ESTIMATION OF SEDIMENT YIELD IN THE UPPER TANA CATCHMENT AREA USING SWAT MODEL

BY

MWANGI JILLIAN WANGARI

DECLARATION

I declare that this project is my own work and has not been submitted by anybody else to the best of my knowledge.

Sign:…………………………………   Date:……………………………………

Name: MWANGI JILLIAN WANGARI

ACKNOWLEDGEMENT

I want to thank the Almighty God for helping me get this far, my supervisor Dr. Nathan Agutu for his continuous support and great help while I did my project and my parents for their continuous support and words of encouragement to push forward.

Abstract

Sediment yield refers to the amount of sediment discharged at an outlet or deposited into a reservoir downstream over a period of time. Globally reservoirs lose their storage capacity due to sedimentation at the rate of 2% annually, the reservoirs at the upper Tana catchment area lose their storage capacity to sedimentation at the rate of 1% annually. This study aimed at simulating sediment yield in the upper Tana catchment using the Soil and Water assessment tool (SWAT) through the following; i.) to derive the hydrological response units, ii.) to estimate the sediment yield by running the swat model for a period of 28 years and iii.) to calibrate and validate the sediment yield output using SWAT_CUP SUFI2 algorithm. Swat model was first run to derive the hydrological response units which are the response units of the catchment and have an effect on sediment and flow using the digital elevation model to first delineate the watershed and the land use/landcover and soil maps to derive the HRUs, the weather data was also input and the SWAT model run to get estimated amount of sediment yield every year. The output was later calibrated and validated using observed data from 1991- 2000. Results from this study showed that; i) high sediment yield data was in the year 2019 as 383276.2 million cm^3 and low sediment yield was from the year 1988 as 600.967million cm^3, ii) the year with the highest derived hydrological response units was 2018 and the lowest year was 1988, iii) the calibrated and validated NSE objective function value was 0.82 indicating the model was a good tool to estimate sediment yield and Iv) the average annual sediment yield yearly was found to be 0.041tonnes. The information could be used to come up with mitigation soil conservation measures such as afforestation in areas highly affected by soil erosion in order to curb sedimentation.

Table of contents

Acronyms and abbreviations

HRUs- hydrological response units.

SWAT- Soil and Water Assessment Tool.

LULC- land use and landcover.

NSE- Nash Sutcliffe

WEPP- Water Erosion Prediction Project

SHETRAN- Systeme Hydrologique Europeen

ANSWERS- Areal Nonpoint Source Watershed Environment Response Simulation

AGNPS- Agricultural Non-Point Source Pollutant model

AnnAGNPS- Annualized Agricultural Non-Point Source Pollutant model

CHAPTER 0NE

INTRODUCTION

Sediment yield refers to the amount of sediment discharged at an outlet over a period of time. It is also referred to as the amount which would enter into a reservoir located at the downstream limit of the basin. (Fan,1988) All the soil particles eroded from a watershed are not transported to the outlet due to their trapping and deposition in the upstream reach.

Estimates of sediment yield are needed for studies of reservoir sedimentation, river morphology and planning of soil and water conservation measures. Past studies revealed that around the globe, storage reservoirs lose their capacity due to sedimentation to the tune of 2% annually (Mahmood, 1987).

Many of the reservoirs in the upper Tana catchment lose their annual capacity at the rate of 0.23m^3. Most of the sediment yield assessment as of present are achieved either by simple regression models or through distributed empirical models. Regression equations often relate sediment yield to basin properties such as drainage area, topography, climate and vegetation characteristics. Most of the time, decision-makers use the basic Universal Soil Loss Equation (USLE); or its variant, which is the Revised Universal Soil Loss Equation (RUSLE) and Modified Universal Soil Loss Equation (MUSLE) (Renard, 1997,1965 and 1978) for sediment yield estimation which produces the best accuracy results. It is reported that these approaches only take into account surface erosion, and are not recommended for soil loss estimation at catchment level. On the other hand, various spatially distributed empirical models have been developed to describe erosion and sediment transport at the basin scale.

These models are often based on sediment transport capacity equation or an estimated Sediment Delivery Ratio to determine sediment transport and sediment yield (Kothyari,2000) Spatially distributed models include CREAMS (Foster, 1981), SEDIMOT (Wilson, 1984), AGNPS (Young, 1987), ANSWERS (Beasley,1980) and WEPP (Nearing, 1989). The Texas agricultural experiment station used Soil and Water Assessment Tool (SWAT) (Arnold, 1995) a physically based distributed hydrological model, to predict the effect of management decisions on water and sediment yield with reasonable accuracy on a large ungauged river basin on a daily time step.

Surface runoff into the rivers that originate from the Aberdares and Mt. Kenya slopes contributes to most of the sedimentation. The high production rates of sediment can be linked with the fact that these rivers pass through intensively cultivated slopes of the highlands associated with the Aberdares and Mt. Kenya. Lack of adequate ground cover on the steep slopes that are often cultivated without carrying out effective soil conservation measures results in increased surface runoff and soil loss with subsequent sedimentation problems of the reservoirs along the Tana River.

Upper Tana Nairobi Water Fund report indicate that the loss of riparian vegetation cover, landslides, quarrying close to rivers, livestock trucks and road run-off and farmlands are major point source contributors of sedimentation in the Tana river area. Although there has been increased surface runoff and soil loss from the upper catchment area, the increased rates of sedimentation of the multipurpose reservoirs constructed along the Tana River has been a subject of debate. (Brown, 1996) estimated that the average input of sediment to the Masinga dam for the first seven years of operation (1981 – 1988) were approximately 0.6-0.9. It was suggested that there is no danger of reservoir lifespan being significantly reduced from that given by the design proposals of 520 years.

However, other reports provide conflicting information. Jacobs (2007), reported that the Masinga Reservoir has a high trap efficiency which ranges between 75% and 98%, which in turn results in an average loss of 23 million m^3 per year of storage. Based on these estimates, complete sedimentation of the Masinga Reservoir is likely to occur in 65 years unless some mitigation measures are implemented. More recently, however, sedimentation studies carried out by (Antonaropoulos and Karavokyris, 2010) estimated rates of 8.03 tonnes per year of sediment inflow into Masinga while Kamburu receives 1.11 tonnes per year from the Thiba catchment. The estimated sediment loads into Masinga are higher than the design sediment load of 3.0 million m^3 per year estimated in 1981 which is about 1% per annum reservoir reduction (Mutua, 2005).

Based on the extent of the upstream catchment area of Masinga reservoir of 7,335km^2 and Kamburu reservoir of 2,185km^2 the average catchment sediment yield for Masinga was calculated as 10.9 tonnes every year while the average catchment sediment yield was calculated as 5.08 tonnes every year for Kamburu reservoir. The sediment yield estimates from the above study are well within the range of values offered by a variety of sources that have studied the sedimentation problem in the upper Tana catchment in the past. Maingi, (1984 &1991) estimated sediment loads of 8.46 and 7.47 tonnes per year respectively. Comparative values of 8.03 tonnes per year were estimated by Droogers, (2011). From these an estimated figure of 8 tonnes per year appears plausible. HR Wallingford (1972), estimate for Masinga which assumes a sediment density of 1.2 tonnes per m^3, amounted to 13.18 tonnes every year, while estimate for Kamburu before the construction of Masinga was 4.28 tonnes every year (Wooldridge, 1984). Most sediment enter reservoirs suspended in water, though below than 50% of the sediment yield can be contributed to uncontrolled runoff along roads, loosened earthworks and culvert discharge into unprotected lands (WRMA, 2010). From this study the amount of sediment yield is estimated using the SWAT model and this will help in coming up with various mitigation processes that will help curb the amount of sediment load that inflows into the reservoirs.

PROBLEM STATEMENT

The upper Tana river catchment is home to around 5.2 million people and its growing population is causing it to be under diverse changes caused by various anthropogenic activities (Hunink & Droogers, 2011).

The catchment is made up of a number of reservoirs such as the seven forks' dams which include Masinga, Kamburu, Kiambere and Gitaru. It is located across parts of six counties namely; Nyeri, Meru, Embu, Tharaka Nthi and Kirinyaga.

The catchment good soils and availability of water from the various rivers such as Mathioya and Matonga have resulted in an increase in anthropogenic activities thus increased amount of sediment yield transported into the reservoirs through the process of soil erosion (Muchena, Hunink, Droogers, Njuguna, Onduru, Muthuri, Macharia & Maingi, 2012). This is a big problem especially in the Masinga dam which is a major dam that produces hydroelectric power of about 40MW, it also acts as a reservoir to store water for various human activities such as fishing and farming and it also regulates the amount of water getting into the other dams.

Masinga dam being the biggest amongst all of them with a capacity of 1.56 million m^3 has been greatly affected by sedimentation hence reduced storage capacity (Bunyasi et al., 2013). If sedimentation continues the existent usefulness of the reservoir will not be there and hence a great effect in both controlling of floods and production of hydroelectric power, therefore management processes to mitigate sedimentation have been put across to help reduce the sediment yield.

A model that is able to keep track on the amount of sediment yield transported into the reservoirs will help manage it since processes such as dredging will not be useful as they are costly. SWAT model which is a soil and water assessment tool in ArcGIS will be useful in determining the amount of sediment yield in the dam and also to predict future sediment yield with high accuracy, unlike the bathymetric surveys that have been conducted previously and are costly and time consuming.

Research identification

The main aim of this research is to use SWAT model for the estimation of sediment yield in the upper Tana catchment area mainly focusing on Kindaruma dam. This is achievable through the following objectives:

GENERAL OBJECTIVE

To estimate the amount of sediment yield in the upper Tana catchment area using swat model for a period of 28 years from 1991 to 2019.

SPECIFIC OBJECTIVES

- To derive hydrological response units.
- To run the model and obtain simulated sediment yield data.

- To calibrate and validate the findings using the SWAT CUP model tool with SUFI2 (sequential uncertainity fitting version 2) program algorithm.

CHAPTER TW0

LITERATURE REVIEW

In recent years there has been a need to understand the problem of sedimentation in reservoirs as it poses a great danger to dam functionality and existence globally. Many studies have been done to help understand and provide a solution to the problem through soil conservation mitigation processes such as afforestation to help curb amount of sediment deposited downstream into reservoirs (Gathagu et al., 2018), (Hunink & Droogers, 2011), (Brown et al., 1996) and (Arnold et al., 1994). Sedimentation is the rate at which sediments transported from a highland area are stored on the bed of reservoirs or waterbodies downstream (fan et al., 1988). Sediment yield is the amount of sediment discharged at an outlet downstream over a period of time (Vanmaercke et al. 2011). In many countries, sustainable use of reservoirs is seriously threatened by their declining storage capacity due to the problem of sedimentation. This process reduces and eventually eliminates the storage capacity required for hydropower generation, flood control and water supply (Hunink et al., 2013).

A common way to assess sediment production and transport is through a mathematical modelling approach. Mathematical models are useful land management decision support tools. For example, sediment yield models are used to determine soil redistribution due to environmental changes. There are many theoretical approaches to sediment modelling (Bunyasi et al., 2013). All these studies point out that, during last decades, development of new models tended to produce conceptual and physically based distributed models. Some examples include; WEPPs, AGNPS, ANSWERS, SHETRAN and SWAT models (Bussi et al., 2014). This is because the sediment cycle is characterized by high complexity and non-linearity. These are phenomena that simple empirical models such as, RUSLE and USLE cannot describe easily. Moreover, the spatial variability of erosion and deposition processes is fundamental for catchment management decision support (Bussi et al., 2014).

A study done by Chandramohan, (2015) evaluated three soil erosion models on a small watershed in India, Kerala basin. The models tested were Modified Universal Soil Loss Equation (MUSLE), Unit Sediment graph (USG) and Water Erosion Prediction Project (WEPP). They incorporated observed rainfall- runoff sediment yield events of three watersheds with varying land use types, topograpphy and drainage density. The results were compared with the observed sediment yield values. It was seen that the USG predicted the rate of sediment yield better than the other two models. Even though WEPP is a physically distributed model, the large and detailed data requirement, which is impractical in studies of this scale, affected its prediction accuracy and MUSLE produced the worst errors as it showed wide predicted value variations as compared to observed sediment data in the basin.

Study carried out in the Orazan watershed, Iran tested the efficiency of WEPP model to predict runoff and sediment yield at catchment scale in a semi-arid area. Continuous simulations were conducted between 1996 and 2005. Comparison between predicted and observed indicated that WEPP underestimated sediment volumes by 23% and overestimated runoff volumes by 27%. Results showed that sediment yield and Runoff outputs are fairly well predicted as the model overestimates and underestimates scenerios (Ahmadi & Feiznia, 2011).

Figueiredo & Bathurst (2007), modelled runoff and sediment yield using SHETRAN in the semi-arid region North-east of Brazil. The study was carried out using data observed at various basin scales. Model parameters were evaluated using field data and techniques based on soil texture. The evaluated parameters were sufficient to represent the characteristics of the region. The results achieved at every basin scale and grid size, and at different time resolutions (daily, monthly and annual) showed that observed runoff and sediment yields were simulated with physically meaningful results. Land-use change effects on simulated runoff and sediment yields were considerable and scale differences were insignificant. For the homogeneous areas, the parameters tested for the plot scale could be used to simulate larger areas with different grid sizes, opening a possibility of simulating relevant processes for ungauged basins, an important requirement in vast region areas.

Another study carried out in the semi-arid area of Cobres basin in the southern Portugal automatically calibrated SHETRAN model using Modified Shuffled Complex Evolution (MSCE). The study incorporated soil data, land use/ landcover data. The Nash Sutcliffe (NSE) values provided satisfactory results for all storm events at a range of 0.69-0.87 in both calibration and validation for sediment yield. The results were satisfactory not only for basin outlet but also for internal gauging stations (Zhang et al., 2013).

Study carried out by Chia & Mbajiorgu, (2018) estimated runoff and sediment yield from the Upper Ebonyi River Watershed, Enugu State, Nigeria using the KINEROS2 and AGNPS models. The models were compared using the observed measured data with that estimated by the models. A sample of 12 rainfall events of different depths and durations were selected between June and July 2013 for the watershed modelling. Six events were used for calibration of both models, while six were used for their validation. Both models showed good capability for simulating peak runoff and sediment yield from an agricultural watershed, the KINEROS2 output data were closer to the measured data than those of AGNPS.

Another study carried out in the Three-Gorge region of the Yangtze River of China assessed the characteristics of soil erosion, sediment and sediment delivery of a watershed using Annualized Agricultural Non-Point Source Pollutant model (AnnAGNPS). The datasets used for the study were soil data, land use/landcover maps, weather data such as temperature and rainfall data and a 25meter Digital Elevation Model (DEM). The model was calibrated using observed monthly runoff from 1988 to 1999 and annual average sediment from 20002002. The NSE results were satisfactory with a range of 0.78-0.93. Post validation simulation of results showed that approximately 48% of the watershed was under soil loss tolerance. However, 8% of the watershed had high soil erosion. Sloping areas and low coverage areas are the main source of soil loss in the watershed (Hua et al., 2014).

A Multi scale approach between soil erosion and sediment yield was done in the Upper Tana catchment area, Masinga dam. The new sediment yield estimate was found to be lower than some previous estimates, but it was sensitive to climatic variation. The study showed that areas of soil erosion and sediment were found to be along tracks and paths (Brown et al., 1996).

Physiographical baseline study carried out by the Water Resource Authority (WRA) in the upper Tana catchment area to estimate sediment load using bathymetric surveys with stations placed at different areas in the catchment was done from 1984 until the year 2011. The model was then validated using available observed data. The results showed high amounts of sediment load in the upper Tana catchment area especially Masinga dam. The study proved that highest sedimentation occurs in Masinga dam (Hunink & Droogers, 2011).

A study on Impact of soil erosion on sedimentation in the Ethiopian highlands assessed landscape sensitivity to erosion using Revised Universal Soil Loss Equation adjusted for sediment delivery ratio. The model predicted high soil loss rates at steep slopes and shoulder positions as well as along gullies (Tamene et al., 2017).

A study estimated sediment yield in Masinga dam using both descriptive and inferential statistics for data analysis. The quantitative data was analyzed using correlation and trend analysis using SPSS and Microsoft Excel soft-wares. The study showed that Masinga dam had lost 13,56% of its design storage capacity as of 2011 from year of construction. This informed a great need for management practices to be done (Bunyasi et al., 2013).

Soil and Water Assessment Tool (SWAT) is a physically distributed model that provides comprehensive good performance results for small and large watersheds to river basin scale. The model incorporates the use of Digital Elevation Models, land use/ landcover maps, soil data and weather data such as rainfall and temperature to estimate sediment yield. It is a widely used model because it provides more accurate results either daily, monthly or yearly. It is a simple model with good user interface and it is an opensource tool (Fao,2018).

The studies below used SWAT model as a tool for sediment yield estimation, hydrology and best soil management practices (BMPs).

A modelling approach to evaluate the impacts of structural conservation measures on water and sediment yield from Thika Chania catchment in Central Kenya was done using the SWAT model. The SWAT model output was calibrated and validated for stream flow and sediment yield at a gauging station in the catchment. The calibrated model was run to create a base scenario for the simulation of structural conservation methods, example; terraces and grassed waterways. The Modeled simulation results indicated that terraces and grassed waterways would significantly impact water and sediment yield at the catchment outlet. Terraces were found to provide the greatest reduction in sediment yield, by 81% from the baseline scenario, while grassed waterways reduced sediment yield by 54%. Terraces indicated a reduction in surface runoff by 30% from the base annual average value of 202 mm. This was attributed to the increased infiltration that was indicated by increase in base flow by 8%. However, grassed waterways did not indicate any significant reduction in water yield. The results of this study show that structural conservation measures could reduce sediment yield from cultivated areas by more than 50% at the sub catchment level. Results also indicated that the effectiveness of structural conservation measures can be increased by implementing more than one method. Structural conservation measures studied in the study were found to have a positive impact in controlling water and sediment yield in the catchment (Gathagu et al., 2018).

A study on hydrology in Simly Dam watershed located in Saon River basin at the north-east of Islamabad was modelled, using the Soil and Water Assessment Tool (SWAT). It simulated the stream flow, established the water balance and estimated the monthly volume inflow to Simly Dam in order to help the managers to plan and handle this important reservoir. The model showed good results of 0.75 in inflow rate into the dam. These results revealed that SWAT model can be used efficiently to support water management policies (Ghoraba, 2015).

A Study carried out in Nigeria, shiroro dam estimated sediment yield using SWAT model, results were later calibrated and validated. The model produced good results for NSE at 0.71 and P- factor value of 0.78, this showed that the model was a good tool to be used for estimation of sediment yield at catchment level and soil management practices could be put forward to save the sustainability of the dam(Daramola et al., 2019).

A study carried out In the Owabi catchment in Ghana assessed the hydro climatic variability due to anthropogenic activities using SWAT. Calibration results of objective functions NSE, PBIAS and R^2 were 0.61, 0.67 and 0.76 respectively. The model proved efficient in determining the catchment hydrology parameters and has potential to be used for further modelling of water quality and pollution to aid in effective water management (Osei et al., 2019).

Study carried out at the Burnaphur basin in the upper Tapi catchment assessed sediment yield at catchment level using SWAT. The model was calibrated and validated using 12 years of observed data for sediment and runoff, the NSE results were 0.91 and 0.99 respectively, which proved the model was a good tool. This study proved that sediment yield should be managed at catchment level so as to curb sedimentation problem in reservoirs (Chandra et al., 2014).

A study carried out in the Sommerville area in Texas assessed sediment yield using SWAT model. The model results were calibrated and validated providing an NSE value of 0.82. this proved that the model was a good tool to estimate sediment yield at catchment level (Sohoulande Djebou, 2018).

A study carried out in the Ganga river which is affected by sedimentation due to anthropogenic activities and climate change in India used SWAT model to estimate the sediment yield and streamflow. The study was assessed the best calibration algorithm in SWAT_CUP using the simulated results. These algorithms include; SUFI2, GLUE, MCMC and parasol, the algorithms were based on their values of NSE, R_factor and P_factor. SUFI2 algorithm was the best performer amongst them as only one calibration interval was required and the values of objective functions provided better results. Parasol was the worst performer amongst the algorithms because it didn't account for all uncertainties and required the highest number of computational parameters for calibration (Shivhare et al., 2018).

A study carried at the upper Danube basin in Germany assessed suspended sediment concentration due to hillslope erosion. The study used SWAT model to obtain suspended sediment concentration in reservoirs and the results were calibrated using hydrological response units unlike reach results. It was found that a lot of hillslope erosion occurred in the HRUs and this lead to increased sediment concentration deposited downstream (Vigiak et al., 2015).

Soil erosion from agricultural field has posed as a great problem of sedimentation in the downstream reservoirs. A study done in the kalaya basin of Morocco used SWAT to calibrate the Best Management Practices (BMP) of the area. SWAT showed good results for the catchment and the BMP values were high showing that a lot of soil erosion was occurring at the agricultural fields and this led to high deposition of sediment in the downstream reservoirs (Briak et al., 2019).

SWAT has proved to be a useful GIS tool to predict and estimate sediment yield with use of satellite imagery showing land use/landcover changes, better land management practices can be implemented in areas that are widely affected by poor management practices and that are leading to high sediment yield being transported (Arnold et al., 1995).

Study done in India, Vaigai reservoir evaluated reservoir loss through sedimentation by incorporating remote sensing and GIS techniques in the field of Civil Engineering. Satellite Remote Sensing (SRS) method predicted reservoir sedimentation using directly the waterspread area of the reservoir at a particular elevation on the date of pass of the satellite. With known area and difference in water level, the capacity and loss in capacity of the reservoir due to sedimentation were estimated in ArcGIS software (Rizvana & Merlin, 2016)

A study assessed reservoir sedimentation of the Patratu Reservoir using Satellite Remote Sensing (SRS). The sedimentation assessment was carried out using satellite data and reservoir water level data from 2006 to 2012. Water spread area was analysed from satellite data. The Normalized Difference Water Index (NDWI) was used to delineate open water features and to enhance the presence of water surface in satellite imagery of the Patratu

Reservoir. Water spread area of the reservoir at a particular elevation on the date of the passing of the satellite was used to develop an elevation-area curve. The linear interpolation/extrapolation technique was employed to assess the water spread area of Patratu Reservoir at different elevations. Further, these areas were used to compute the live storage capacity of the reservoir between two elevations by the Prismoidal formula. From the study, it was found that due to sedimentation, the live storage capacity of Patratu Reservoir had reduced from 101.95 to 89.96 hm3, thus showing capacity loss of 11.76% in a span of 44 years (Pandey et al., 2016).

This study uses SWAT model as it is the best to simulate large, complex watersheds with a higher accuracy as compared to the other models. It is the most comprehensive environmental model available and it operates at a daily time step on basin scale with an objective to predict long-term impacts of management practices (Fao, 2020). Widely accepted soil loss estimating tools such as Modified Universal Sediment Loss Equation (MUSLE) that provides high accuracy than the other soil erosion modelling tools, has been integrated into SWAT model and broad parameters incorporated in the model cater for all scenario uncertainities (Sohoulande Djebou, 2018). This makes SWAT a more preferred model for estimation and prediction of sediment yield.

STUDY AREA

The Upper Tana catchment starts 50 km northeast of Nairobi and covers an area of approximately 17,000 squared kilometres. It occupies parts of the following counties; Nyeri, Muranga, Embu, Tharaka-Nithi, Meru, Kitui, Kiambu, Kirinyaga, oMachakos, Nyandarua, Laikipia and Isiolo.

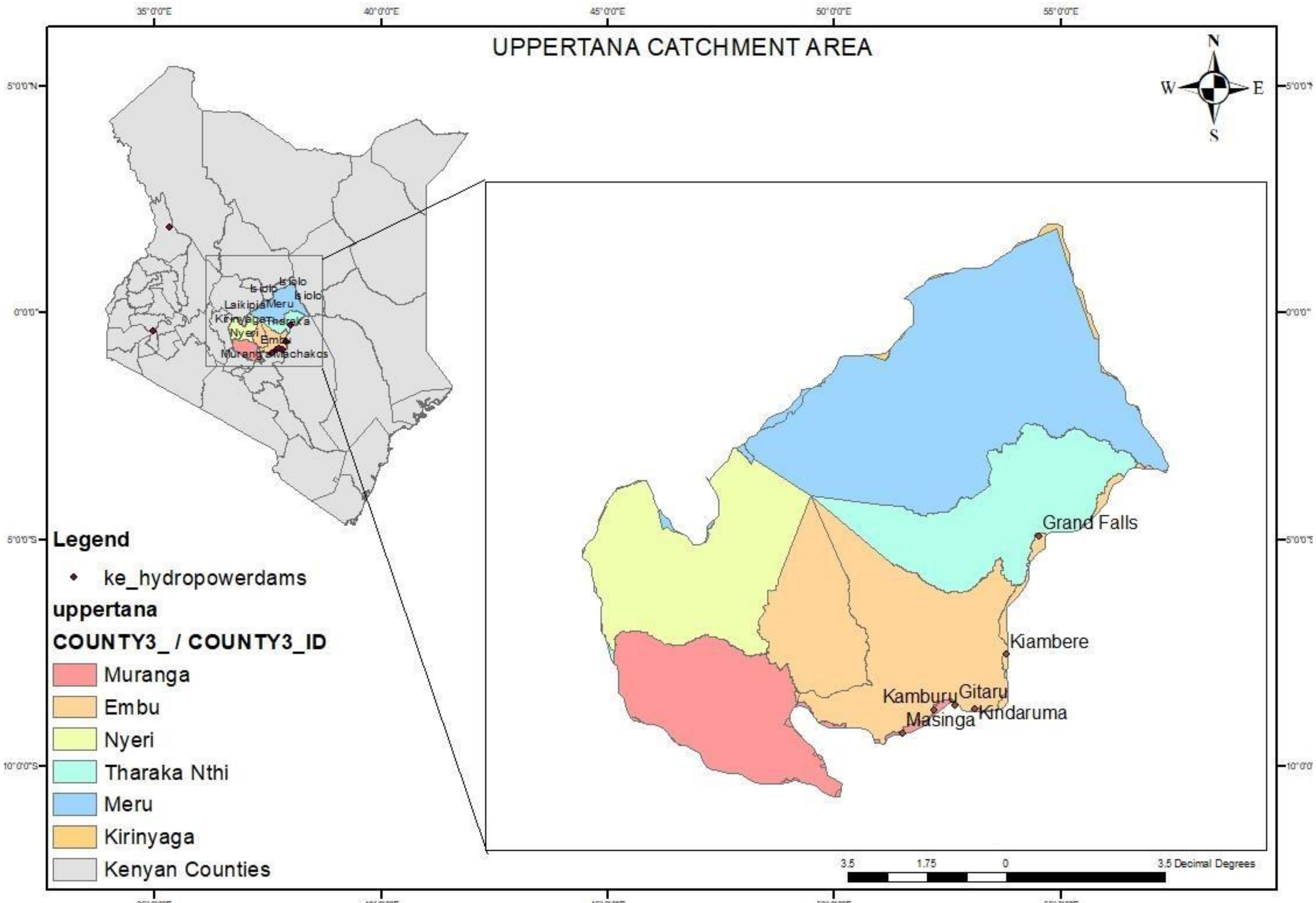

Fig.1; Upper Tana catchment study area.

CHAPTER THREE DATA AND METHODOLOGY

DATA

Table 1 provides a summary of dataset used in this study.

Table 1: Datasets used and their sources.

DATASETS	SOURCE	RESOLUTION	YEAR
RAINFALL	https://data.chc.ucsb.edu/products/CHIRPS-2.0/africa_daily/tifs/	0.05 degrees	1988-2019
TEMPERATURE	https://data.chc.ucsb.edu/products/CHIRTSdaily/v1.0/africa_netcdf_p05/	0.05 degrees	1988-2019
LANDSAT 5 AND 7	https://code.earthengine.google.com/	30METERS	1988-2018
DIGITAL ELEVATION MODEL(DEM)	https://opendata.rcmrd.org/datasets/	30METERS	2018
SOIL DATA	http://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/	90METERS	1971-1981

DATASETS

SWAT model incorporates various data so that it can run, the data include; DEM, land use/landcover maps, soils maps, meteorological data such as rainfall and temperature.

DEM

The digital elevation model of the entire country (Kenya) of a spatial resolution of 30 meters was downloaded from https://opendata.rcmrd.org/datasets/ and uploaded to the Arcmap. The DEM was extracted by mask using the study area shapefile already projected and then saved. The masked DEM of the study area was used in the watershed delineation step.

WEATHER DATA

Weather data downloaded for daily rainfall and temperature from 1988 to 2018 was first clipped to the study area using R and then extracted according to the meteorological weather station name and number, the data was then saved to as csv file. This csv file was used to create rainfall and temperature text files understood by SWAT model, in the swat weather generator preprocessed Microsoft access and excel sheets (WGEN_user). First the rainfall data from the csv file was copy pasted to the PCP excel sheet according to the stations namely; Nyeri, Meru and Embu. The same was done for temperature. Then the preprocessed SWAT access file was opened and the PCP and TMP files in excel form uploaded to it and it generated swat coded text files that could be used later as input files for weather.

SOILS

The soils data downloaded from the FAO database http://www.fao.org/soils-portal/datahub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/ in the form of a shapefile, was then uploaded in Arcmap for further clipping with the study area. It was then projected to the same projection of the study area which was Arc_1960_UTM_37s and then clipped according to the study area, upper Tana. The attribute table was opened so as to see the soil names and codes used to create the look up table for the swat model. Fig.A1; shows the soil map.

L*ANDUSE/LANDCOVER MAPS*

The land use landcover maps were prepared using Google earth engine code editor. The code engine editor performed cloud masking, classification and accuracy assessment. The shapefile of the study area was first uploaded and imported so that the area could be visually seen on the map. The code for image collection was first input, this showed the Landsat collection for use in this case LANDSAT 5 and 7 were chosen for easy change detection later. Later the code for cloud masking and image composite was input followed by the classification and accuracy assessment. For LANDSAT 7 destripping was also performed in the cloud masking part so that the fill gaps present in it were removed. Training points were manually collected according to the six classes namely; forest, vegetation, water, urban, snow and ice and bare land. The code performed accuracy assessment and the classified image in TIFF format was saved and downloaded from the google drive. The tiff files were uploaded to Arcmap so that they could be used to create land use/landcover maps. First each tiff image properties of unique values were added under the symbology tab and this showed the various landcover type number that was used as each number signified a land use name. The lookup table was created from the landcover class and number. This lookup table was used in SWAT to understand the land use/ landcover type. Fig.A2, A3, A4, A5, A6 and A7 show the land use/landcover maps for the years 1988, 1994, 2000, 2006, 2012 and 2018.

METHODOLOGY

The methodology adopted for this study is summarized in fig.2, below.

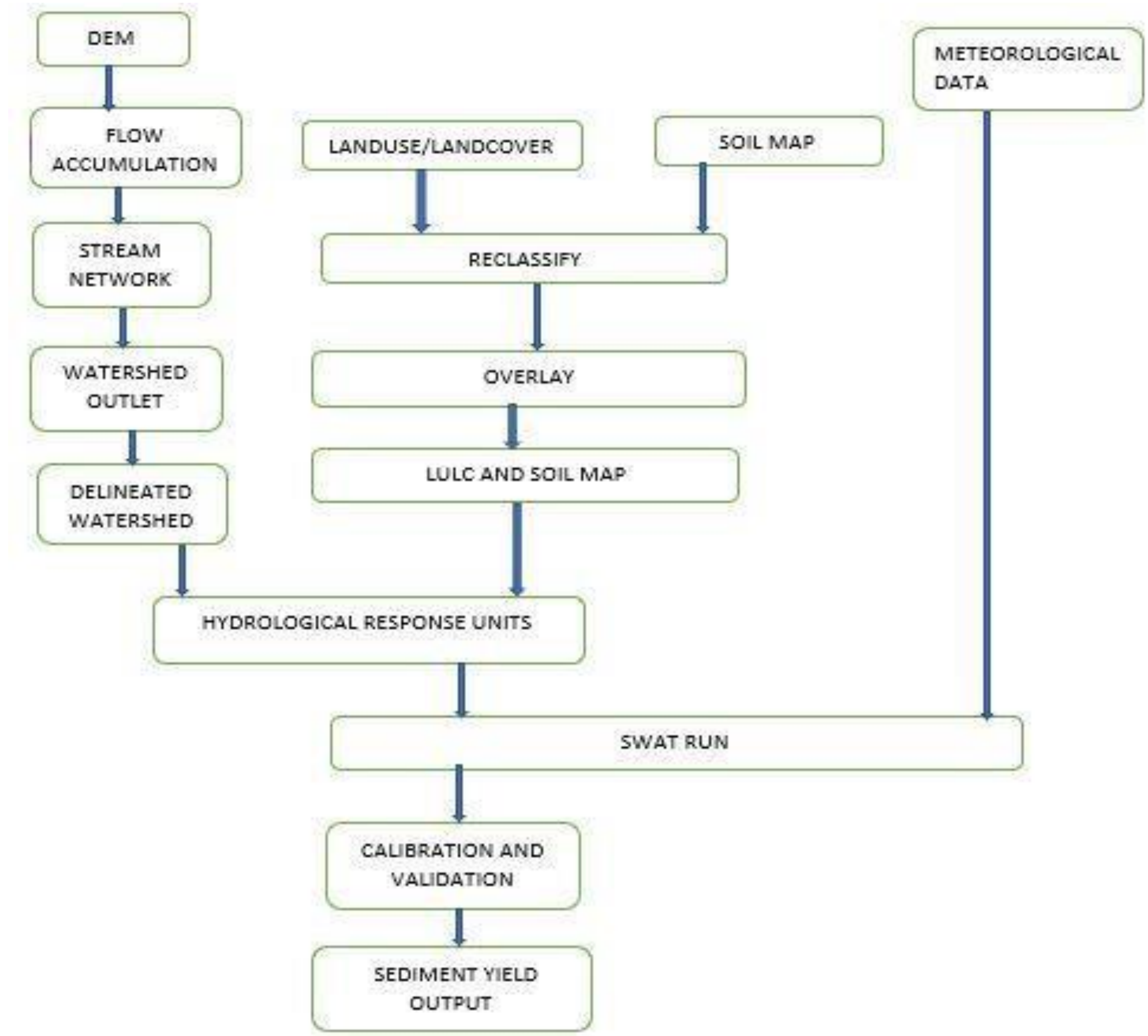

Fig. 2; Methodology flow chart.

Watershed Delineation

In the swat model a new swat project was created and the DEM was input, the next step was to perform flow accumulation and stream network of the entire DEM masked to the study area. The outlets were added manually by selecting the lower points of interest putting in consideration the hydropower dams. Then model selected outlets according to those manually selected, the area had six outlets which were the watershed outlet. The final step was watershed delineation and the generation of a watershed report.

Hydrological response units

This are areas of the same hydrological events and land use, slope and soil properties. They are areas highly responsive in the watershed and the affect flow and sediment. They were derived by inputting the land use maps, soil maps and slope percent limits. The HRUs were created and the full report on the number and their properties included in the report. This report could be used in the final output of sediment yield as it tells us which basin and the units in the area that are likely to undergo changes. Also, in this stage the subbasin number is generated in the report and each subbasin contains certain number of the units.

The watershed had 14 subbasins.

Weather data

The weather data was input in the write input tables where the user weather generator (WGEN_user) is selected and it counts the number of stations in the study area considered by swat. Rainfall and temperature data was input by uploading the pcp and tmp text files initially created. These text files were used by SWAT to generate the rainfall and temperature data from the database. The other weather data such as windspeed was simulated.

Run SWAT

SWAT model is run by first inputting the years of simulation that are being simulated and the number of skip years to be used as warm up period for the model. The simulated number of years was 8 years but the warm up period was 3 years therefore the data that was considered by SWAT model was the one for 5 years.

Calibration and validation

The SWAT CUP was used for the calibration and validation of the model. The simulated data was calibrated from 1991 to 1995 and validated from 1996 to 2000. In swat cup a new project was created by choosing the txtInOut folder in the scenerios output folder of swat, the project was given a name and the program to be used selected as SUFI2. The calibration input parameters were also input in accordance to sediment. The parameters used were SPCON, SPEXP, CH_COV1 and CH_ERODMO this are sediment related parameters that are sensitive to sediment data. The calibration was done using observed sediment yield data already created in an excel sheet. The calibration had one variable that was sediment. RCH (reach) and HRUs (hydrological response units) were used to correctly determine the amount of the sediment yield on the subbasin of interest which was subbasin number 11 and the objective function was set as NSE (Nash Sutcliffe). It was then calibrated for the years 1991- 1999 and the output was validated using the observed data for the years 2000-2006. Sensitivity analysis was also performed on the parameters to compare which responds more efficiently to the data used.

CHAPTER FOUR

RESULTS

Sediment yield estimation was carried out using the Soil and Water Assessment Tool (SWAT). ArcSWAT interfaced with the ESRI ARCGIS 10.4 version was used to obtain simulated sediment yield data of the model and the SWAT_CUP was used for calibration and validation of observed and simulated sediment yield data. The hydrological response units derived are the response units of the catchment and they affect sediment and flow in the catchment. The results are as shown in fig. 3, 4, 5, 6, 7 and 8. The results show that there was a high number 744 in hydrological response units in the year 2018 and the lowest number 682 in 1988. There was the highest increase in hydrological response units in 1988 to 1994 of 43 units. The response units can be seen highly distributed in the middle and far south east regions of the watershed in all years.

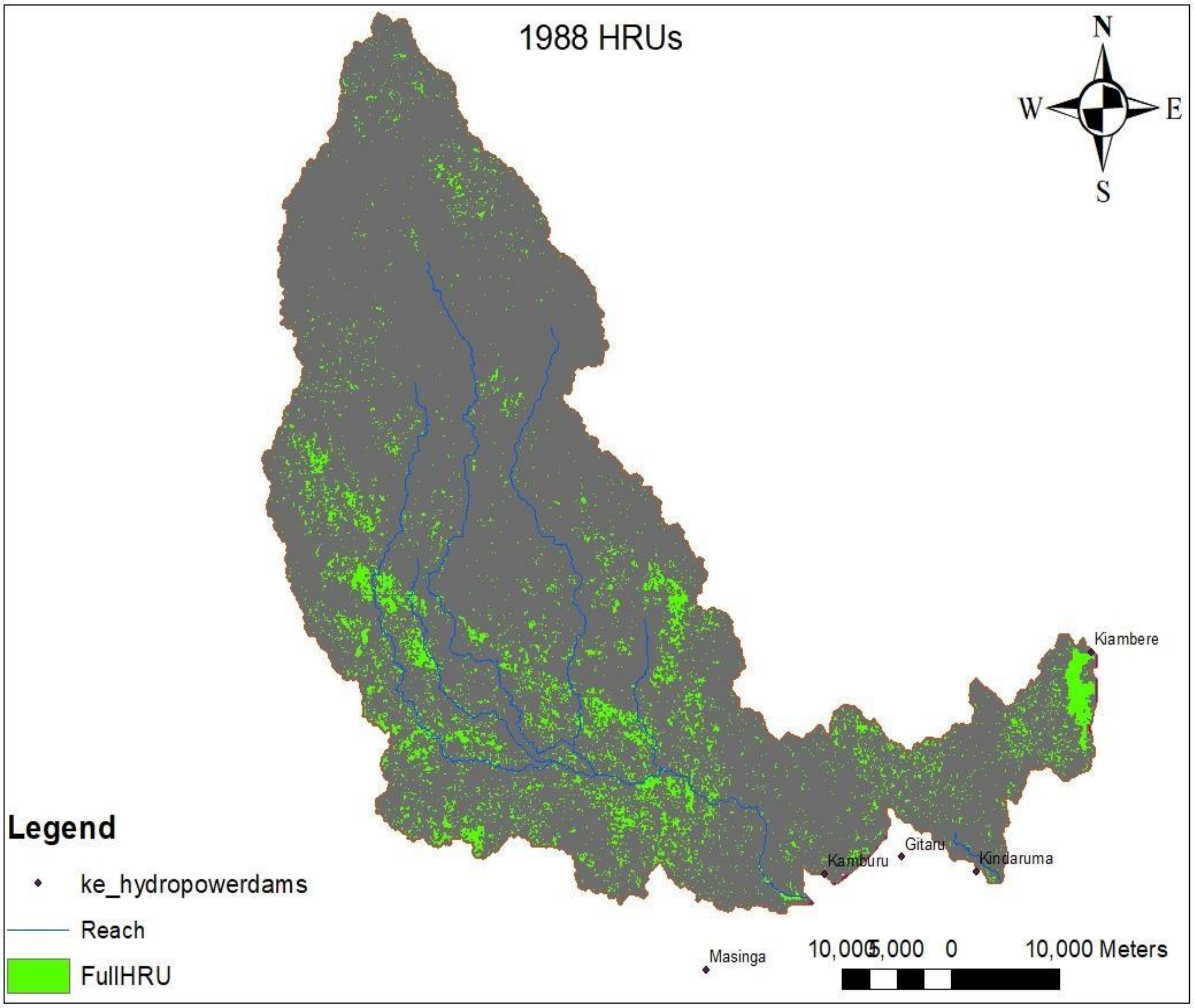

Fig.3; Derived hydrological response units from land use/landcover and soil map of the year 1988. It shows 682 response units of the year.

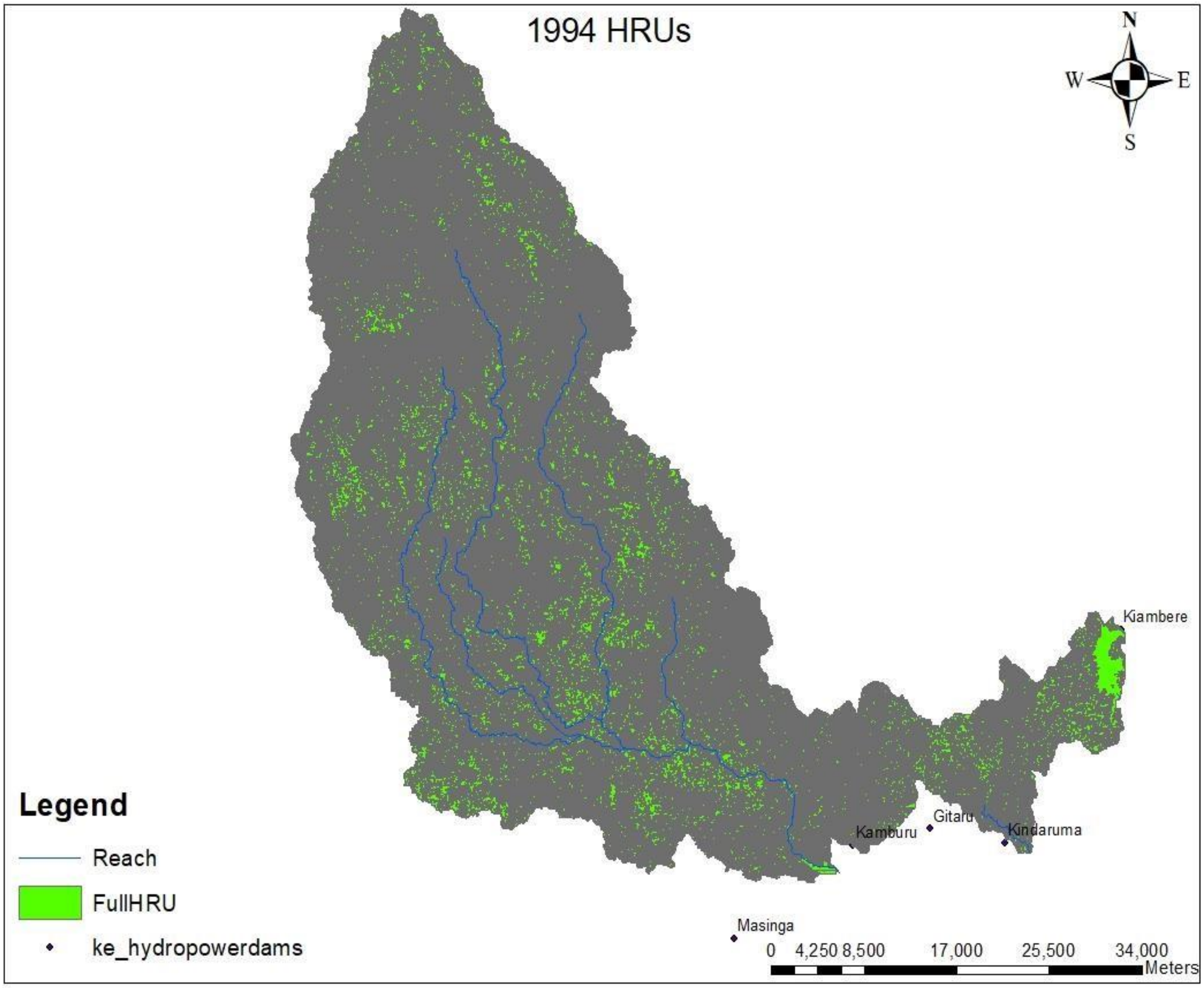

Fig.4; Derived hydrological response units from the land use/landcover and soil map of the year 1994. Its shows 725 response units number of the year.

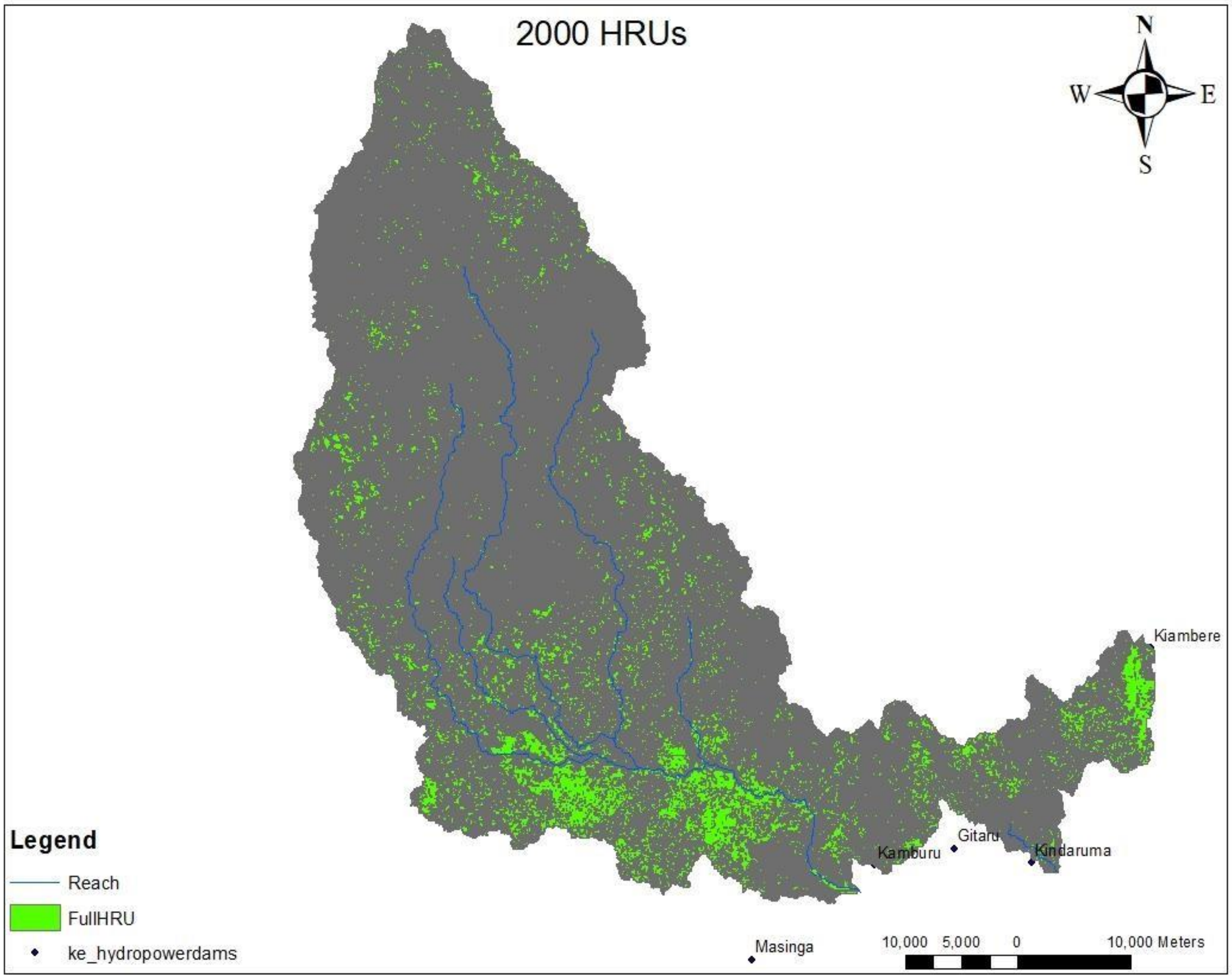

Fig.5; Derived hydrological response units from the land use/ landcover and Soil map of the year 2000. It shows 735 response units of the year.

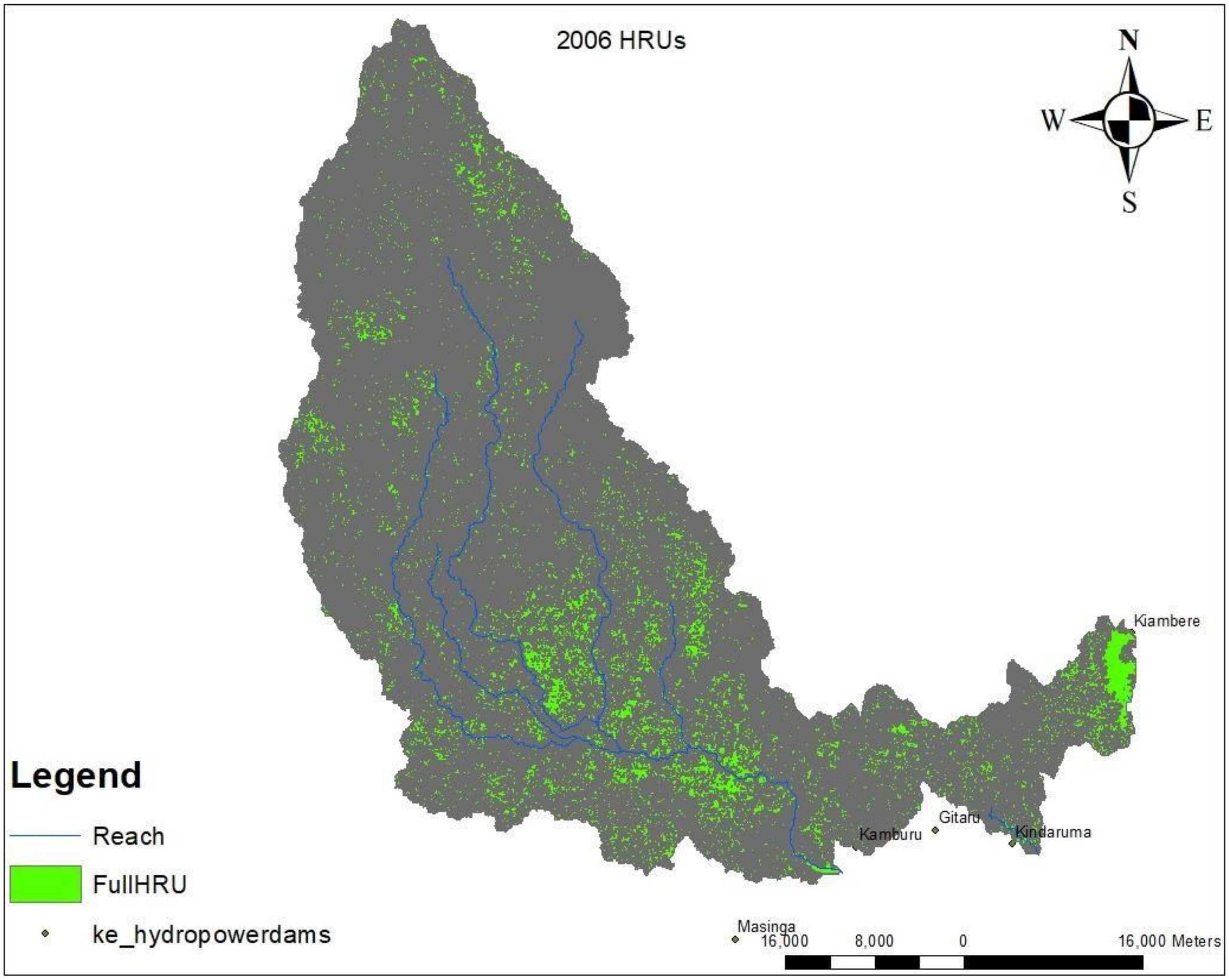

Fig.6; Derived hydrological response units from land use/landcover and Soil map of the year 2006. Its shows 736 response units number of the year.

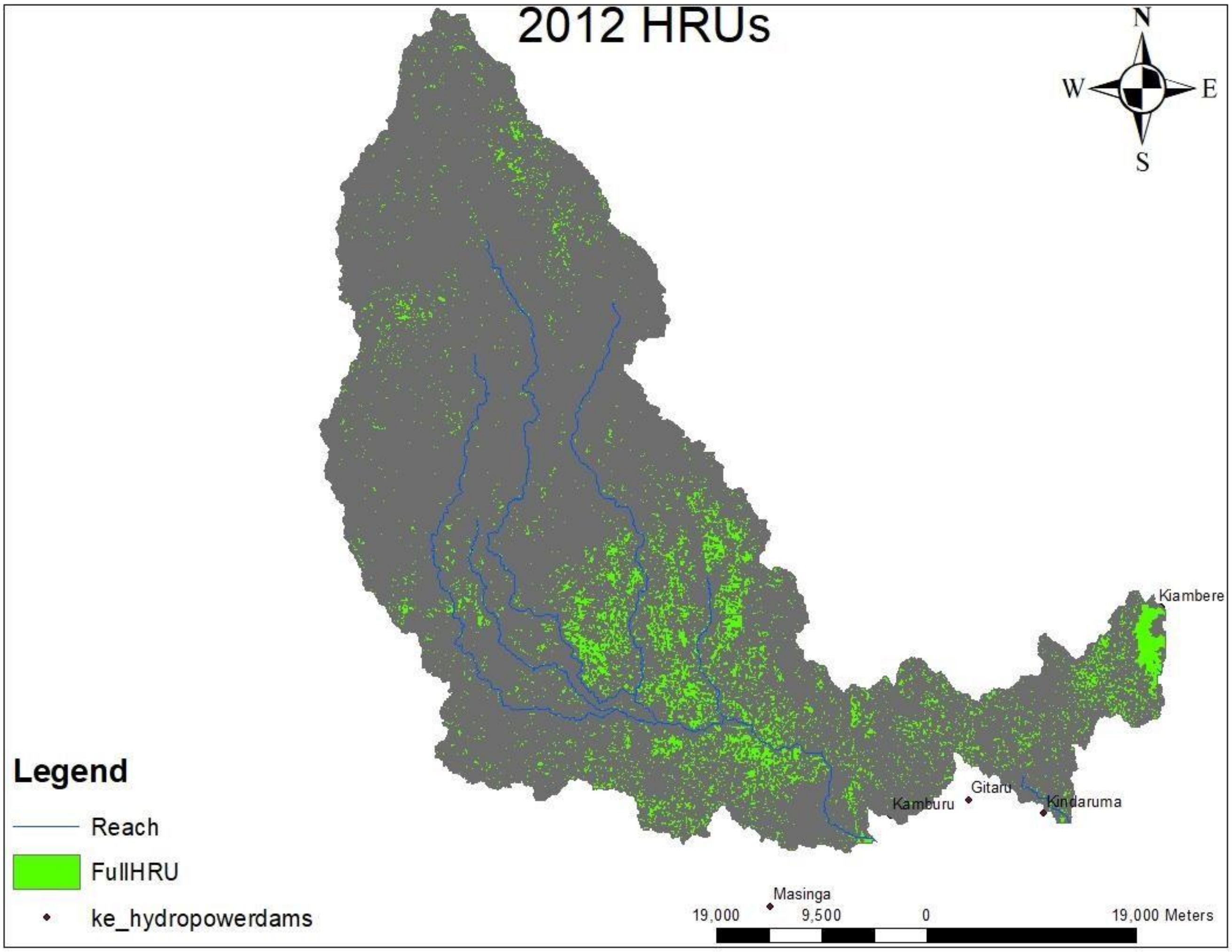

Fig.7; Derived hydrological response units from the land use/landcover and soil map of the year 2012. Its shows 738 response units for the year.

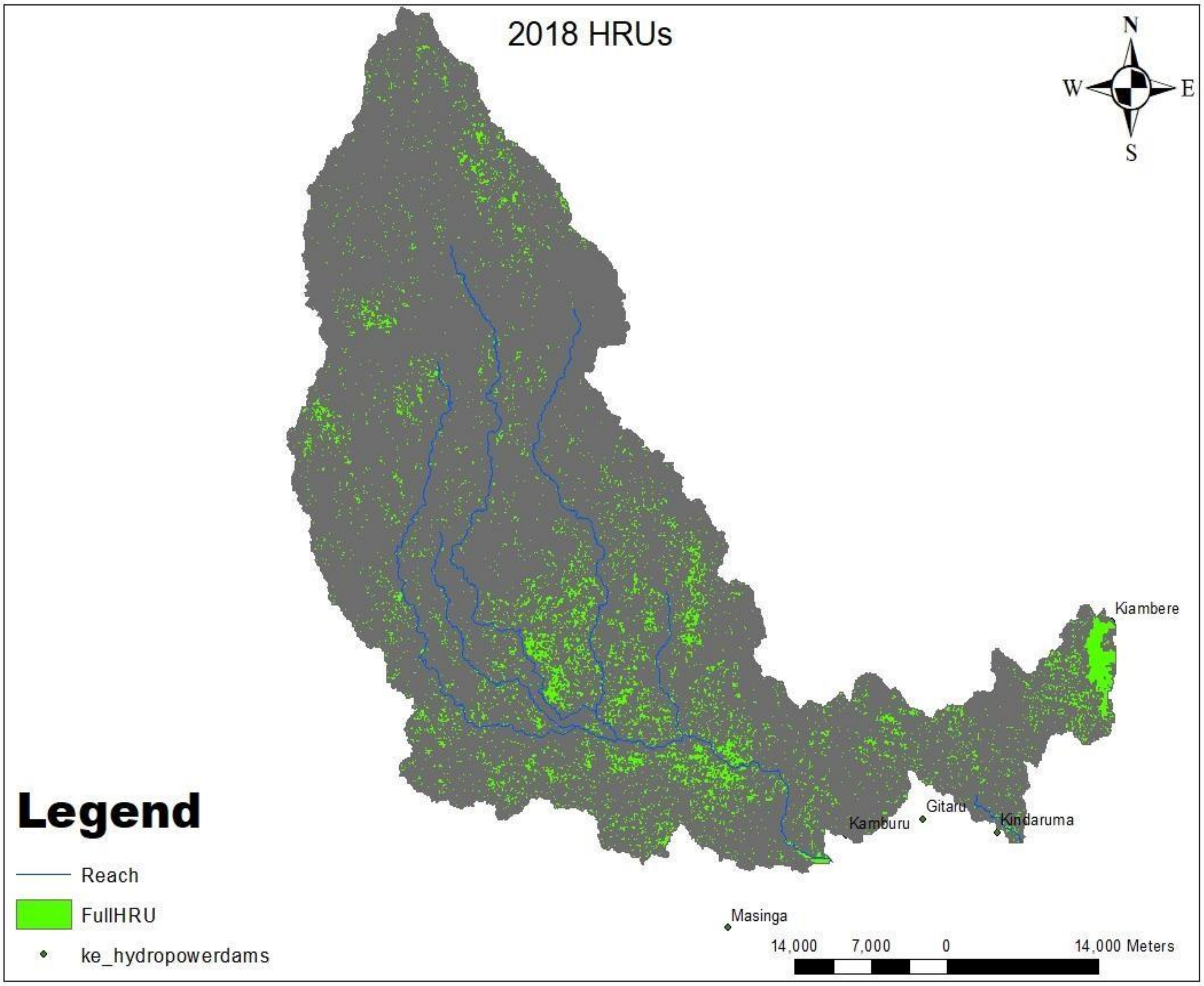

Fig.8; Hydrological response units derived from the land use/landcover and soil map of the year 2018. It shows 744 response units number of the year.

SWAT model was run for a period of 28 years from 1991 to 2019. The warm up period was three years and the average annual simulated sediment yield data in tonnes from the reach number 11 was saved into an excel file as shown in Table 2. Highest sediment yield is seen in the year 2019 and the lowest sediment yield year as 1988. The landuse/ landcover maps were used to derive hydrological response units and the model simulated sediment yield at a period of five years for every lulc map used yearly.

Table 2: Simulated sediment yield data from SWAT model for the period of 28 years.

DATE	SIMULATED SED_OUT MILLION CUBIC CENTIMETERS	HRUs NO.	LULC YEAR	REACH NO.
1/1/1991	600.967	682	1988	11
1/1/1992	1097.78	682	1988	11
1/1/1993	1143.89	682	1988	11
1/1/1994	1811.23	682	1988	11
1/1/1995	1917.33	682	1988	11
1/1/1996	1981.03	725	1994	11
1/1/1997	2600.611	725	1994	11
1/1/1998	3345.63	725	1994	11
1/1/1999	3412.7	725	1994	11
1/1/2000	3598.89	725	1994	11
1/1/2001	7934.65	735	2000	11
1/1/2002	8645.17	735	2000	11
1/1/2003	11179.87	735	2000	11
1/1/2004	13987.76	735	2000	11
1/1/2005	17567.85	735	2000	11
1/1/2006	17856.93	736	2006	11
1/1/2007	21308.8	736	2006	11
1/1/2008	21998.5	736	2006	11
1/1/2009	23697.43	736	2006	11
1/1/2010	37398.34	736	2006	11
1/1/2011	43152.7	738	2012	11
1/1/2012	45678.9	738	2012	11
1/1/2013	47683.43	738	2012	11
1/1/2014	71238.32	738	2012	11
1/1/2015	74323.47	738	2012	11
1/1/2016	100456.2	744	2018	11

1/1/2017	101345	744	2018	11
1/1/2018	109649.4	744	2018	11
1/1/2019	383276.2	744	2018	11
Average annual per year in cubic centimeters	40685.83			
Average annual per year in tonnes	0.041			

Simulated sediment yield data from reach number 11 and subbasin number 11 was calibrated using SWAT_CUP SUFI 2 algorithm with that of observed from 1991 to 1994 and validated from 1996 to 2000. The Nash Sutcliffe objective function was 0.82 in the calibration and 0.82 in the validation as shown in the summary statistics Table 3 below. The table shows the Nash Sutcliffe is 0.82, standard deviation of simulated data as 72,478.66 and the mean value of simulated sediment data as 40685.83. The number of simulations done was 8 and the best simulation number was 1. Fig.9, shows the correlation between simulated sediment yield data and the observed data.

Table 3; Summary statistics table of the objective functions

Variable	p-factor	r-factor	R^2	NS	br^2	MSE	SSQR	PBIAS	KGE	RSR	VOL_FR
SED-OUT 11	0.85	0	0.82	0.82	0.6744	2.60E+10	2.20E+09	5.9	0.88	0.43	1.06
best_sim_no = 1											

Best_goal= 0.8175867										
Behavioral threshold= o.ooo										
Number of behavorial simulations= 8										
mean= 40685.83										
Standard deviation- 72478.66										

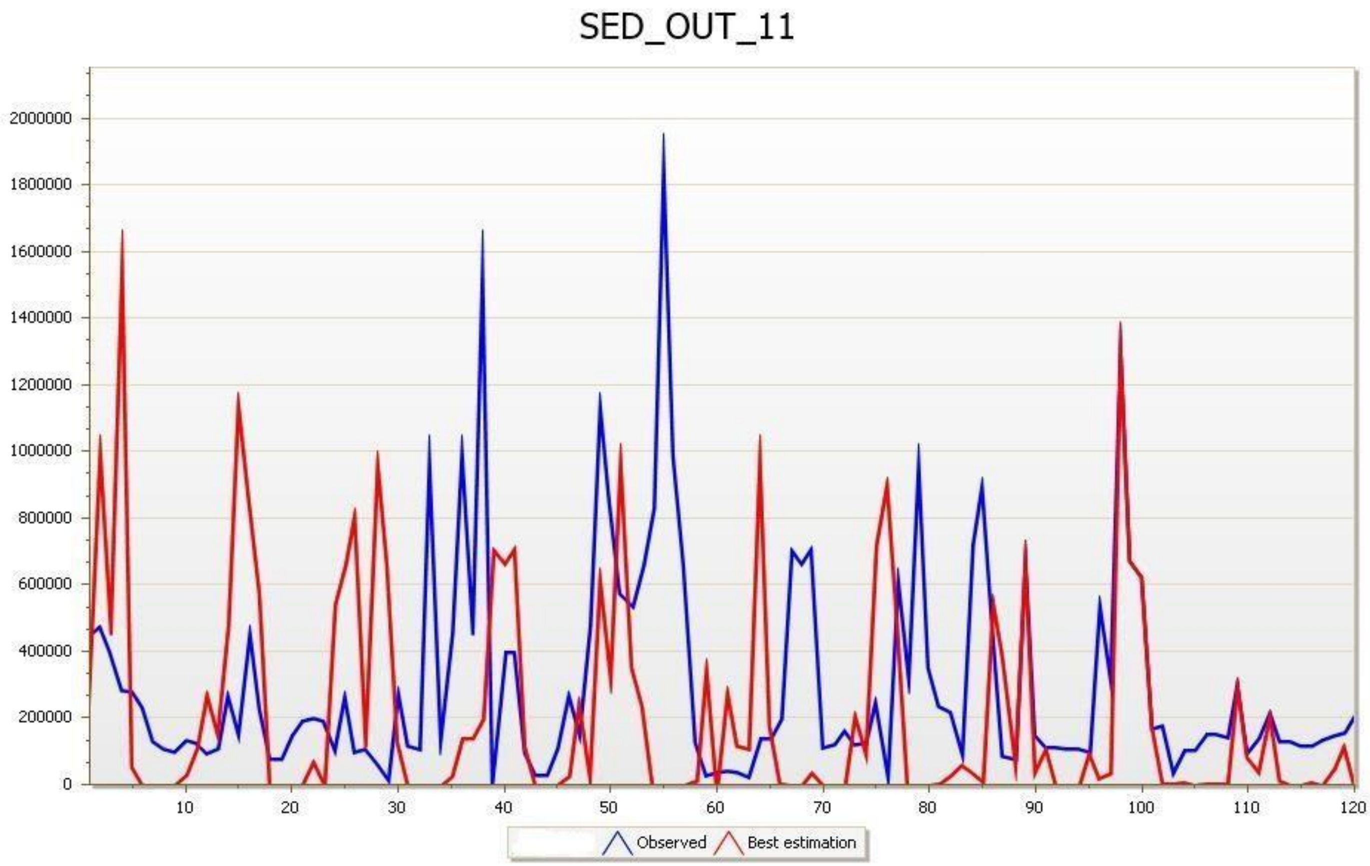

Fig.9; 82% estimation of best simulation plot of both observed and simulated sediment yield data.

DISCUSSION

Hydrological response units

The hydrological response units (HRUs) are areas of the same soil type, land use/ landcover and hydrology properties. According to previous studies carried out, hydrological response units are attributed to hillslope soil erosion and changes in landuse/ landcover maps (Vigiak et al., 2015) and (Gadrani et al., 2018). The reponse units number increases with an increase in soil erosion due to an increase in poor soil conservation practices (Gadrani et al., 2018). These units provide the responsive units of the catchment and affect flow and sediment yield amounts. In this study they were used to obtain sediment yield of the area.

The year 1988 had a total HRU number of 682 which was the lowest number, this showed that little soil erosion was occurring hence low amount of response units. 1994 had a total HRU number of 725, 2000 had a total of 735 HRUs, 2006 had 736 HRU number, 2012 had a total HRU number of 738 and 2018 had the highest total of 744 HRU number. The number of vulnerable response units is increasing over the years due to increased land use/ landcover changes in the catchment. 2018 had the highest amount and this is seen with the fact that most land use was encroached by urban areas hence high soil erosion occurring due to less soil conservation measures as seen in fig. A7.

They are highly distributed in the middle and far south east parts of the region as seen in fig. 3, 4, 5. 6, 7 and 8 these are areas affected by soil erosion which could be attributed to availability of water from reservoirs downstream hence more land use/ landcover activities occurring in these areas such as crop farming of coffee (Hunink et al., 2013).

Estimated sediment yield

The annual average sediment yield was found to be 41000 million cm^3 per year by the model which was high as compared to that of observed data 30000 million cm^3 per year, this was caused by the difference in methodology and datasets used in undertaking the study. Observed data was derived from the echo sounding survey methodology which aimed at looking at the depth of water in the Kindaruma reservoir dam in different stations placed across the reservoir. The year 1988 sediment yield data was 600.967 million cm^3, this was a low amount of sediment yield attributed to low soil erosion as there was little anthropogenic activities being carried out (Hunink et al., 2013). This is also seen in the land use/ landcover map in fig. A2 where the area of forested land was more than that of urban, vegetation, bare land and water. Sediment yield is seen to increase all through the years, this could be attributed to a lot of human activities such as intensive cultivation in the highland areas with minimal land conservation measures leading to soil erosion and the sediments being transported downstream to the reservoirs. Hunink (2013), determined target intervention areas where soil erosion highly occurred in the region. The study found that the erosion rates increased over the years from 1981- 2010. The years 1981-1990 had minimal rates of erosion but the rate increased thereafter. A similar study also shows this trend of erosion (Archer et al., 1996) where most erosive areas are those cultivated for coffee, maize and subsistence crops respectively.

The year 2019 had the highest sediment yield of 383276.2 million cm^3 attributed to high soil erosion occurring in the area due to an increase in anthropogenic activities in the area as seen in the land use/ landcover map in fig. A7 where urban land had the largest coverage as compared to forested, vegetation, bare land and water. Sediment yield is seen to be increasing over the years and according to previous studies carried out in the area this could be proved (Arnold et al., 1994), (Brown et al., 1996) and (Gathagu et al., 2018). The high production rates of sediment could be linked to the fact that the rivers pass through the intensively cultivated slopes of the highlands associated with the Aberdares and Mt. Kenya (Hunink & Droogers, 2011). Lack of adequate ground cover on the steep slopes often cultivated without carrying out effective soil conservation measures results in increased surface runoff and soil loss with subsequent sediment yield (Muchena, Hunink, Droogers, Njuguna, Onduru, Muthuri, Macharia & Maingi, 2012). Increasing sediment yield amounts yearly shows that there is need for soil mitigation measures to reduce the problem of sedimentation in reservoir (Muchena, Hunink, Droogers, Njuguna, Onduru, Muthuri, Macharia & Maingi, 2012).

Calibration and validation

Nash Sutcliffe (NSE) objective function is the best function to calibrate for sediment yield, and it tends from 1 to less than -infinity. Results from calibration of the model showed that the NSE value of 0.82 which is a good value in correspondence to previous studies that indicated that sediment yield objective function of NSE greater than 0.50 and RSR less than 0.70 (Moriasi et al., 2007).

The NSE value of 0.82 shows that the observed and the simulated were well correlated. In a previous study the value of NSE obtained was 0.75 (Hunink et al., 2013) which was a smaller value than that of this study attributed to a difference in data used in the SWAT model where wind, solar radiation and humidity meteorological data was used.

CONCLUSION

This study estimated sediment yield in the upper Tana catchment using the soil and water assessment tool (SWAT) for a period of 28 years. The simulated data was calibrated and validated using observed sediment yield data in SWAT_CUP for the years 1991 to 2000.

The study found that;

- The year with the highest number of hydrological response units was 2018 with a total number of 744 HRUs and the year with the lowest was 1988 with a total number of 682 HRUs.
- High sediment yield year was 2019 due to high soil erosion attributed to increased land use/ landcover change and high number of hydrological response units.
- Low sediment yield year was 1991 due to less soil erosion attributed to minimal land use/ landcover change and low total number of hydrological response units.
- The sediment yield annual average yearly was found to be 0.041tonnes.
- The calibration and validation objective function value of NSE was found to be 0.82 which showed the model was a good tool for sediment yield estimation.

RECOMMENDATION

This study recommends integration of other climatic data such as solar radiation, humidity and wind to enhance the performance of the model. Similar study can be carried out around the other reservoir areas in relation to the observed sediment data from bathymetric survey studies currently being undertaken.

REFRENCES

J.Hunink, E.Njuguna, F.Muchena, P. Maingi & P. Muchiri (2012). Upper Tana-Nairobi Water Fund Technical Report. www.etc-international.org

Briak, H., Mrabet, R., Moussadek, R., & Aboumaria, K. (2019). Use of a calibrated SWAT model to evaluate the effects of agricultural BMPs on sediments of the Kalaya river basin (North of Morocco). International Soil and Water Conservation Research, *7*(2), 176–183. https://doi.org/10.1016/j.iswcr.2019.02.002

Brown, T., Schneider, H., & Harper, D. (1996). Erosion and Sediment Yield: Global and Regional Perspectives (Issue 236).https://www.researchgate.net/2365689.1996.04.011

Bunyasi, M. M., Onywere, S. M., & Kigomo, M. K. (2013). Sustainable Catchment Management: Assessment of Sedimentation of Masinga Reservoir And its Implication on the Dam's Hydropower Generation Capacity. In International Journal of Humanities and Social Science (Vol. 3, Issue 9). www.ijhssnet.com

Bussi, G., Francés, F., José, J., & Julien, P. Y. (2014). Environmental Modelling & Software Distributed sediment yield modelling : Importance of initial sediment conditions. Environmental Modelling and Software, *58*, 58–70. https://doi.org/10.1016/j.envsoft.2014.04.010

Chandra, P., Patel, P. L., Porey, P. D., & Gupta, I. D. (2014). Estimation of sediment yield using SWAT model for Upper Tapi basin. ISH Journal of Hydraulic Engineering, *20*(3), 291–300. https://doi.org/10.1080/09715010.2014.902170

Chandramohan, T., Venkatesh, B., & Balchand, A. N. (2015). Evaluation of Three Soil Erosion Models for Small Watersheds. Aquatic Procedia, *4*, 1227–1234. https://doi.org/10.1016/j.aqpro.2015.02.156

Chia, V., & Mbajiorgu, C. C. (2018). A Comparative Study of KINEROS2 and AGNPS Models in the Estimation of Runoff and Sediment Yield from an Agricultural Watershed A Comparative Study of KINEROS2 and

AGNPS Models in the Estimation of Runoff and Sediment Yield from an Agricultural Watershed. January 2019.https://www.researchgate.net/330565969.2019.001

Daramola, J., Ekhwan, T. M., Mokhtar, J., Lam, K. C., & Adeogun, G. A. (2019). Estimating sediment yield at Kaduna watershed, Nigeria using soil and water assessment tool (SWAT) model. *Heliyon*, *5*(7). https://doi.org/10.1016/j.heliyon.2019.e02106

Figueiredo, E. E. D. E., & Bathurst, J. C. (2007). Runoff and sediment yield predictions in a semiarid region of Brazil using SHETRAN. November 2007, 20–22.https://www.researchgate.net/309030

Gadrani, L., Lominadze, G., & Tsitsagi, M. (2018). F assessment of landuse/landcover (LULC) change of Tbilisi and surrounding area using remote sensing (RS) and GIS. Annals of Agrarian Science, *16*(2), 163–169. https://doi.org/10.1016/j.aasci.2018.02.005

Gathagu, J. N. ang a., Sang, J. K., & Maina, C. W. (2018). Modelling the impacts of structural conservation measures on sediment and water yield in Thika-Chania catchment, Kenya. International Soil and Water Conservation Research, *6*(2), 165–174. https://doi.org/10.1016/j.iswcr.2017.12.007

Ghoraba, S. M. (2015). Hydrological modeling of the Simly Dam watershed (Pakistan) using GIS and SWAT model. Alexandria Engineering Journal, *54*(3), 583–594. https://doi.org/10.1016/j.aej.2015.05.018

Hua, L., He, X., Yuan, Y., & Nan, H. (2014). Assessment of Runoff and Sediment Yields Using the AnnAGNPS Model in a Three-Gorge Watershed of China. May 2012. https://doi.org/10.3390/ijerph9051887

Hunink, J. E., & Droogers, P. (2011). Physiographical baseline survey for the Upper Tana catchment: erosion and sediment yield assessment Client Water Resources Management Authority (WRMA), Kenya. www.futurewater.nl

Hunink, J. E., Niadas, I. A., Antonaropoulos, P., Droogers, P., & de Vente, J. (2013). Targeting of intervention areas to reduce reservoir sedimentation in the Tana catchment (Kenya) using SWAT. Hydrological Sciences Journal, *58*(3), 600–614. https://doi.org/10.1080/02626667.2013.774090

Osei, M. A., Amekudzi, L. K., Wemegah, D. D., Preko, K., Gyawu, E. S., & Obiri-Danso, K. (2019). The impact of climate and land-use changes on the hydrological processes of Owabi catchment from SWAT analysis. Journal of Hydrology*:* Regional Studies, *25*. https://doi.org/10.1016/j.ejrh.2019.100620

Pandey, A., Chaube, U. C., Mishra, S. K., & Kumar, D. (2016). Assessment of reservoir sedimentation using remote sensing and recommendations for desilting Patratu Reservoir , India. Hydrological Sciences Journal, *61*(4), 711–718. https://doi.org/10.1080/02626667.2014.993988

Rizvana. N, Nininja. R & Adlin. R. Sedimentation study in a reservoir using Remote Sensing. (2016). *14*(4), 296–304.http://dx.doi.org/10.15666/aeer/1404_296304

Shivhare, N., Dikshit, P. K. S., & Dwivedi, S. B. (2018). A Comparison of SWAT Model Calibration Techniques for Hydrological Modeling in the Ganga River Watershed. Engineering, *4*(5), 643–652. https://doi.org/10.1016/j.eng.2018.08.012

Sohoulande Djebou, D. C. (2018). Assessment of sediment inflow to a reservoir using the SWAT model under undammed conditions: A case study for the Somerville reservoir, Texas, USA. International Soil and Water Conservation Research, *6*(3), 222–229. https://doi.org/10.1016/j.iswcr.2018.03.003

Tamene, L., Adimassu, Z., Aynekulu, E., & Yaekob, T. (2017). Estimating landscape susceptibility to soil erosion using a GIS-based approach in Northern Ethiopia. International Soil and Water Conservation Research, *5*(3), 221–230. https://doi.org/10.1016/j.iswcr.2017.05.002

Vigiak, O., Malagó, A., Bouraoui, F., Vanmaercke, M., & Poesen, J. (2015). Adapting SWAT hillslope erosion model to predict sediment concentrations and yields in large Basins. Science of the Total Environment, *538*, 855–875. https://doi.org/10.1016/j.scitotenv.2015.08.095

Watershed, O. (2011). Runoff and sediment yield modeling using WEPP in a semi-arid environment (Case study : Orazan Watershed). June 2014.www.mdpi.com/journal/water12

Zhang, R., Santos, C. A. G., Moreira, M., & Freire, P. K. M. M. (2013). Automatic Calibration of the SHETRAN Hydrological Modelling System Using Automatic Calibration of the SHETRAN Hydrological Modelling System Using MSCE. September. https://doi.org/10.1007/s11269-013-0395-z

APPENDIX

Preliminary Results

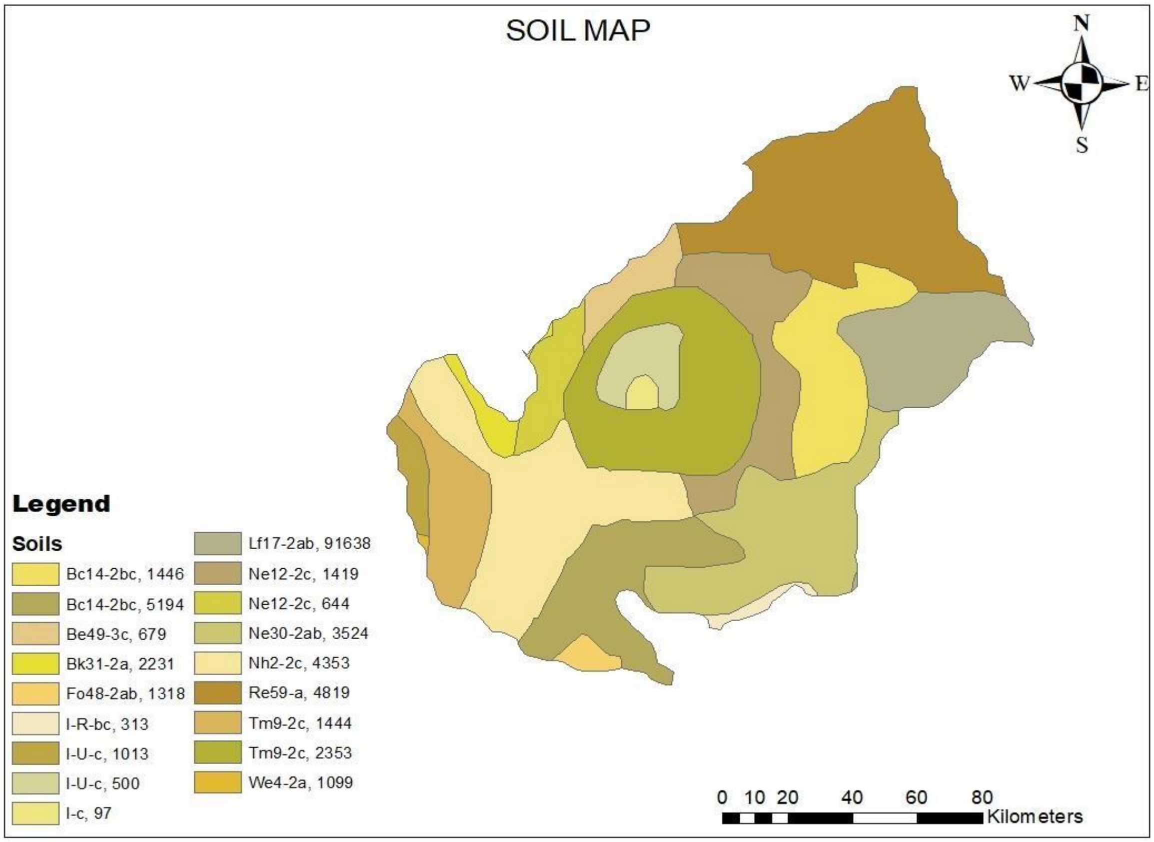

Fig.A1; soil map of the upper Tana catchment area

The soil classes are represented by the slope class and soil unit name. The first part of the soil class represents the soil name for example Bc14-2bc, BC14 represents Chromic Cambisols type of soil and the last part is the percentage of mapping unit and the slope class for example in the example above 2bc represents, 2 is the texture class of the soil type with slope classes b and c.

The soils in the catchment are, Cambisols, Ferralsols, Lithosols, Luvisols, Nitrosols, Regosols, Andosols and Planosols. The slope classes for the area are a, b and c and the texture classes are 2 and 3.

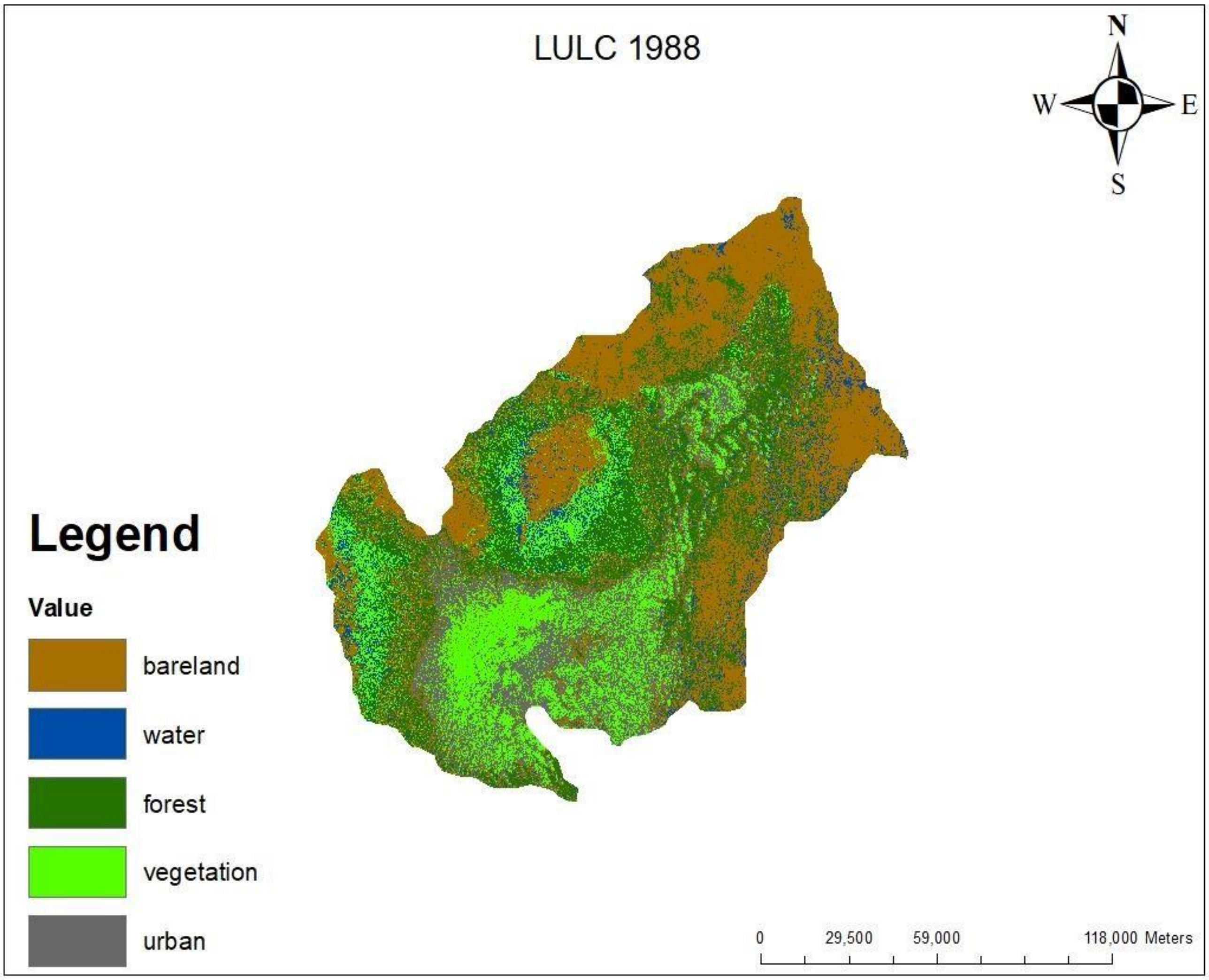

Fig.A2; land use/landcover map of the year 1988.

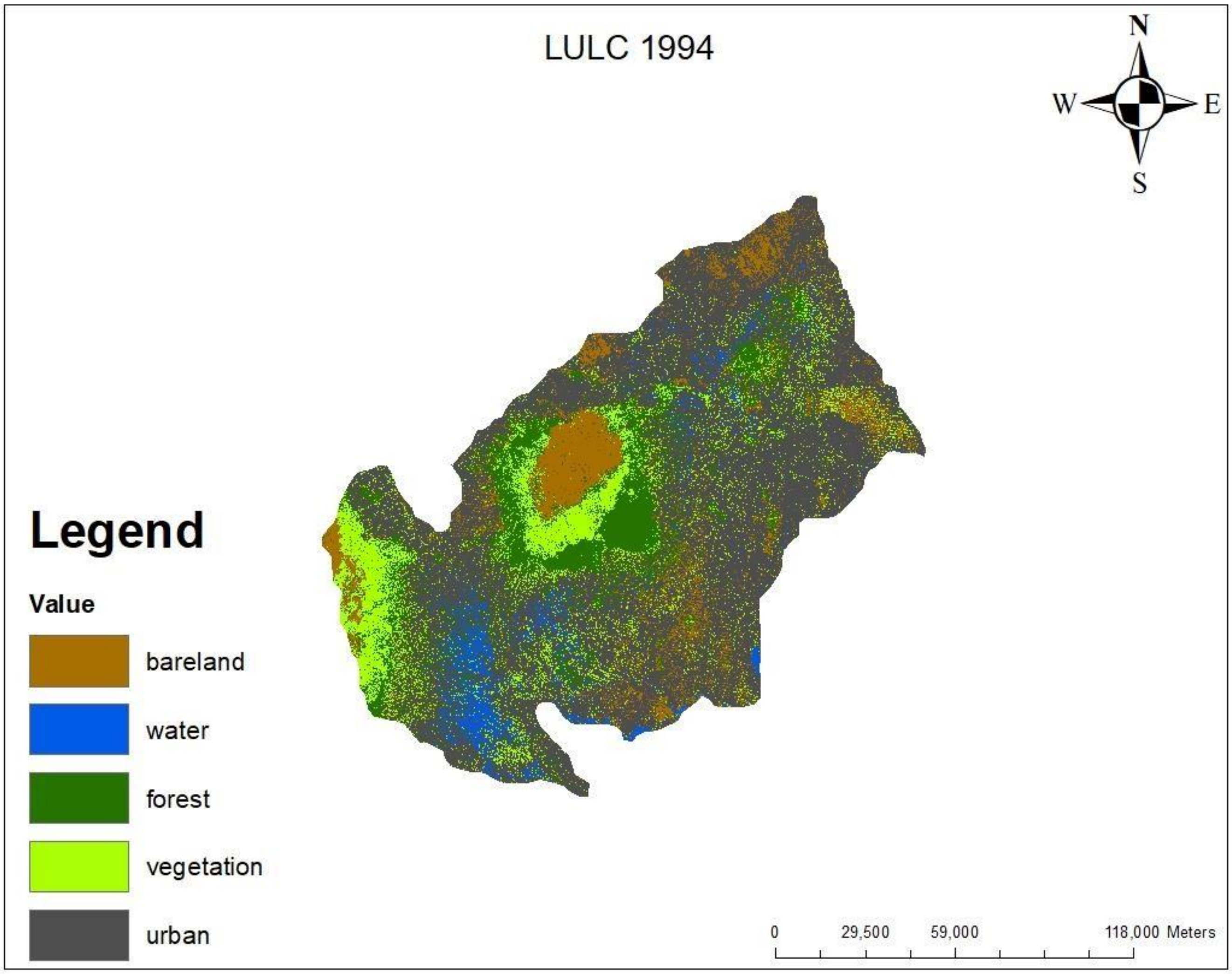

Fig.A3; land use/landcover map of the year 1994

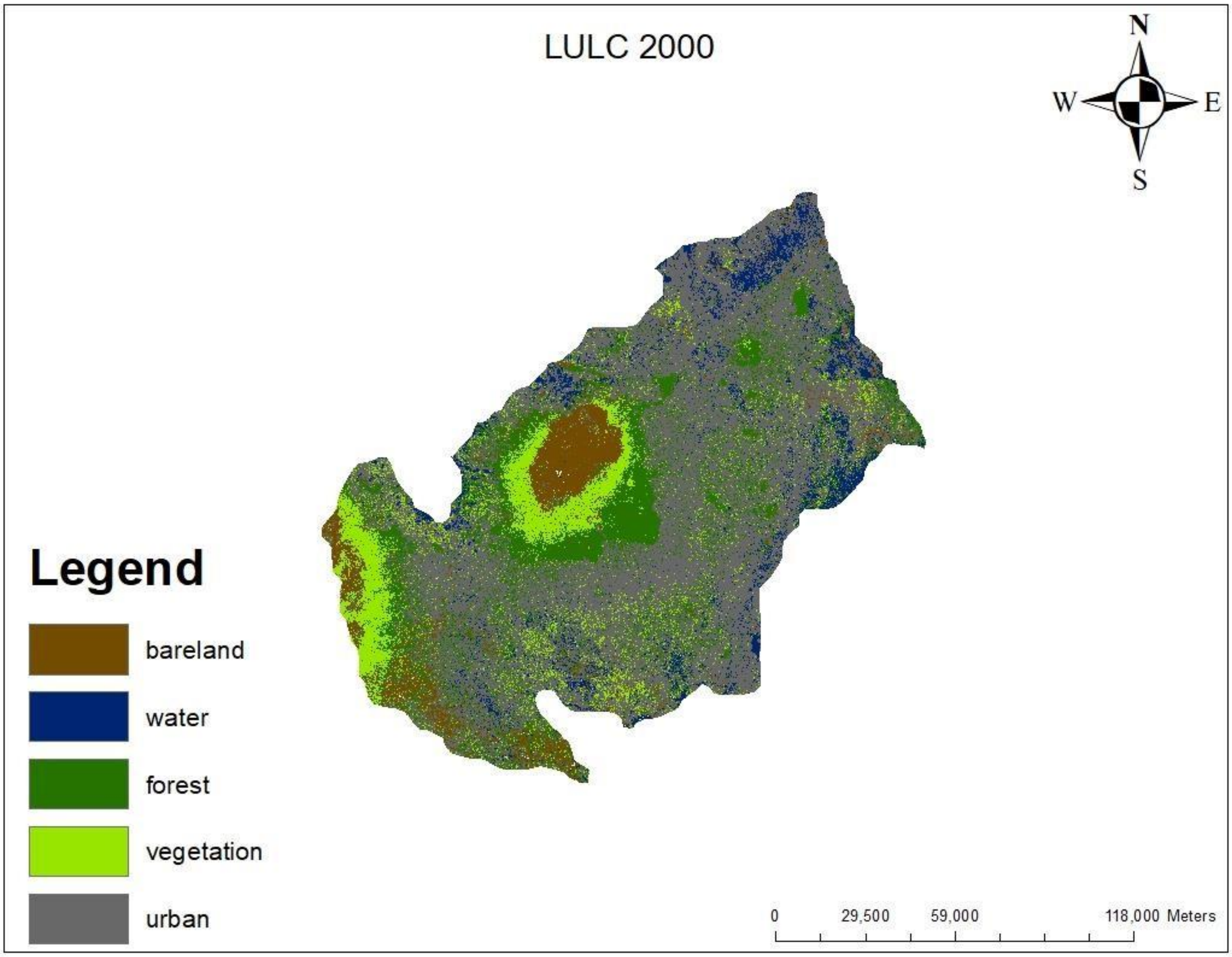

Fig.A4; land use/landcover map of the year 2000

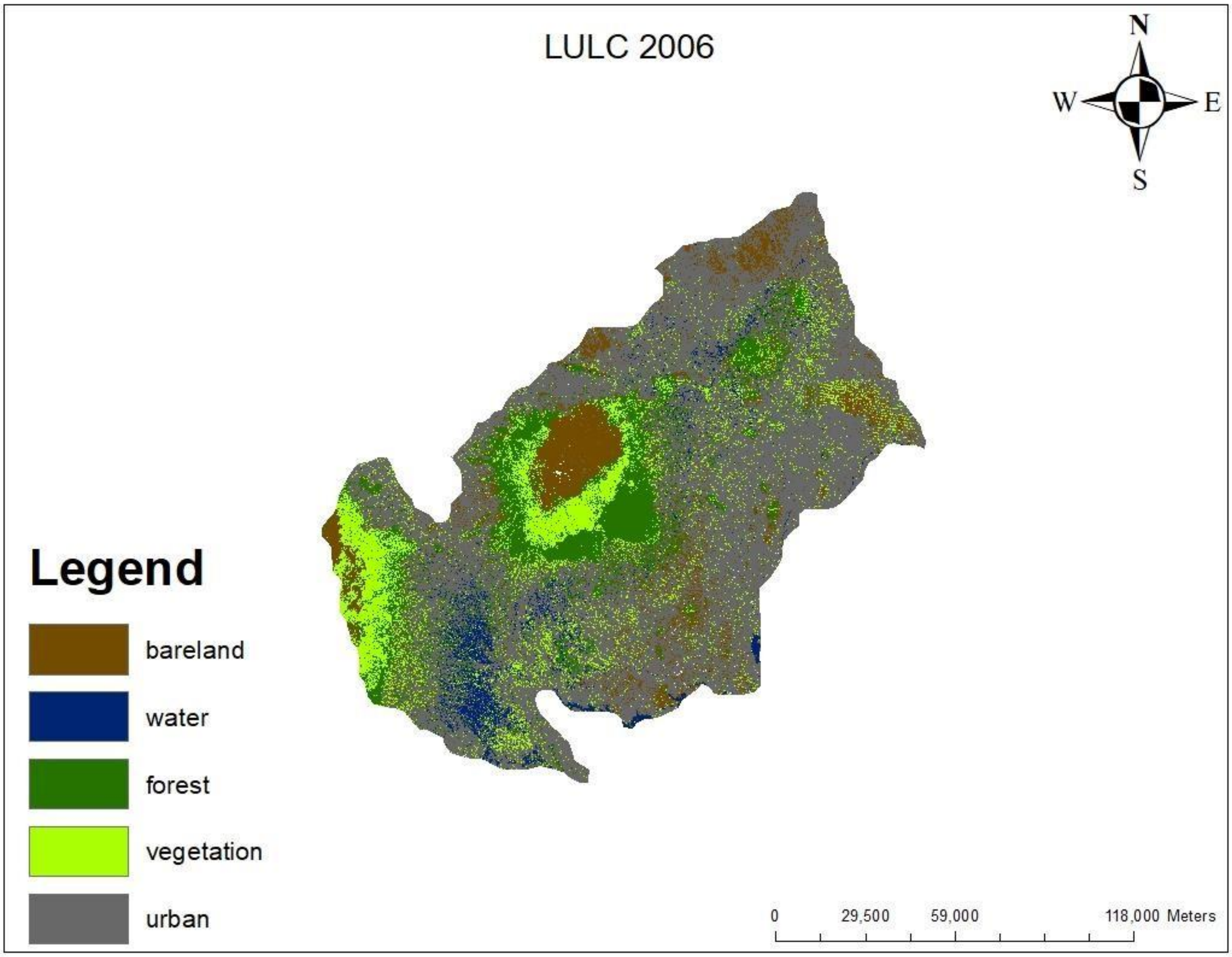

Fig.A5; land use/landcover map of the year 2006

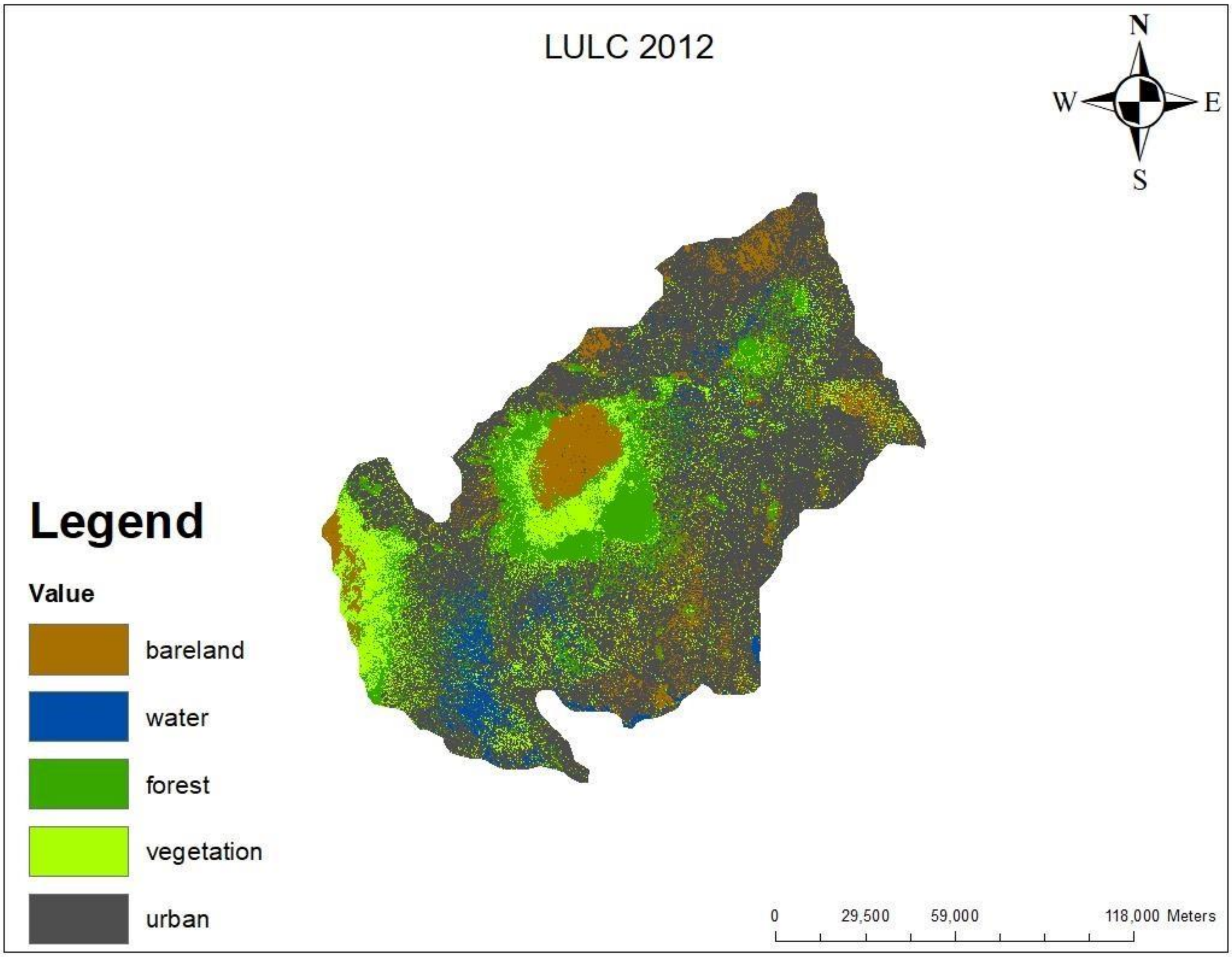

Fig.A6; land use/landcover map of the year 2012

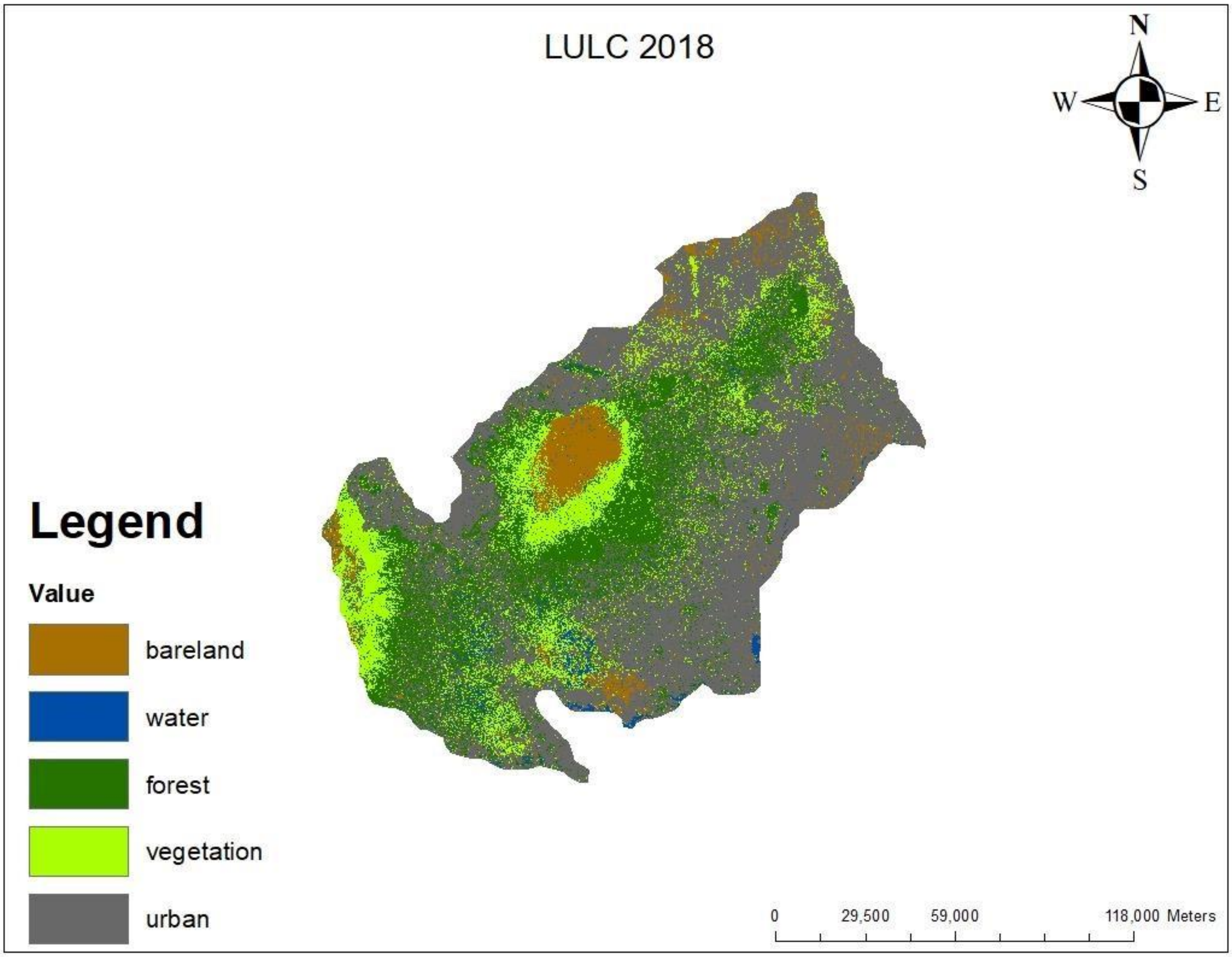

Fig.A7; land use/landcover map of the year 2018

Table A1: confusion matrix of the land use/landcover map of the year 1988

	Bareland	Water	Forest	Vegetation	Urban	SUM	
Bareland	184	0	0	0	0	184	
Water	0	91	0	0	0	91	
Forest	1	1	58	0	0	60	
Vegetation	0	0	2	85	0	87	
Urban	0	0	2	2	74	78	
SUM	185	92	62	87	74	500	500
	34040	8372	3720	7569	5772	59473	536
				overall	0.9817		
				kappa	0.97694		

Table A2: confusion matrix for the land use/landcover map of the year 1994.

	Bareland	Water	Forest	Vegetation	Urban	SUM	
Bareland	213	0	0	0	0	213	
Water	0	92	0	0	0	92	
Forest	0	0	44	1	2	47	
Vegetation	0	0	1	88	0	89	
Urban	0	0	2	0	128	130	
SUM	213	92	47	89	130	571	565
	45369	8464	2209	7921	16900	80863	571

					overall	0.9865	
					kappa	0.9825	

Table A3: confusion matrix for the land use/ landcover map of the year 2000

	Bareland	Water	Forest	Vegetation	Urban	sum	
Bareland	204	0	0	0	0	204	
Water	0	98	0	0	0	98	
Forest	0	0	57	0	0	57	
Vegetation	0	0	0	84	0	84	
Urban	0	0	2	0	125	127	
sum	204	98	59	84	125	570	570
	41616	9604	3363	7056	15875	77514	605
					overall	0.9869	
					kappa	0.9835	

Table A4: confusion matrix for the land use/landcover map of the year 2006.

	Bareland	Water	Forest	Vegetation	Urban	sum	
Bareland	195	0	0	0	0	194	
Water	0	93	0	0	0	93	
Forest	0	0	57	3	1	61	
Vegetation	0	0	0	92	3	95	
Urban	0	0	0	2	116	118	
sum	194	93	57	97	120	561	561
	37636	8649	3477	9215	14160	73137	601
					overall	0.982	
					kappa	0.9775	

Table A5: confusion matrix for the land use/landcover map of the year 2012

	Bareland	Water	Forest	Vegetation	Urban	SUM	
Bareland	203	0	0	0	0	203	
Water	0	85	0	0	0	85	
Forest	0	0	54	1	1	56	

Vegetation	0	0	0	82	0	82	
Urban	0	0	0	1	116	117	
SUM	203	85	54	84	117	543	543
	41209	7225	3024	6888	13689	72035	590
					overall	0.9916	
					kappa	0.9894	

Table A6: confusion matrix for the land use/ landcover map of the year 2018

	Bareland	Water	Forest	Vegetation	Urban	SUM	
Bareland	191	0	0	0	0	191	
Water	0	93	0	0	0	93	
Forest	0	0	47	0	2	49	
Vegetation	0	0	0	86	0	86	
Urban	0	0	0	1	115	116	
SUM	191	93	47	87	117	535	535
	36481	8649	2303	7482	13572	68487	577
					overall	0.9914	
					kappa	0.9892	

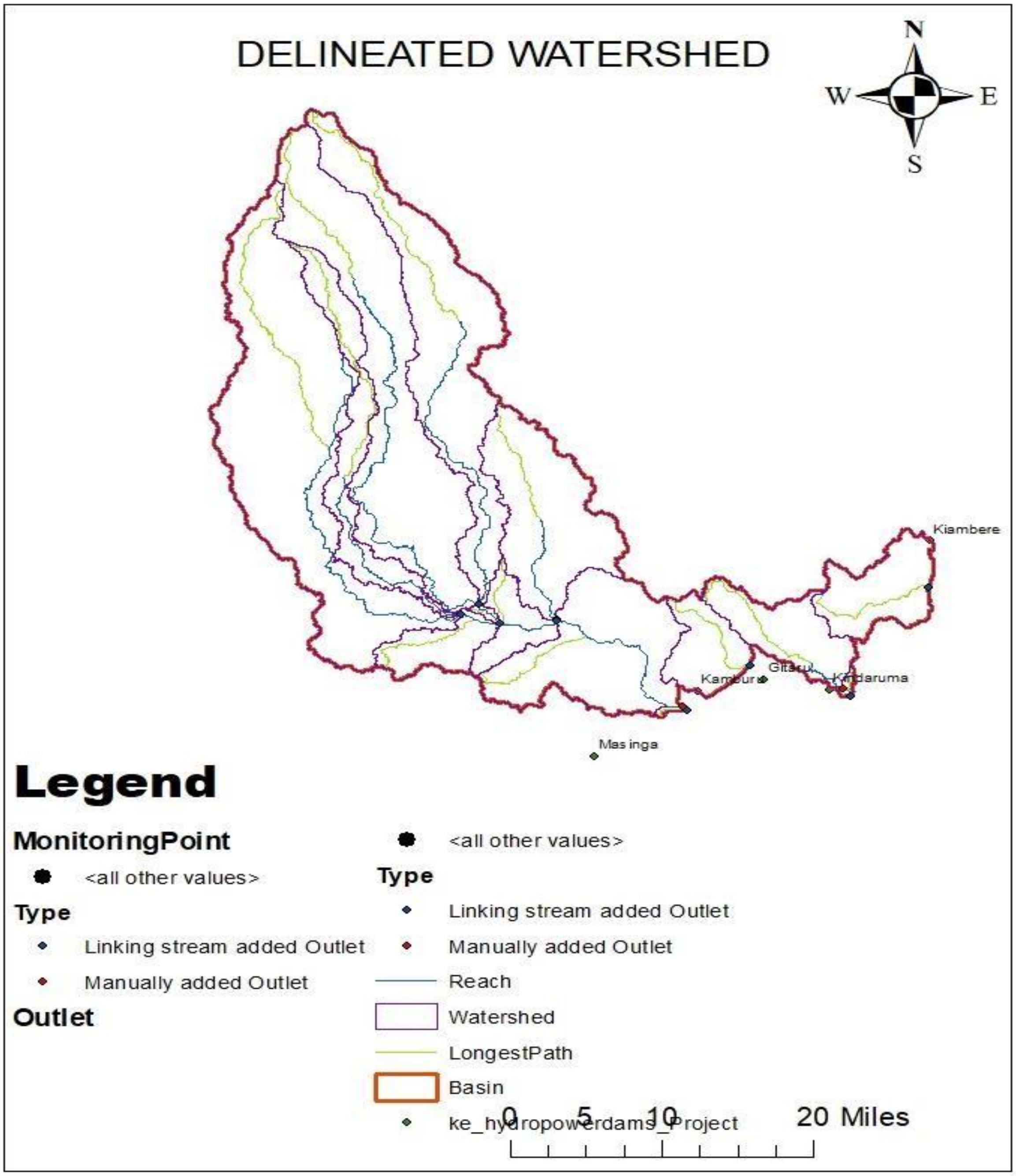

Fig.A8; delineated watershed of the study area derived from the Digital Elevation Model.

Land use/ landcover maps were generated from Landsat 5 and Landsat 7. They were generated using google earth engine code editor. The training points were manually collected and the classes named as; forest, vegetation, urban, bare land and water. The codes used are as shown below.

Code for Landsat 5

```
Map.addLayer(roi,{},'upperTana',true);
var roi = ee.FeatureCollection("users/jillianwangari98/upperTana"); var
landsatCollection = ee.ImageCollection("LANDSAT/LT05/C01/T1")
  .filterDate('2011-07-01', '2013-05-31')
  .filterBounds(roi)
  .map(function(image){return image.clip(roi)}) ; //
Make a cloud-free composite.
var composite = ee.Algorithms.Landsat.simpleComposite(landsatCollection);
Map.addLayer(composite, {bands: ['B3', 'B2', 'B1'], max: 128}, 'composite'); //
Merge the three geometry layers into a single FeatureCollection.
var    newfc         =
snow.merge(bareland).merge(water).merge(forest).merge(vegetation).merge(agriculture).
merge(urban);

// Use these bands for classification.
var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7'];
// The name of the property on the points storing the class label.
var classProperty = 'landcover';

// Sample the composite to generate training data.  Note that the
```

```
// class label is stored in the 'landcover' property.
var training = composite.select(bands).sampleRegions({
collection: newfc,   properties: [classProperty],   scale:
30
});
// Train a CART classifier.
var classifier =
ee.Classifier.smileCart().train({   features:
training,   classProperty: classProperty,
});
// Print some info about the classifier (specific to CART).
print('CART, explained', classifier.explain());

// Classify the composite. var classified =
composite.classify(classifier);
Map.centerObject(newfc);
Map.addLayer(classified, {min: 0, max: 2, palette: ['red', 'green', 'blue']});

// Optionally, do some accuracy assessment.  Fist, add a column of //
random uniforms to the training dataset.
var withRandom = training.randomColumn('random');

// We want to reserve some of the data for testing, to avoid overfitting the model. var
split = 0.7;  // Roughly 70% training, 30% testing.
```

```
var trainingPartition = withRandom.filter(ee.Filter.lt('random', split)); var
testingPartition = withRandom.filter(ee.Filter.gte('random', split));

// Trained with 70% of our data. var trainedClassifier =
ee.Classifier.smileRandomForest(5).train({   features:
trainingPartition,   classProperty: classProperty,
inputProperties: bands
});

// Classify the test FeatureCollection. var test =
testingPartition.classify(trainedClassifier); //Classify the
image with the same bands used for training. var result =
composite.select(bands).classify(classifier)
Map.addLayer(result.randomVisualizer(), {}, 'classfied')

// Print the confusion matrix.
var train_accuracy = classifier.confusionMatrix() print('confusionMatrix',train_accuracy)
print('Overallaccuraccy',train_accuracy.accuracy())
print('kappa',train_accuracy.kappa())
//Export results
Export.image.toDrive({ image:
result,
description: '2012classifiednew',
region: roi, scale:
30,
```

```
});
```

Code for Landsat 7

```
Map.addLayer(roi,{},'upperTana',true); var roi =

ee.FeatureCollection("users/jillianwangari98/upperTana");

/**

* Function to mask clouds based on the pixel_qa band of Landsat SR data.

* @param {ee.Image} image Input Landsat SR image

* @return {ee.Image} Cloudmasked Landsat image

 */ var cloudMaskL457 =

function(image) {   var qa =

image.select('pixel_qa');

  // If the cloud bit (5) is set and the cloud confidence (7) is high

  // or the cloud shadow bit is set (3), then it's a bad pixel.

  var cloud = qa.bitwiseAnd(1 << 5)

                  .and(qa.bitwiseAnd(1 << 7))

                  .or(qa.bitwiseAnd(1 << 3));

  // Remove edge pixels that don't occur in all bands
  var mask2 = image.mask().reduce(ee.Reducer.min());

return image.updateMask(cloud.not()).updateMask(mask2);
```

```
};


var imagery = ee.ImageCollection('LANDSAT/LE07/C01/T1_SR')
                .filterDate('2018-01-01', '2018-12-31')
                .map(cloudMaskL457);


var visParams = {
bands: ['B3', 'B2', 'B1'],
min: 0,   max: 3000,
gamma: 1.4,
};
Map.centerObject(roi, 10); var
upperTana=imagery.median().clip(roi)
Map.addLayer(upperTana, visParams);
// Merge the three geometry layers into a single FeatureCollection. var newfc =
snow.merge(bareland).merge(Water).merge(forest).merge(vegetation).merge(agriculture)
.merge(urban); var label = 'landcover' var bands = ['B1',
'B2', 'B3', 'B4', 'B5', 'B6','B7'] var sample =
upperTana.select(bands).sampleRegions({
  'collection': newfc,
```

```
  'properties': [label],
  'scale': 30
})
var Sample = sample.randomColumn()


var split = 0.7


var training = Sample.filter(ee.Filter.lt('random', split)) var
validation = Sample.filter(ee.Filter.gte('random', split))


//Train a randomforest classifier with default parameters.
var classifier = ee.Classifier.smileRandomForest(10).train(training, label, bands)
//Classify the image with the same bands used for training.
var result = upperTana.select(bands).classify(classifier)
Map.addLayer(result.randomVisualizer(), {}, 'classfied') var
train_accuracy = classifier.confusionMatrix()
print('confusionMatrix',train_accuracy)
print('Overallaccuraccy',train_accuracy.accuracy())
print('kappa',train_accuracy.kappa())
```

```
//Export results

Export.image.toDrive({

image: result, description:

'Classified2018', scale: 30,

region: roi,

});
```

The soil and water assessment tool (SWAT) used lookup text files that could communicate the data it should use from the SWAT database. The lookup text files were created for the soils, land use/ landcover and weather data. The text files contained the value and codes of the data as in the SWAT database. The lookup text files are as shown below.

```
"VALUE","NAME"
0,WATR
1,BARR
2,WATR
3,FRST
4,RNGE
6,URBN
```

Table A7; look up text file for the land use/landcover maps. The codes represent water, bare land, forest, vegetation and urban respectively.

```
"VALUE","NAME"
987,X7-2ab-987
247,Re59-a-247
246,Re59-2c-246
473,Bk31-2a-473
20,Be49-3c-20
155,Ne12-2c-155
440,Bc14-2bc-440
948,Tm9-2c-948
155,Ne12-2c-155
737,Lf17-2ab-737
848,Nh2-2c-848
665,I-U-c-665
977,We4-2a-977
948,Tm9-2c-948
99,I-c-99
665,I-U-c-665
824,Ne30-2ab-824
76,I-R-bc-76
440,Bc14-2bc-440
941,Tm10-2bc-941
76,I-R-bc-76
76,I-R-bc-76
42,Fo48-2ab-42
580,Fr7-2a-580
42,Fo48-2ab-42
76,I-R-bc-76
```

Table A8; lookup text file for the soil data representing all the soil types in the study area and there codes according to the SWAT database.

```
ID,NAME,LAT,LONG,ELEVATION
1,NYERI_TMP,-0.43,36.97,1759.00
2,EMBU_TMP,-0.5,37.45,1493.00
3,MERU_TMP,0.083,37.65,1554.00
```

Table A9; lookup text file for the temperature data containing the weather station data of elevation, latitude, longitude and the station name. the text file is used by SWAT model to obtain information regarding temperature of the stations from the SWAT user weather generator database.

```
ID,NAME,LAT,LONG,ELEVATION
1,NYERI_PCP,-0.43,36.97,1759.00
2,EMBU_PCP,-0.5,37.45,1493.00
3,MERU_PCP,0.083,37.65,1554.00
```

Table A10; lookup text file for precipitation used by SWAT model to obtain precipitation information data in the SWAT weather user generator. The text file contains the weather stations data elevation, station name, latitude and longitude.