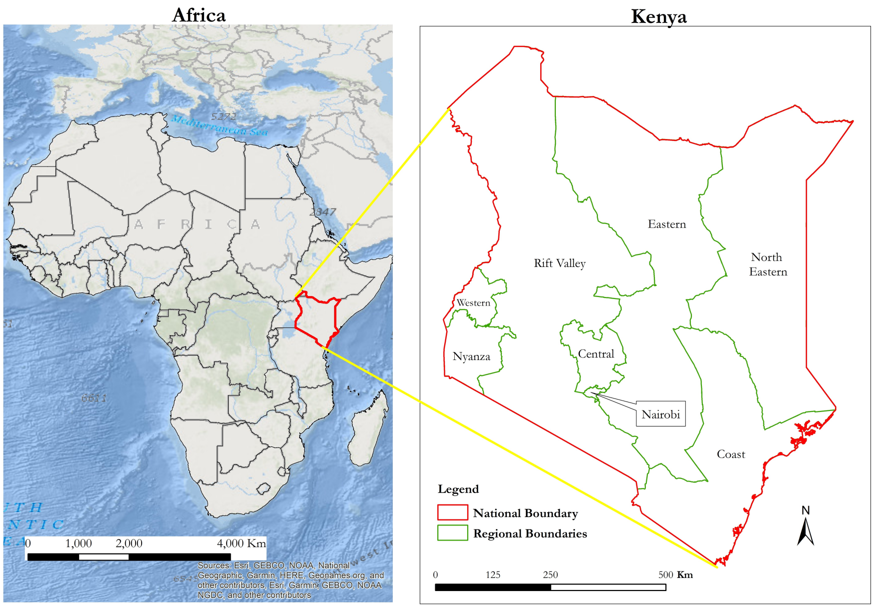

Figure 2.1: Study area map

Zone (ITCZ) although it is considerably varying across the country. The north and north-eastern regions are normally very dry thus classified as Arid and Semi-Arid Lands (ASALs) and it is approximately 85% of the country total land (Barrow & Mogaka, 2007). The coastal region is characterized by hot and humid temperature. The western, central and coastal regions mainly receive a high amount of rainfall coming in two wet seasons; long rains as from April to June and short rains as from October to December (NEMA, 2015).

The country's main natural resources are forests, aquatic and marine life, water and agricultural land though they are now under intense pressure due to over-utilization to cater for ever-growing population, excessive deforestations, degradation, coastal modification and poor management (NEMA, 2015). Thus leading to erratic climate changes and extremes, which are threatening the support of livelihoods services such as food security and drop of the valuable benefits obtained from the utilization of resources (Government of Kenya, 2010).

3. DATA AND METHODS

3.1 Data

The datasets include land cover maps and other ancillary data representing biophysical properties that were considered to influence land cover changes (Table 3.1). The biophysical properties were; slope derived from DEM, distance from the roads, distance to the cities and major towns, and soil type (Li & Yeh, 2002). The 2004 and 2009 land cover maps had 21 classes derived from ENVISAT MERIS sensor (Table 3.2). The maps were automatically generated with an overall accuracy of about 70% (Bontemps et al., 2011) while 2016 land cover map has 9 classes derived from Sentinel 2A sensor and was also automatically generated at an overall accuracy of 65% (Ramoino, Pera, & Arino, 2018).Refer to Table 3.2 for details on reclassification and harmornization of the land cover classess.

Dataset	**Spatial Resolution**	**Source**	**Data Type**	**Purpose**
Land cover maps for 2004 and 2009	300m	ESA/ESA GlobCover Project (http://due.esrin.esa.int/page_globcover.php)	raster	Derive LC classes, compute the LC neighborhood functions and estimation of carbon stock
Land cover map for 2016	20m	ESA Africa Land cover Project (http://2016africalandcover20m.esrin.esa.int)		
Road networks		WRI (https://www.wri.org/resources/data-sets/kenya-gis-data)	vector	Computation of the distance to roads, cities and major urban centers, respectively
Cities		ILRI, (https://data.ilri.org)		
Major town centers				
Soil				Derive soil classes
DEM	1 Arc-second	USGS (https://earthexplorer.usgs.gov)	raster	Computation of slope

Table 3.1: Datasets

3.2 Methods

The overall methodology of this was executed in two major steps; land cover change modelling using Random Forest decision trees algorithm (Breiman, 2001) and thereafter, carbon stock estimation using InVEST carbon model (Figure 3.1).

3.2.1 Data Preprocessing

After acquiring prerequisite data, preprocessing was carried out to generate compatible data for land cover change modelling in a machine learning environment (Charif, Omrani, Abdallah, & Pijanowski, 2017). The following procedures were carried out;

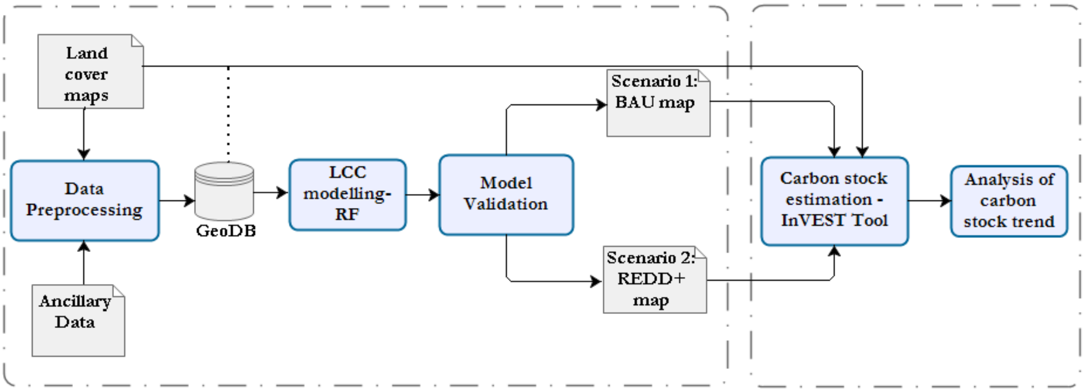

Figure 3.1: Flowchart of the overall methodology

- **Data preparation**: using QGIS 3.4 (QGIS Development Team, 2018), all the datasets were projected to UTM coordinates systems, clipping to the same extent, rasterization of vector data, generation of slope from DEM, post-classification enhancement of the land cover maps using majority filter, land cover classes simplification (through reclassification) to the same class legend (Table 3.2) and lastly, resampling of the data to the same spatial resolution.
- **Decomposing of multi-label land cover classes to binary rasters**: using to_categorical function in Keras library (François Chollet et al., 2015), the land cover classes were decomposed to binary to get six rasters, one for each land cover class. The binary rasters were important for computation of neighbourhood functions and modelling as described in subsection 3.2.2
- **LC classes neighborhood function computation**: using QGIS 3.2 raster proximity distance tool, the neighborhood proximity for each pixel to the nearest target pixel distances were computed. The neighborhood proximities are analogous to the number of neighborhood pixels considered in the cellular automata for land cover change modelling (Shafizadeh-Moghadam, Asghari, Tayyebi, & Taleai, 2017). The raster proximity distances were also computed for distance to the road network, distance to cities and major towns.
- **Normalization/Rescaling**: using RStudio (R Core Team, 2018), the minimum-maximum normalization was applied to the neighborhood proximity rasters, distance rasters and slope. This transformed all the variable values to be in the range between 0 and 1 for compatible manipulation in machine learning (Li & Yeh, 2002).

All the raster layers were then imported to python 3.6 (Python Software Foundation,

Code	Land cover category	Simplified legend	Reclassified class code
11	Post-flooding or irrigated croplands (or aquatic)	Cropland	1
14	Rainfed croplands	Cropland	1
20	Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)	Cropland	1
30	Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)	Non-Forest Vegetation	3
40	Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m)	Forests	2
50	Closed (>40%) broadleaved deciduous forest (>5m)	Forests	2
60	Open (15-40%) broadleaved deciduous forest/ woodland (>5m)	Forests	2
70	Closed (>40%) needleleaved evergreen forest (>5m)	Forests	2
90	Open (15-40%) needleleaved deciduous or evergreen forest (>5m)	Forests	2
100	Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)	Forests	2
110	Mosaic forest or shrubland (50-70%) / grassland (20-50%)	Non-Forest Vegetation	3
120	Mosaic grassland (50-70%) / forest or shrubland (20-50%)	Non-Forest Vegetation	3
130	Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m)	Forests	2
140	Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)	Non-Forest Vegetation	3
150	Sparse (<15%) vegetation	Non-Forest Vegetation	3
160	Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily) - Fresh or brackish water	Forests	2
170	Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water	Forests	2
180	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water	Non-Forest Vegetation	3
190	Artificial surfaces and associated areas (Urban areas >50%)	Built-up areas	4
200	Bare areas	Bare areas	5
210	Water bodies	Water	6
Land cover 2016			
1	Tree cover areas	Forests	2
2	Shrubs cover areas	Non-Forest Vegetation	3
3	Grassland	Non-Forest Vegetation	3
4	Cropland	Cropland	1
5	Vegetation aquatic or regularly flooded	Forests	2
6	Lichens Mosses / Sparse vegetation	Non-Forest Vegetation	3
7	Bare areas	Bare areas	5
8	Built up areas	Built-up areas	4
10	Open Water	Water	6

Table 3.2: Simplified land cover classes; adopted from Niquisse, Cabral, Rodrigues, and Augusto (2017)

2016) for land cover change modelling.

3.2.2 Land cover change modelling

Before processing data for LCC modelling, descriptive and correlation analysis using the interquartile range boxplot and correlation heat map were carried out to understand and determine relationships between the variables, establish the of the existence of the missing data and outliers, and the distribution within the datasets (Apendix A.2).

Random Forest (RF) decision trees algorithm were used to develop a model for the production of 2028 land cover maps under BAU and REDD+ scenarios. The two scenarios maps were produced based on 2004 land cover map, ancillary data, and 2016 land cover map. Random forest algorithm is a non-parametric technique that generates estimators which fits a number of decision trees on various sub-sets of the training data set (Breiman, 2001). The subsets are randomly drawn/split in from the training data (Rahmati, Pourghasemi, & Melesse, 2016). To improve the accuracy and avoid overfitting the average votes of the popular class that maps an input is carried out(Pedregosa et al., 2011). According to (Breiman, 2001), RF allows measuring of the level of importance for each variable in contribution to prediction accuracy. Estimation of the importance of variable n, is done by random permutation of all its' values in the out-of-bag subset for each estimator, if the out-of-bag error increases, it indicates the importance of the variable (Gislason, Benediktsson, & Sveinsson, 2006). The measure of the level of importance is necessary to understand variable interactions.

The procedures followed were; importing of all pre-processed rasters, reshaping raster arrays, exploratory data analysis, training and fine-tuning of the model and lastly, model validation.

The imported raster datasets were converted into numeric column vectors using gdal (Warmerdam, 2008) and numpy libraries (Oliphant, 2006). Using to_categorical function in Keras library (François Chollet et al., 2015), One-Verses-All (OVA) encoding technique (Tayyebi & Pijanowski, 2014) was applied to the land cover map for 2004 and soil column vectors and thus producing a matrix of 6 columns and 5 columns for 6 land cover classes and 5 soil types respectively. Encoding of land cover and soil classes allows for the conversion of categorical variables to numerical variables i.e. 0 and 1 (Tayyebi & Pijanowski, 2014). The 2004 land cover matrix, the soil matrix and column vectors for the biophysical properties were combined to form a larger matrix 22 columns which represented 21 independent variables and 1 dependent variable (2016 land cover). The rows represented data for each pixel and columns represent the variables. The 21 independent variables consisted of 6 binary vectors for 2004 land cover classes, and 15 biophysical driving variables (5 binary vectors for soil classes, slope, distance to roads, distance to major cities, distance to major towns and 6 land cover classes neighborhood proximity raster).

The training data was set by randomly sampling 18000 rows (3000 for each land cover

class). The training data was split into 2 by a ratio of 67% to 33% for training and testing respectively. The splitting of training data was randomly shuffled 10 times which iteratively trained the model (10 fold cross-validation)(Kohavi, 1995). This helps to avoid the underfitting/overfitting of the model.
Using RF classifier in the scikit-learn library (Pedregosa et al., 2011), the LCC model was developed. The sensitivity analysis was carried out by severally changing the number of estimators and the criteria of measuring of quality of the trees while keeping other parameters default. The criteria of measuring of quality of trees for information gain can either be Gini impurity index or entropy (Rahmati et al., 2016). In this study, 200 number of estimators and entropy criterion produced the highest accuracy, thus were used to build the prediction model.
After building the model, it was applied to all independent variables to generate modelled values for 2016 which were compared with the 'actual' 2016 values for model validation.

3.2.3 Model validation

After training the model, the performance evaluation of the model was conducted to assess for under-fit or over-fit of the model (Francois Chollet, 2017). In this case, a confusion matrix (Congalton, 1991) was used to provide the validation parameters; overall accuracy, user accuracies, producer accuracies, kappa value. A confusion matrix was derived from a pixel-by-pixel comparison of 2016 modelled and 'actual' values. Thus, giving information on agreement between the two set of values.
Additionally, relative operating characteristic (ROC) was used to compare the outcomes of modelled values with the 'actual' (reference) values of 2016 land cover (Pontius Jr & Batchu, 2003). ROC calculates the proportion true-positive rate (TPR) and false positive rate (FPR) for a number of thresholds and relates them to each other in a graph. It then measures the area under the curve which should vary between 0.5 (random fit) and 1 (perfect fit) (Mustafa, Cools, Saadi, & Teller, 2017).

3.2.4 Land cover 2028 scenarios development

After model fine-tuning, validation of the model, the independent variable matrix for modelling 2028 land cover was constructed. The construction was done by applying one-hot encoded to 2016 land cover column vector and combining them with the biophysical variables and neighborhood proximity rasters for 2016 land cover classes. Using the trained model, the 2028 land cover values were predicted based on the constructed independent variables matrix. The values were reshaped back to the dimensions of the rasters and given UTM georeferenced coordinated system to form 2028 BAU land cover map.

Similarly, for the development of the alternative scenario map; REDD+, a mask was applied to cropland, forest and built-up areas constraining them from changing, however, other classes could change to these classes. Additionally, to allow the forest to increase, land cover classes were restrained from changing to cropland. This is by the fact that cropland give competition to the forest as reported by (Houghton & Nassikas, 2017). The masking was done by sampling training data which forest, cropland and built-up areas pixel values for 2004 land cover remained the same in 2016 land cover while other classes were sampled regardless on the classes they were in 2016 except those pixels that changed to cropland. The sampled data was then used to re-trained the model. The procedure described under BAU was repeated but predicted carried out using the retrained model to produce 2028 REDD+ scenario land cover map. The accuracy assesses for REDD+ scenario was not carried because the procedure was considered already biased but the training predicted accuracy was observed to be higher than the training accuracy under BAU as it was expected.

3.2.5 Carbon stock estimation and change analysis

In considerations that all the land cover maps were produced in the same working environment; all in UTM coordinate system and spatial resolution. Hence, preprocessing was not be necessary at this stage.

InVEST carbon model was used to estimate carbon stock and value carbon sequestration (Sharp et al., 2014). The tool essentially requires the current land cover map which is used as a baseline year for the computation of carbon stock, carbon pools, and future land cover map(s), price per ton for carbon sequestration valuation. In addition, the REDD+ scenario land cover map for the reference standards assessment and is considered as one of the future land cover maps. The carbon pool consisted of a table contains all fundamental carbon storage media i.e. above-ground biomass, below-ground biomass, soil organic matter and dead organic matter for each particular land cover class (Table 3.3). The dead organic matter was assigned zero to omit it from this study analysis. Additionally, water was assigned zero since it was considered to have negligible carbon content. For the valuation of carbon sequestration, price value of US$232 per ton of carbon was used based on an avoided costs (Tol, 2009). The carbon stock is computed in megatons per hectare.

As implemented by Niquisse et al. (2017), the carbon stock change analysis for the study area at the national and regional level was done for each epoch; 2004, 2009, 2016 and 2028 BAU scenario and 2028 REDD+ scenario. The calculation was done based previous set baseline situation and carbon stock change indices according to equation 3.1.

LC code	**LC Name**	**C above-ground**	**C below-ground**	**C in soil**
1	Cropland	44.43	29.3	10.47
2	Forest	132.44	26.14	15.62
3	Non-Forest	32.2	1.1	18.67
4	Built-up Areas	3	0.6	13.5
5	Bare Areas	3.5	0.35	16.7
6	Water	0	0	0

Table 3.3: Carbon pool assignment adopted from Zhang, Zhan, Zhang, Yao, and Liu (2017)

$$\Delta C_{t+1} = \left[\frac{C_{t+1} - C_t}{C_t}\right] \times 100 \tag{3.1}$$

Where ΔC_{t+1} is the change in carbon stock percentage at time $t+1$, C_t is the carbon sock at time t, and Ct+1 is carbon stock at time $t+1$.

The output results were imported into ArcMap 10.6 (ESRI, 2018) for further analysis and visualization.

4. RESULTS

4.1 Land change modelling and validation

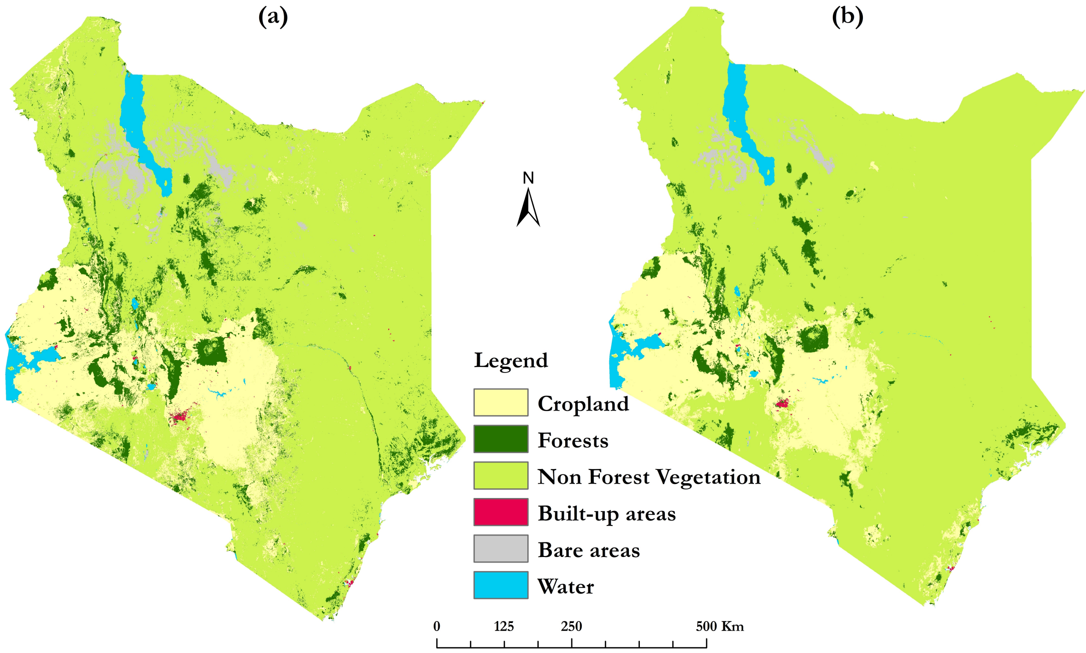

Figure 4.1: Actual (a) and simulated (b) land cover maps for 2016

		Modelled Land Cover							Commission error	Users accuracy
		1	2	3	4	5	6	Sum	(%)	(%)
Actual Land cover	1	1216175	82244	240725	3270	157	1111	1543682	21.2	78.8
	2	41763	255989	76698	232	642	370	375694	31.9	68.1
	3	487697	348502	6635377	1920	88994	4919	7567409	12.3	87.7
	4	270	121	254	4051	4	17	4717	14.1	85.9
	5	71	1318	14187	4	43659	189	59428	26.5	73.5
	6	1657	1826	5014	76	1452	814115	824140	1.2	98.8
Sum		1747633	690000	6972255	9553	134908	820721	10375070		
Omission error	(%)	30.4	62.9	4.8	57.6	67.6	0.8		**Kappa value**	0.7
Producer accuracy	(%)	69.6	37.1	95.2	42.4	32.4	99.2		**Overall accuracy**	86.5

Table 4.1: Land cover model confusion matrix

The comparison between the actual and simulated land cover map for 2016 modelling produced an overall accuracy of 86.5% and a kappa value of 0.7 indicating a good level of agreement (Table 4.1 and Figure 4.1). The user accuracies computed for each class were above 65% (Table 4.1) and Figure 4.2. Additionally, the ratio between the true positive rate (TPR) and false positive rate (FPR) showed that water had the highest value of 1 indicating almost perfect fit between simulated and actual land cover while bare areas and forest had the least ROC value of 0.66 and 0.68 respectively. The high ROC value for water can be attributed to its' distinct characteristics from other classes making it easy to be discriminated during modeling. However, some classes had a low producer accuracy exhibiting existence of certain bias. The low user and producer's accuracies, particularly for bare areas and forest can be attributed to the aggregation of land cover classes from the original land cover dataset as shown in Table 3.2. For instance, land cover code number 30 and 120 going by legend name; *Mosaic vegetation (grassland/shrubland/forest) (50-70%)/cropland (20-50%) and Mosaic grassland (50-70%) / forest or shrubland (20-50%)* is quite hard to generalize them into one class without causing the misclassification during modelling and even causing the occurrence of other land cover class where they less likely to occur. Moreover, if the land cover classes exhibit a high tendency of change to other classes then the model does not easily stabilize.

The ROC value indicated the independent variables were adequate to model land cover since the values were greater than 0.5 thus the model can be reliable for prediction (Kindu et al., 2018).

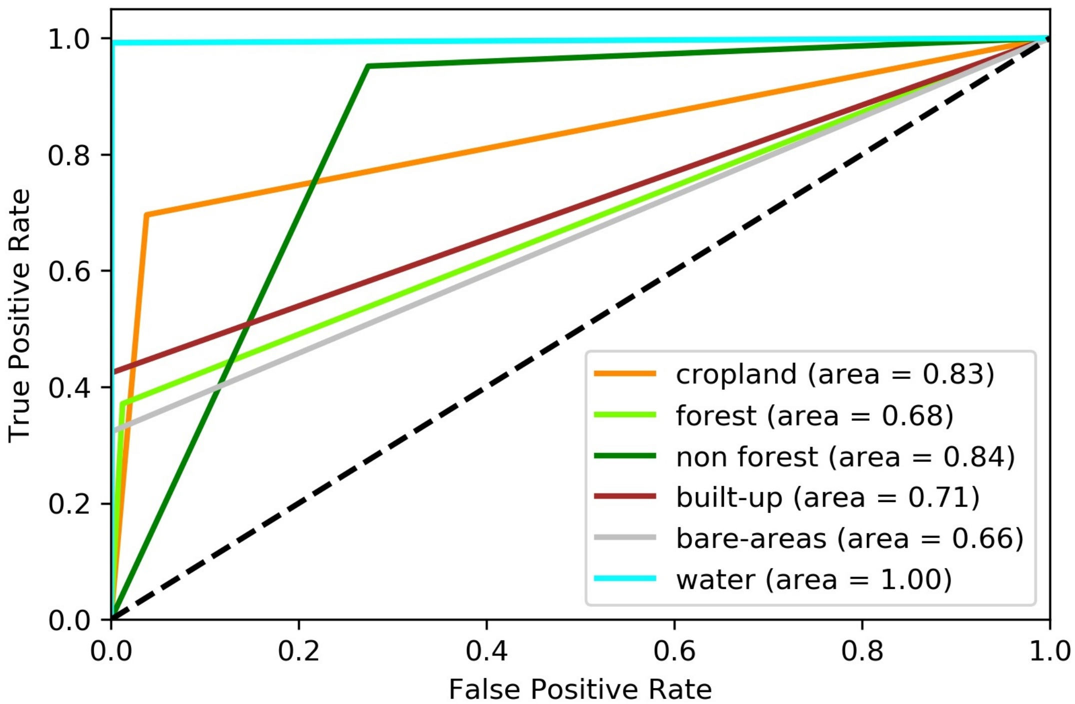

Figure 4.2: Relative operating characteristic curves for each land cover class

4.2 LCC under BAU and REDD+ scenarios

Having reclassified and modelled the land cover for the study period, it can be observed that cropland gradually increased as from 2006 to 2016, then it will slightly decrease or remain constant between 2016 and 2028 under BAU. Forest gradually decreased from 15.7% to 6.9% between 2004 and 2016 and it is expected to go down further to less than 2% by 2028 under the BAU scenario. The non-forest vegetation is expected to increase from 63.4% to 79% between 2004 and 2028 under BAU (Table 4.2). The increase of non-forest vegetation as forest decrease between 2016 and 2028 under BAU can possibly be caused by deforestation and forest degradation since other land cover classes remain more or less the same coverage (Belay et al., 2015). The implementation of REDD+ scenario will increase forest coverage by 5% between 2016 and 2028 while cropland will be kept the same as 2016.

	2004		2009		2016		Modelled 2028 under different scenarios			
							BAU		REDD+	
Land cover class	**Km2**	**%**	**Km2**	**%**	**Km2**	**%**	**Km2**	**%**	**Km2**	**%**
Cropland	51412.92	8.64	72056.95	12.10	102399.76	17.20	101204.66	17.00	102399.76	17.20
Forest	93452.28	15.70	61181.29	10.28	39407.87	6.62	10638.24	1.79	41219.25	6.92
Fon-forest	377287.00	63.37	393227.60	66.05	429274.40	72.10	471086.70	79.13	439156.70	73.76
Built-up	584.55	0.10	501.03	0.08	748.58	0.13	639.45	0.11	737.28	0.12
Bare-areas	60546.51	10.17	56379.11	9.47	11645.94	1.96	40.11	0.01	40.98	0.01
Bater	12069.39	2.03	12006.70	2.02	11876.16	1.99	11743.48	1.97	11798.68	1.98

Table 4.2: Area of coverage for each land cover class under all the scenarios

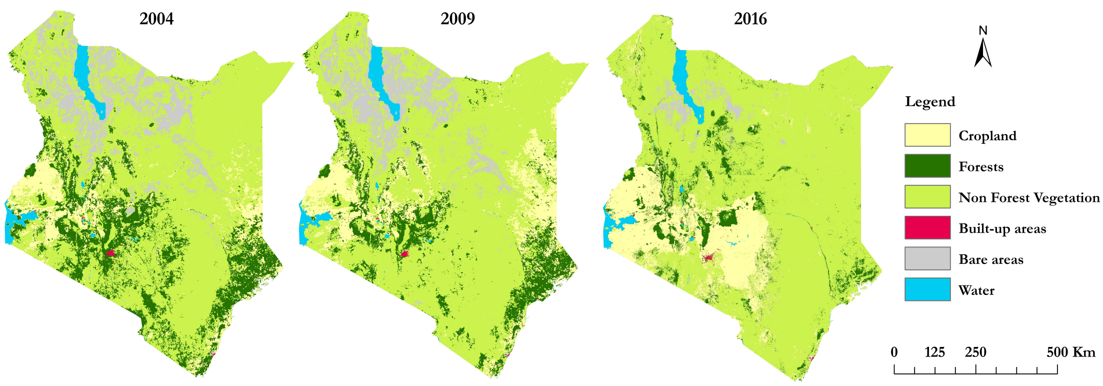

Figure 4.3: Historic land cover maps

In the analysis for the LCC for both historical and future scenarios (Figure 4.3 and Figure 4.4), it can be observed; there was a significant conversion from forest to cropland in the western and central parts of Kenya than other regions. The North Eastern and Coast regions showed a significant conversion of forest to non-forest vegetation.

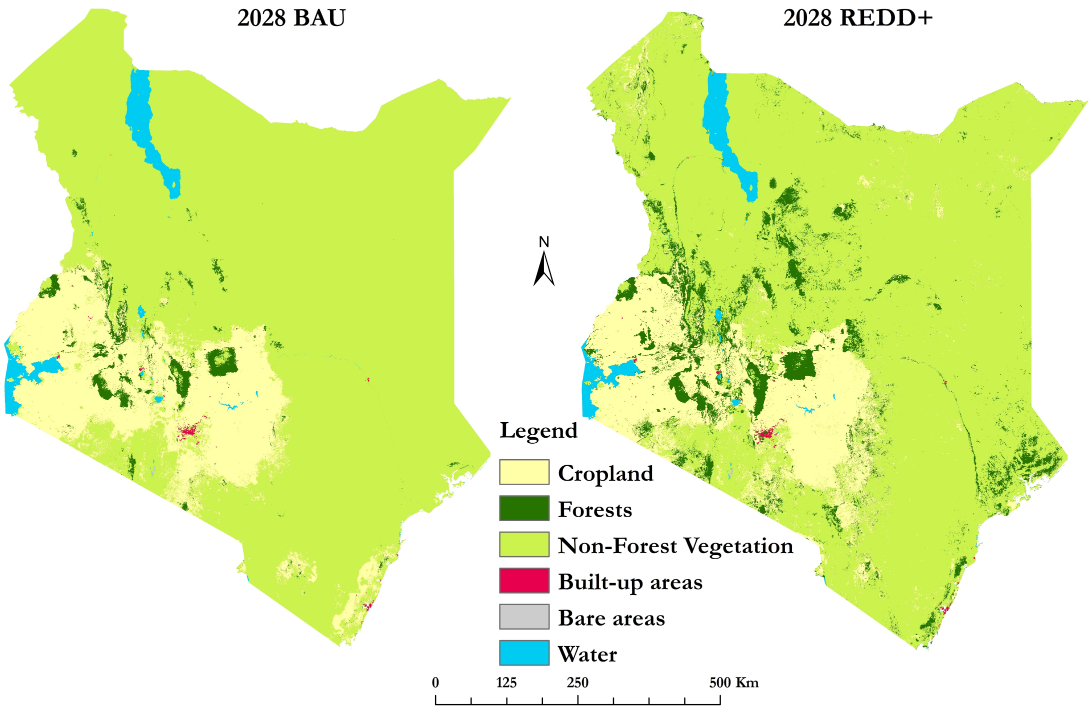

Figure 4.4: Land cover maps for 2028 under BAU and REDD+ scenarios

4.3 Impact of land cover change on carbon stock

From Table 4.3 and Figure 4.5, the total amount of carbon sequestration was valued at npv of 14.6 billion, 1 billion and 6 billon US$ for 2004-2009, 2009-2016 and 2016-2028 respectively. There was an 8% decrease in carbon stock between 2004 and 2016 and is expected to decrease even further by 8% by 2028 under BAU which was consistent with the observed decrease in the forest coverage despite the increase in cropland and non-forest vegetation covers. The decrease highlights the importance of forest in carbon sequestration. However, If REDD+ scenario was to be implemented between 2016 and 2028, emissions worth costs 1 billion US$ could be avoided.

Year	Total carbon stock estimate (Mtons)	Carbon stock change (%)	Net present value (npv) (US$)
2004	4,150	-	-
2009	3,830	-7.6	-14,600,000,000
2016	3,800	-0.7	-1,000,000,000
2028 BAU	3,490	-8.4	-6,000,000,000
2028 REDD	3,860	1.6	1,000,000,000

Table 4.3: Carbon stock change between 2004 and 2028 under all scenarios

Figure 4.6 provides a national level outlook of the spatiotemporal carbon stock changes as a result of land cover changes. The red regions indicate the conversion from forest whereas the deep green represents the contrary. It can be observed; between 2004-2016

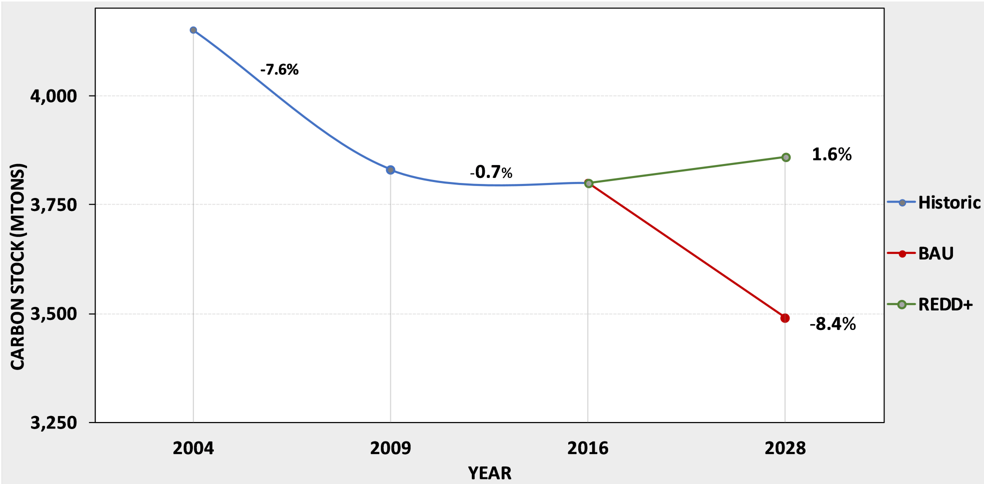

Figure 4.5: Carbon stock change between 2014 and 2028

central region and coastal, the highest coverage area of forests under BAU scenario. In addition, regional level analysis is important for understanding unique ecosystem

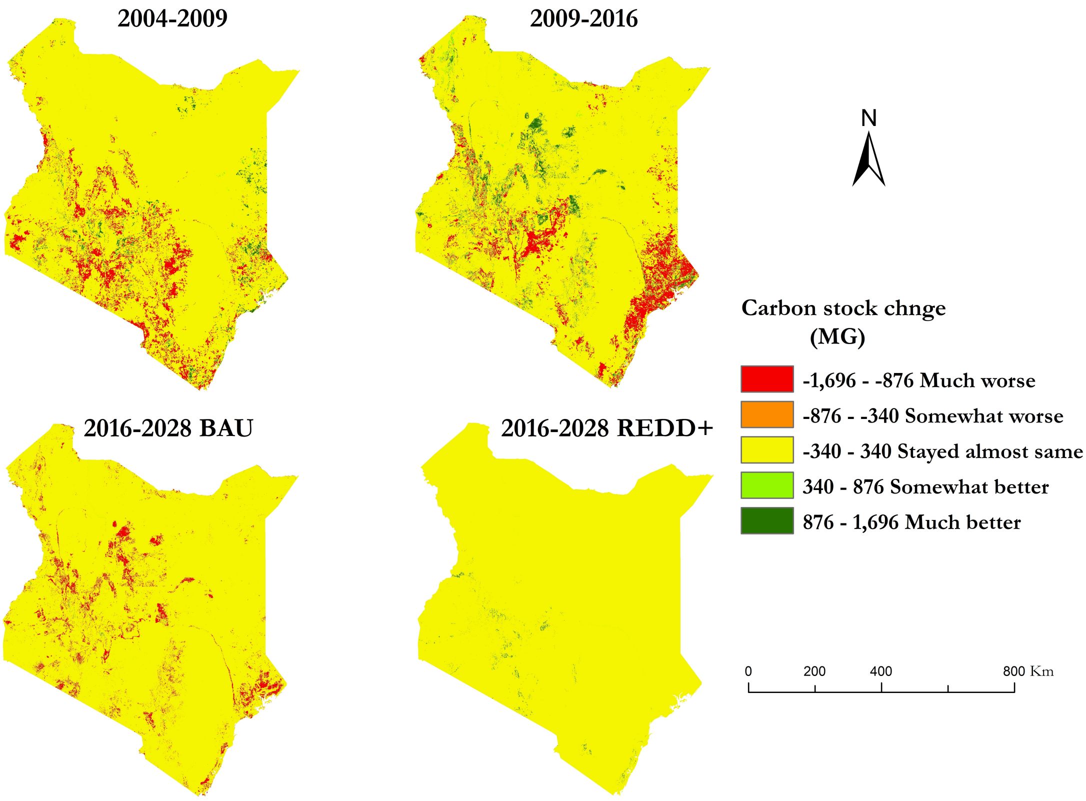

Figure 4.6: Carbon stock changes between 2004 and 2028

within the regions. Table 4.4 and Figure 4.7 illustrate carbon stock variations within the different regions for each different epoch under study. Nairobi region exhibited the greatest losses between 2004-2009 and 2016-2028 by about 38% and 15% respectively. The REDD+ scenario once again indicates a positive move to restore the carbon sequestration, Nairobi been the most benefiting region with a gain rate of 18%. Figure 4.7 shows the spatial-temporal variations across the study area regions.

Region	Area (Km2)	Rate of change (%)			
		2004-2009	2009-2016	2016-2028 BAU	2016-2028 REDD+
Western	8,363	1.3	5.5	-5.9	0.1
Nyanza	15,934	-18	20.3	-3.9	0.6
Rift Valley	184407	-14.1	9.4	-10.1	2.6
Nairobi	740	-37.7	1.6	-15.2	18.2
Central	13356	-3.1	-14.1	-5.7	7.5
Eastern	159036	-10.0	23.5	-10	1.9
North Eastern	130080	7.0	-21.3	-5.3	0.07
Coast	83310	-9.5	-21.7	-6.8	0.7

Table 4.4: Carbon stock change rate by region between 2004 and 2028

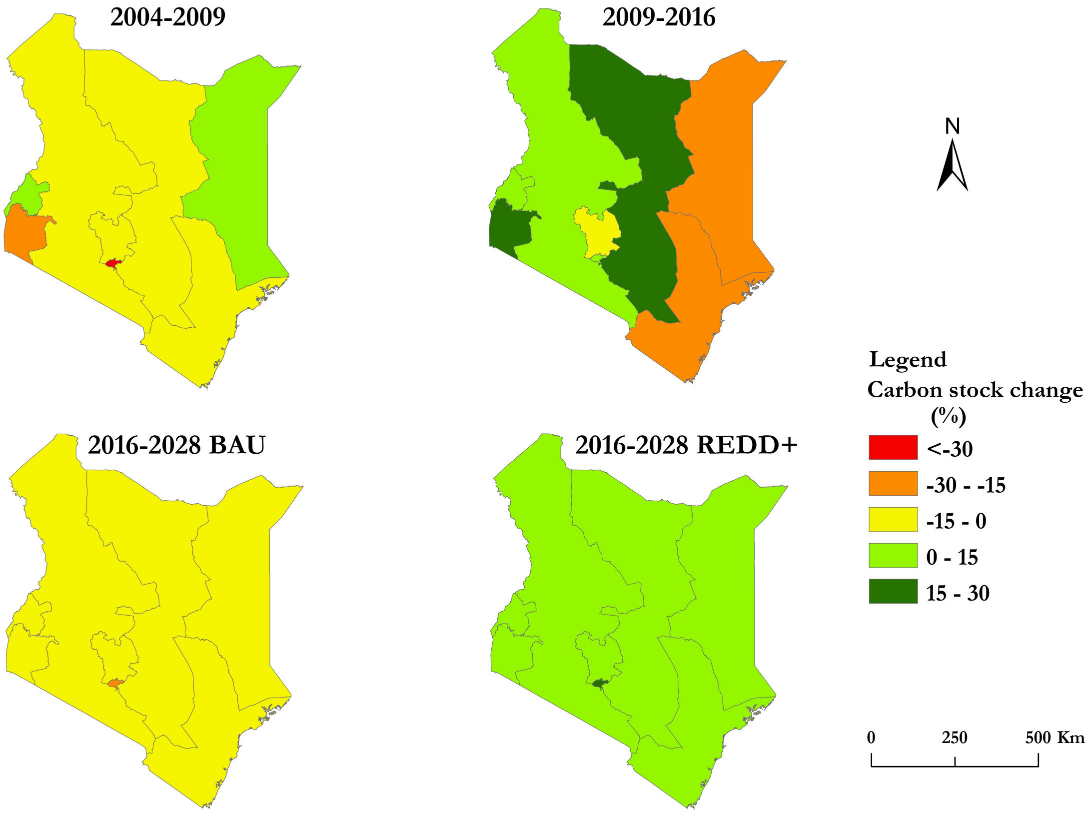

Figure 4.7: Carbon stock change rate (%) by region between 2004 and 2028

5. DISCUSSION

Understanding the spatial-temporal dynamics of carbon stock trends as a result of land cover change is core to the sustainable utilization of natural resources. Using open data and software, land cover changes for Kenya over the next decade were successfully modelled under different alternative scenarios. The modelled land cover maps were keys for estimation and valuation of carbon stock in InVEST carbon model. These estimations and analysis provides insight carbon stock trajectories for benchmarking of land use policies particularly in Kenya where there are heterogeneous and diverse ecosystems but with no or little information at national or regional scale.

From the results, at a national level, the carbon stock is continuously declining. The decline was mainly as a result of the decrease in forest coverage. As observed by (Belay et al., 2015), the decline of forest indicated deforestation and degradation within the country since there was an increase of grassland and shrubland. These findings call for urgent reinforcement of REDD+ supporting policies such as sustainable intensification (Mbow, Reisinger, Canadell, & O'Brien, 2017) to enable restoration of forests.

In the regional level, Central and Coastal regions showed a continuous decline in carbon stock in comparison with other regions. It is important to conserve the forests in these regions because they hold sensitivity ecosystems that support even life beyond the regions. For instance, Coastal region has the mangrove forests and distinct ecosystem that support in cultural value, erosion regulation, provision of fishing grounds etc. (Huxham et al., 2015) while the Central region has Aberdare ranges which is one of the regional water catchment towers and basically provides Nairobi metropolitan with water and other supporting services (Ark, Group, et al., 2011). In addition, Nairobi region exhibited unique results from other regions with the greatest loss. This highlights the importance of dedicating separate studies on the region since it is an urban set-up which supports a huge population with extensive heterogeneity. It is also recommended to conduct extensive region-wise analysis in order to account contributions of specific landscapes such as Mount Elgon which have previously shown declining protected forests (Petursson, Vedeld, & Sassen, 2013).

Estimation of the carbon stock as a result of land cover change requires consistent climatic season, accurate land cover data and reliable modelling with less uncertainty.

However, these aspects are sometimes impeded by inherent model inadequacies or lack of data (Hamel & Bryant, 2017), especially at a national scale. In this study, it was assumed that the climatic season did not significantly affect the data from which land cover maps were derived. The land cover maps for 2004 and 2009 from GlobCover project were produced at an overall accuracy of 73% with a spatial resolution of 300m and 2016 land cover from Africa land cover project produced at an accuracy of 65% at a spatial resolution of 20m. The accuracies are a bit way lower than the recommended accuracy of 85% (Niquisse et al., 2017). The differences in spatial resolution is also a challenge where resampling introduce some errors. In addition, aggregation of land cover classes in GlobCover introduces uncertainty that causes during misclassification during LCC modelling. However, the results can be improved by using more accurate and update data, generation of more robust algorithms (Mendoza-Ponce, Corona-Núñez, Kraxner, Leduc, & Patrizio, 2018) and incorporating extensive land cover change drivers (Kindu, Schneider, Teketay, & Knoke, 2016).

Finally, this study analyzed carbon stock as a result of land cover change under different alternative scenarios thereby providing opportunities and constraints for reinforcing of spatial planning development. However, it is recommended to understand further other factors such as the socio-economic impact on the carbon stock since the implementation of REDD+ scenarios involve multiple processes, governance actors and geographic scales (Corbera & Schroeder, 2011). Additionally, this study focussed on carbon stock supply, hence there is a need to analyze from the aspect of demand versus supply to enable more beneficial resource management.

6. CONCLUSION

This study contributes by presenting useful insights for understanding the impact of land cover change on carbon stock reservoirs as an indicator of climate regulation ecosystem services in a spatial and temporal explicit manner. By modelling land cover under 'what if' scenarios provide the basis to anchor land use planning decision making and policy development in order to systematical implement balance between economic and sustainable developments. More importantly, this study provide a methodology for transforming the United Nation REDD+ policy to a map for spatially setting out the reduction emission of carbon dioxide to the atmosphere through forest conservation.

The findings under business as usual, showed a decline of carbon stocks mainly as because of the decline in forest coverage. This serves as an early warning for further deteriorating climate regulating ecosystem service. Therefore, it points out to the responsible authority to conserve and restore forests.

Lastly, the use of open data and software such as InVEST carbon model provide an affordable but reliable means for carbon stock estimation and valuation particular on large geographic scales thus contribute to the maintenance of balances and tradeoffs in ecosystem services functioning.

Bibliography

Ark, R., Group, K. F. W. et al. (2011). Environmental, social and economic assessment of the fencing of the aberdare conservation area-executive summary.

Barrow, E., & Mogaka, H. (2007). Kenya's drylands–wastelands or an undervalued national economic resource'. *IUCN–The World Conservation Union. Nairobi.*

Belay, K. T., Van Rompaey, A., Poesen, J., Van Bruyssel, S., Deckers, J., & Amare, K. (2015). Spatial analysis of land cover changes in eastern tigray (ethiopia) from 1965 to 2007: Are there signs of a forest transition? *Land Degradation & Development*, *26*(7), 680–689.

Bontemps, S., Defourny, P., Bogaert, E. V., Arino, O., Kalogirou, V., & Perez, J. R. (2011). Globcover 2009-products description and validation report.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32.

Charif, O., Omrani, H., Abdallah, F., & Pijanowski, B. (2017). A multi-label cellular automata model for land change simulation. *Transactions in GIS*, *21*(6), 1298–1320.

Chollet, F. [Francois]. (2017). *Deep learning with python*. Manning Publications Co.

Chollet, F. [François] et al. (2015). Keras.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, *37*(1), 35–46.

Corbera, E., & Schroeder, H. (2011). Governing and implementing redd+. *Environmental science & policy*, *14*(2), 89–99.

ESRI. (2018). Arcgis [internet]. Environmental Systems Research Institute. Retrieved from https://www.esri.com

Gibbs, H. K., Brown, S., Niles, J. O., & Foley, J. A. (2007). Monitoring and estimating tropical forest carbon stocks: Making redd a reality. *Environmental Research Letters*, *2*(4), 045023.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, *27*(4), 294–300.

Gómez-Baggethun, E., & Barton, D. N. (2013). Classifying and valuing ecosystem services for urban planning. *Ecological Economics*, *86*, 235–245.

Government of Kenya. (2010). *National climate change response strategy.*

Hamel, P., & Bryant, B. P. (2017). Uncertainty assessment in ecosystem services analyses: Seven challenges and practical responses. *Ecosystem Services*, *24*, 1–15.

Houghton, R., & Nassikas, A. A. (2017). Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochemical Cycles*, *31*(3), 456–472.

Huxham, M., Emerton, L., Kairo, J., Munyi, F., Abdirizak, H., Muriuki, T., ... Briers, R. A. (2015). Applying climate compatible development and economic valuation to coastal management: A case study of kenya's mangrove forests. *Journal of environmental management*, *157*, 168–181.

Kindu, M., Schneider, T., Döllerer, M., Teketay, D., & Knoke, T. (2018). Scenario modelling of land use/land cover changes in munessa-shashemene landscape of the ethiopian highlands. *Science of the Total Environment*, *622*, 534–546.

Kindu, M., Schneider, T., Teketay, D., & Knoke, T. (2016). Changes of ecosystem service values in response to land use/land cover dynamics in munessa–shashemene landscape of the ethiopian highlands. *Science of The Total Environment*, *547*, 137–147.

Leh, M. D., Matlock, M. D., Cummings, E. C., & Nalley, L. L. (2013). Quantifying and mapping multiple ecosystem services change in west africa. *Agriculture, ecosystems & environment*, *165*, 6–18.

Li, X., & Yeh, A. G.-O. (2002). Neural-network-based cellular automata for simulating multiple land use changes using gis. *International Journal of Geographical Information Science*, *16*(4), 323–343.

Mbow, H.-O. P., Reisinger, A., Canadell, J., & O'Brien, P. (2017). Special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (sr2).

Mendoza-Ponce, A., Corona-Núñez, R., Kraxner, F., Leduc, S., & Patrizio, P. (2018). Identifying effects of land use cover changes and climate change on terrestrial ecosystems and carbon stocks in mexico. *Global Environmental Change*, *53*, 12–23.

Mustafa, A., Cools, M., Saadi, I., & Teller, J. (2017). Coupling agent-based, cellular automata and logistic regression into a hybrid urban expansion model (huem). *Land Use Policy*, *69*, 529–540.

Nakicenovic, N., Alcamo, J., Grubler, A., Riahi, K., Roehrl, R., Rogner, H.-H., & Victor, N. (2000). *Special report on emissions scenarios (sres), a special report of working group iii of the intergovernmental panel on climate change*. Cambridge University Press.

NCPD. (2018). *State of kenya population june 2018*. National Council of Population, Government of Kenya. Nairobi, Kenya. Retrieved from http://www.ncpd.go.ke/wp-content/uploads/2018/12/State-of-Kenya-Population-June-2018-2.pdf

NEMA. (2015). *Government of kenya second national communication to the united nations framework convention on climate change*. National Environment Management Authority, Government of Kenya. Nairobi, Kenya. Retrieved from https://unfccc.int/resource/docs/natc/kennc2es.pdf

Niquisse, S., Cabral, P., Rodrigues, Â., & Augusto, G. (2017). Ecosystem services and biodiversity trends in mozambique as a consequence of land cover change. *International Journal of Biodiversity Science, Ecosystem Services & Management*, *13*(1), 297–311.

Oliphant, T. E. (2006). *A guide to numpy*. Trelgol Publishing USA.

Parry, M., Parry, M. L., Canziani, O., Palutikof, J., Van der Linden, P., & Hanson, C. (2007). *Climate change 2007-impacts, adaptation and vulnerability: Working group ii contribution to the fourth assessment report of the ipcc*. Cambridge University Press.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *Journal of machine learning research*, *12*(Oct), 2825–2830.

Pellikka, P., Heikinheimo, V., Hietanen, J., Schäfer, E., Siljander, M., & Heiskanen, J. (2018). Impact of land cover change on aboveground carbon stocks in afromontane landscape in kenya. *Applied Geography*, *94*, 178–189.

Petursson, J. G., Vedeld, P., & Sassen, M. (2013). An institutional analysis of deforestation processes in protected areas: The case of the transboundary mt. elgon, uganda and kenya. *Forest Policy and Economics*, *26*, 22–33.

Pfeifer, M., Platts, P. J., gess, N. D., Swetnam, R., Willcock, S., Lewis, S., & Marchant, R. (2013). Land use change and carbon fluxes in east africa quantified using earth observation data and field measurements. *Environmental conservation*, *40*(3), 241–252.

Pontius Jr, R. G., & Batchu, K. (2003). Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in india. *Transactions in GIS*, *7*(4), 467–484.

Python Software Foundation. (2016). *Python language reference version 3.6*. Retrieved from https://www.python.org/

QGIS Development Team. (2018). *Qgis geographic information system*. Open Source Geospatial Foundation Project. Retrieved from https://qgis.org

R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from https://www.R-project.org

Rahmati, O., Pourghasemi, H. R., & Melesse, A. M. (2016). Application of gis-based data driven random forest and maximum entropy models for groundwater potential mapping: A case study at mehran region, iran. *Catena*, *137*, 360–372.

UN-REDD. (2019). Un redd programme. Accessed: 2019-02-13. United Nation Development Programme (UNDP), United Nation Environment Programme (UNEP), Food, and Agriculture Organization of United Nation(FAO). Retrieved from https://www.un-redd.org

Shafizadeh-Moghadam, H., Asghari, A., Tayyebi, A., & Taleai, M. (2017). Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth. *Computers, Environment and Urban Systems*, *64*, 297–308.

Sharp, R., Tallis, H., Ricketts, T., Guerry, A., Wood, S., Chaplin-Kramer, R., ... Olwero, N., et al. (2014). Invest user's guide. *The Natural Capital Project: Stanford, CA, USA*.

Stocker, T. (2014). *Climate change 2013: The physical science basis: Working group i contribution to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press.

Tayyebi, A., & Pijanowski, B. C. (2014). Modeling multiple land use changes using ann, cart and mars: Comparing tradeoffs in goodness of fit and explanatory power of data mining tools. *International Journal of Applied Earth Observation and Geoinformation*, *28*, 102–116.

Tol, R. S. (2009). The economic effects of climate change. *Journal of economic perspectives*, *23*(2), 29–51.

Trexler, M. C., & Kosloff, L. H. (1998). The 1997 kyoto protocol: What does it mean for project-based climate change mitigation? *Mitigation and Adaptation Strategies for Global Change*, *3*(1), 1–58.

Vashum, K. T., & Jayakumar, S. (2012). Methods to estimate above-ground biomass and carbon stock in natural forests-a review. *J. Ecosyst. Ecogr*, *2*(4), 1–7.

Wangai, P. W., Burkhard, B., & Müller, F. (2016). A review of studies on ecosystem services in africa. *International journal of sustainable built environment*, *5*(2), 225–245.

Warmerdam, F. (2008). The geospatial data abstraction library. In *Open source approaches in spatial data handling* (pp. 87–104). Springer.

Willcock, S., Phillips, O. L., Platts, P. J., Swetnam, R. D., Balmford, A., Burgess, N. D., ... Doody, K., et al. (2016). Land cover change and carbon emissions over 100 years in an a frican biodiversity hotspot. *Global Change Biology*, *22*(8), 2787–2800.

World Bank. (2019). Kenya[internet]. Accessed: 2019-01-27. Retrieved from http://data.worldbank.org/country/Kenya

Zhang, F., Zhan, J., Zhang, Q., Yao, L., & Liu, W. (2017). Impacts of land use/cover change on terrestrial carbon stocks in uganda. *Physics and Chemistry of the Earth, Parts A/B/C*, *101*, 195–203.

A. APPENDICES

A.1 Land cover change modelling python script

Refer to model script in https://github.com/nicnri/Land-cover-change-modelling

A.2 Exploratory data analysis

Symbol	Variable name	Description
Numbers 1-6	cropland, forests, non-forest vegetation, built-up areas, bare areas and water respectively	Land cover classes
SL, dC, dT, dRD	Slope, cities distance,major towns distance, and road distance respectively	Biophysical variable; slope, computed distance to cities, major towns, and roads.
d1, d2, d3,d4, d4 & d6	Neighborhood proximities for cropland, forests, non-forest vegetation, built-up areas, bare areas and water respectively	Computed proximity (raster distances) for each land cover class

Table A.1: Description of modelling independent variables

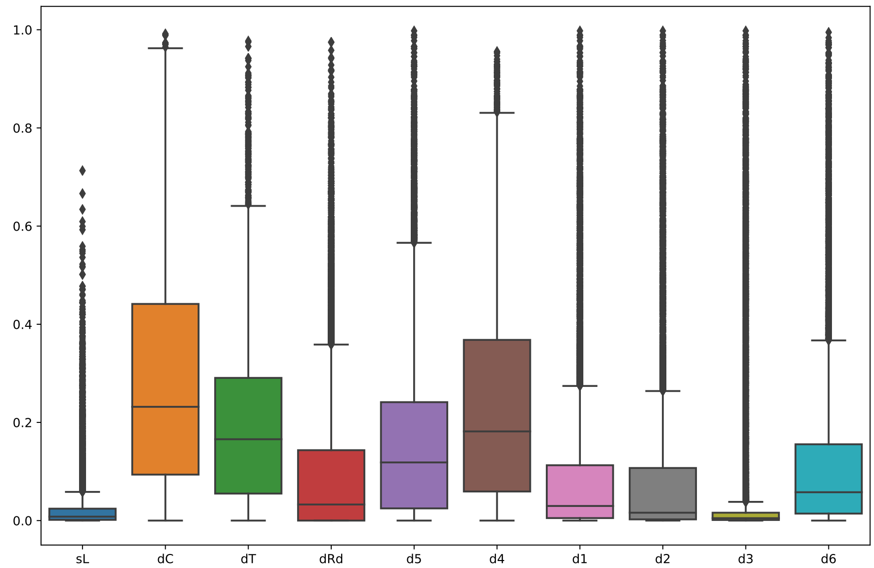

Figure A.1: Boxplots of land cover proximity distances and biophysical variables

From Figure A.1, it can be observed that slope and non-forest vegetation neighborhood proximity have the lowest spread in contrast to the distance to cities and built-up areas neighborhood proximity raster (the variable names refer Table A.1). Additionally, the ranges for distance to cities and built-up areas neighborhood proximity raster are comparably similar as expected since most urban developments occur within the proximity of cities. From the boxplot, it can be deduced that all the variables have some data that are out of the third quartile which show the existence of outliers but since the represent actual measurements, they were not removed from the processing data.

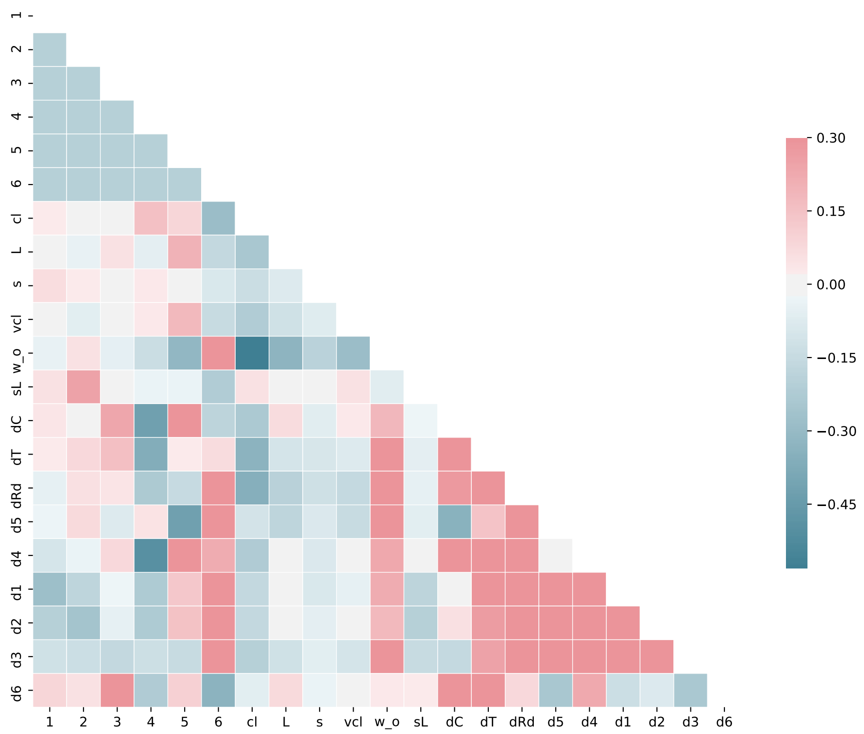

Figure A.2: Land cover modelling independent variables correlation heat map

Figure A.2 shows the land cover classes are relatively negative correlated, additionally, distance to cities, towns, roads are to some extent exhibiting a positive correlation to land cover neighborhood proximities but the correlation coefficient is not sufficient to conclude the correlation between the variables. Therefore,the variables are independent.

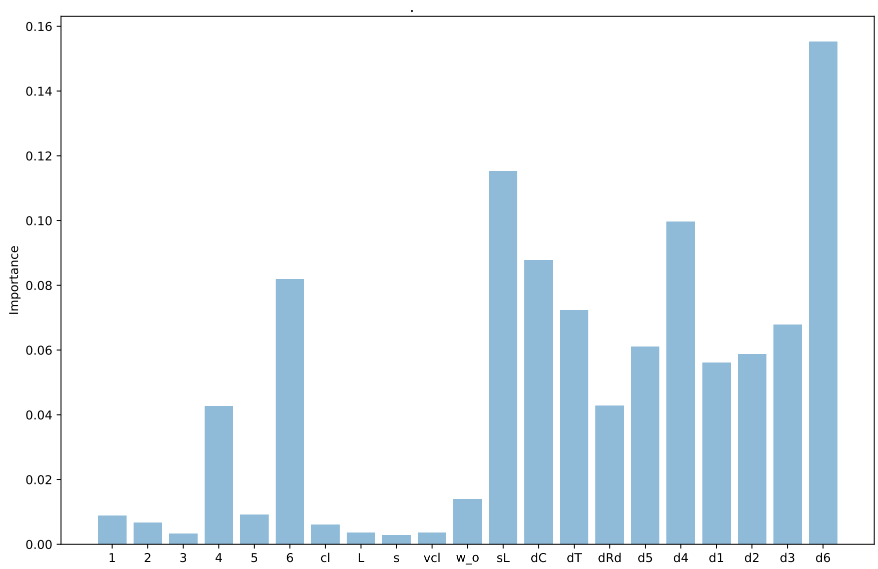

Figure A.3: RF measure of level of importance of each independent variable in prediction

Figure A.3 shows the contribution of each variable to the discrimination features during prediction. The level of importance is given on the scale between 0 and 1. It can be deduced; the biophysical and neighborhood variables are relatively more important, with slope and water neighborhood proximity distance contributing the most than other variables by about 12% and 15.5% respectively. Moreover, non-forest vegetation and sandy soil type contribute the least, showing non-forest vegetation cover can occur indiscriminately of the physical attributes. It can also be deduced that sandy soil type supports almost all land cover types.

Masters
Program
in **Geospatial**
Technologies

IMPACT OF LAND COVER CHANGES ON CARBON STOCK TRENDS IN KENYA USING FREE OPEN DATA

Nicodemus Ontweka Nyamari

Dissertation submitted in partial fulfilment of the requirements for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS

UNIVERSITAT JAUME•I

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER

2019

IMPACT OF LAND COVER CHANGES ON CARBON STOCK
TRENDS IN KENYA USING FREE OPEN DATA

Nicodemus Ontweka Nyamari

Masters
Program
in **Geospatial**
Technologies

Supported by:
Education and Culture
ERASMUS MUNDUS