3-22-2010

Modeling the Effect of Land Use and Climate Change Scenarios on the Water Flux of the Upper Mara River Flow, Kenya

Liya M. Mango
Florida International University, lmang001@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

MODELING THE EFFECT OF LAND USE AND CLIMATE CHANGE SCENARIOS

ON THE WATER FLUX OF THE UPPER MARA RIVER FLOW, KENYA

A thesis submitted in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

in

ENVIRONMENTAL STUDIES

by

Liya Masiga Mango

2010

To: Dean Kenneth Furton
    College of Arts and Sciences

This thesis, written by Liya Masiga Mango, and entitled Modeling the Effect of Land Use and Climate Change Scenarios on the Water Flux of the Upper Mara River Flow, Kenya, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this thesis and recommend that it be approved.

_____
Jeffrey A. Onsted

_____
Michael E. McClain

_____
Assefa M. Melesse, Major Professor

Date of Defense: March 22, 2010

The thesis of Liya Masiga Mango is approved.

_____
Dean Kenneth Furton
College of Arts and Sciences

_____
Interim Dean Kevin O'Shea
University Graduate School

Florida International University, 2010

DEDICATION

I dedicate this thesis first and foremost to God, my family, to the memories of my late

father Mukasa Mango, and uncle Martin Obuyu. I also dedicate this work to the people of

the Mara River Basin.

ABSTRACT OF THE THESIS

MODELING THE EFFECT OF LAND USE AND CLIMATE CHANGE SCENARIOS

ON THE WATER FLUX OF THE UPPER MARA RIVER FLOW, KENYA

by

Liya Masiga Mango

Florida International University, 2010

Miami, Florida

Professor Assefa M. Melesse, Major Professor

Increasingly erratic flow in the upper reaches of the Mara River, has directed attention to land use change as the major cause of this problem. The semi-distributed hydrological model SWAT and Landsat imagery were utilized in order to 1) map existing land use practices, 2) determine the impacts of land use change on water flux; and 3) determine the impacts of climate change scenarios on the water flux of the upper Mara River.

This study found that land use change scenarios resulted in more erratic discharge while climate change scenarios had a more predictable impact on the discharge and water balance components. The model results showed the flow was more sensitive to the rainfall changes than land use changes but land use changes reduce dry season flows which is a major problem in the basin. Deforestation increased the peak flows which translated to increased sediment loading in the Mara River.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

ATCOR – Atmospheric Correction

BMP – Best Management Practices

CN- Curve Number

DEM – Digital Elevation Model

EM – Electromagnetic Spectrum

FAO – Food and Agriculture Organization of the United Nations

FEWS – Famine Early Warning Systems

GIS – Geographical Information System

GPS – Global Positioning System

HRU – Hydrological Response Unit

IWRM – Integrated Water Resource Management

LULCCS – Land-Use/ Land Cover Classification System

NSE – Nash Sutcliffe Efficiency

NTU – Nephelometric Turbidity Unit

RFE – Rainfall Estimate

SCS – Soil Conservation Service

SOTER-Soil Terrain Database of East Africa

SRTM - Shuttle Radar Topography Mission

SWAT – Soil and Water Assessment Tool

TIFF – Tagged Image File Format

TM – Thematic Mapper

TSS – Total Suspended Solids

USGS – United States Geological Survey

UTM – Universal Transverse Mercator coordinate system

1.0 INTRODUCTION

1.1 Background

According to FAO (2005), the world's water resources are estimated to be about 43,750 km^3/ year distributed globally. Continentally, the largest share of freshwater lies in America which has 45 percent, followed by Asia with 28 percent, Europe with 15.5 percent and Africa has the least amount with 9 percent. In terms of per capita water resources in each continent, America has 24,000 m^3/year, Europe 9,300 m^3/year, Africa 5,000 m^3/year and Asia 3,400.1 m^3/year. These figures exhibit the irregularity of distribution of global water resources and proper management is therefore essential to ensure the supply of this precious resource in adequate quantity and quality to all users in order to fulfill their basic and secondary needs.

Water is an extremely important resource in Kenya and Tanzania and is the lifeline of their ecosystems. It is used for agriculture, industry, power generation, livestock production, and many other important activities. The lack of this crucial resource therefore, is a major impediment to development in both countries.

About 1.9 percent of Kenya is covered by water (SoK 2003). The largest lakes in Kenya are Lake Victoria, Lake Turkana, Lake Naivasha and Lake Baringo. Most of the water is supplied by the country's rivers. The permanent rivers are mostly found in the highlands and the seasonal or intermittent rivers are found in the rangelands. In terms of water supply and demand, Kenya receives marginal rainfall in a larger part of its area (about 80%) with more of the rainfall received in the highland areas. The annual average rainfall of 630 mm, is relatively low for a country on the equator with levels varying significantly both seasonally and annually (FAO 2005). The available total renewable

water resources are 936 cubic meters per person per year and with a total population of 32.8 million (Kenya CBS, 2006), population growth alone will significantly reduce the per capita availability of water of the country. Kenya is categorized as a water scarce country based on the average per capita water availability (WRI, 2007) and this is a major challenge to the country in several ways.

On the other hand, Tanzania has an adequate if not abundant supply of water resources to meet the requirements of users. According to WRI (2003), Tanzania had 2,466.9 and 2,291.2 m^3 per capita of renewable water resources for 2006 and 2007 respectively. As a result of population growth, this figure is expected to drop to 1,500 m^3 per capita by 2025 and will still be above the 1,000 m^3 per capita mark for a country to be considered water scarce. Tanzania's water resources however, are unevenly distributed with rainfall being highly variable both spatially and temporally with some areas receiving around 1,600mm of rainfall annually, while the central drier areas receive lower amounts of around 600 mm annual rainfall. This variability in the supply causes localized and intermittent water shortages that impede and limit various economic activities. The perceived water shortages can be attributed to inadequacies in the national capacity to manage water resources and to inadequate water storage capacity in terms of adequate reservoirs and artificial storage facilities. Establishment of these facilities would be a significant step in harnessing the available water resources and achieving water security in the country as a whole (URT, 2007). In terms of natural storage however, Tanzania is not lacking and has plenty in the form of perennial rivers, lakes and groundwater however, these sources are mostly inaccessible to the greater population. The lack of storage and harvesting

infrastructure greatly reduces the potential for and access to the country's plentiful supply of water resources.
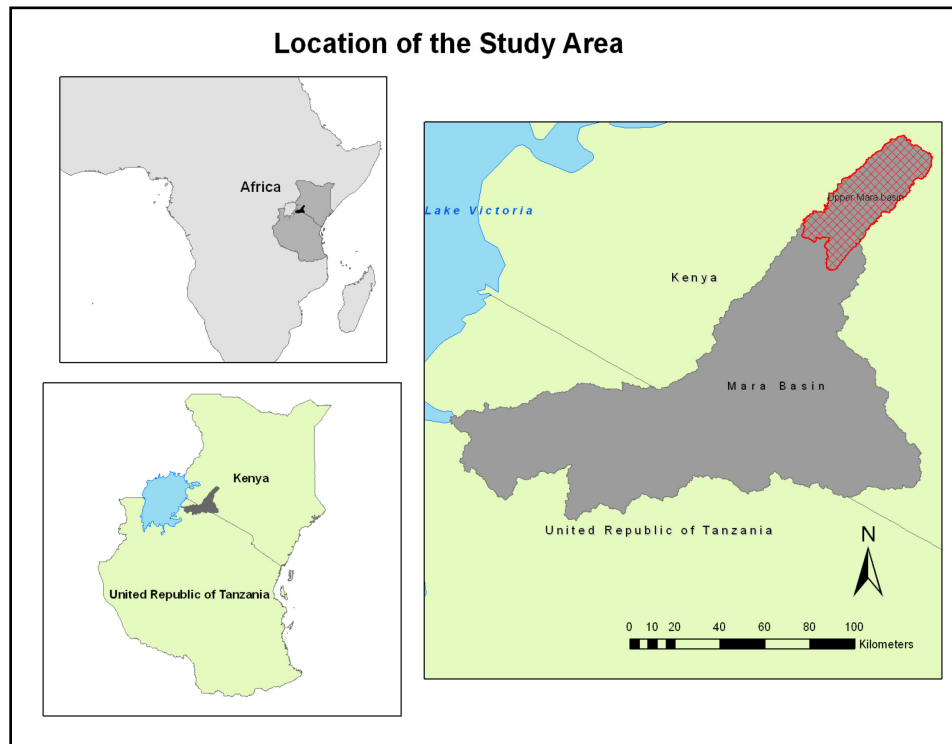
Figure 1-1. Location of the Study Area

1.2 Justification of the Study

The transboundary Mara River Basin is shared between Kenya and Tanzania and is located between longitudes 33.88372^0 and 35.907682^0 West, latitudes -0.331573^0 and -1.975056^0 South. It covers about 13,750 km^2 (Mati et al., 2005) and is characterized by different types of land cover and land uses as a result of different human activities carried out by the stakeholders in various parts of the basin. The land uses include; urban settlements and villages, subsistence and large scale agriculture, forestry, livestock, fisheries, tourism, conservation areas, mining and other industries. The Mara River flows from its catchment in the high altitude Mau Forest in Kenya across different landscapes and finally drains into Lake Victoria at Musoma Bay in Tanzania.
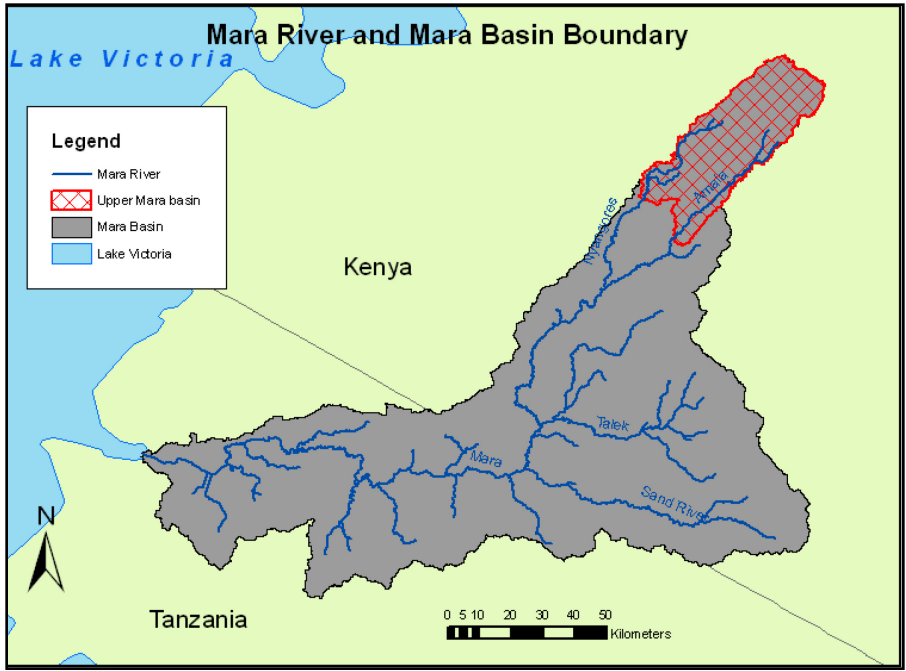
Figure 1-2. Mara River and Mara Basin Boundary
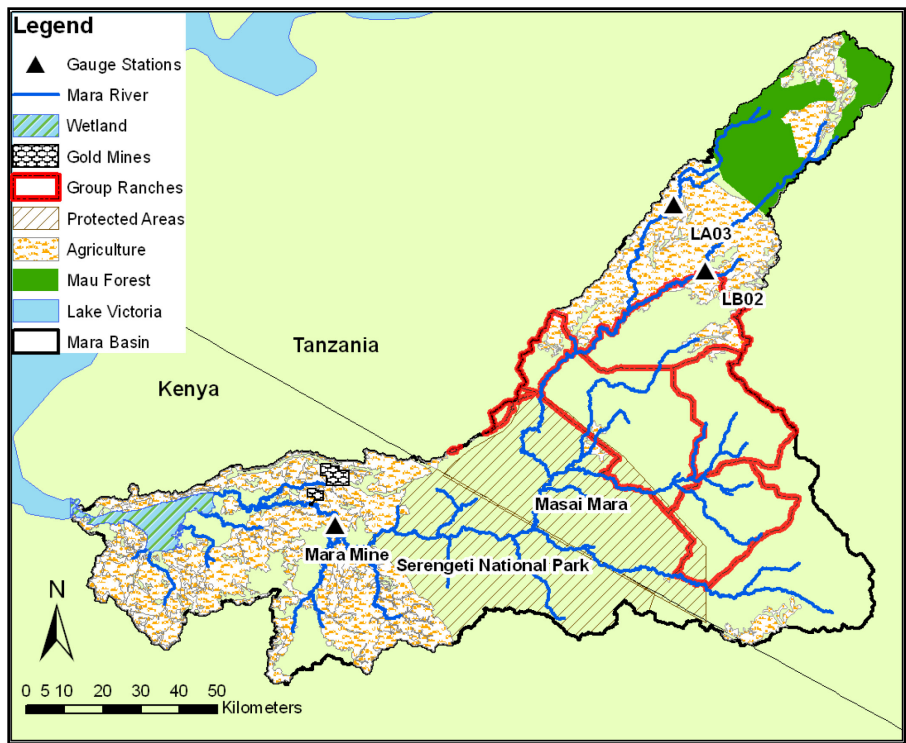
Figure 1-3. Mara Basin Land use

The Mara River is a lifeline to the different ecosystems that it supports especially the Masai Mara Game reserve and the Serengeti National Park which are world renowned biodiversity conservation areas and the large numbers of livestock and game that inhabit the Mara-Serengeti ecosystem. Mounting challenges in terms of water resources present an uncertain future for the Mara Basin. There is need for management guidance to ensure equitable water resource distribution among the stakeholders to fulfill their different needs and to ensure the stability and sustainability of water and other natural resources within the Mara Basin.

Natural resource related challenges faced in the Mara River Basin are: the destruction of catchment forest hence reduced dry season flow, conversion of natural land to agricultural land, pollution from urban effluents, soil erosion resulting from unsustainable agricultural practices, increasing large scale irrigation practices, and increasing domestic water demand driven by population growth. These processes occurring within the basin present serious uncertainties for the water resources in terms of quality and quantity and therefore quantification of these resources is very important for their allocation to different activities and users throughout the vast expanse of the Mara basin.

The Mara River basin is supported by a number of streams and rivers some of them seasonal and few perennial ones that flow all year round. The Mara River is about 395 km long (Mutie et al., 2006) and is formed by the convergence of its two main perennial tributaries, the Amala and Nyangores Rivers that flow from the Enapuiyapui swamp in the Mau Escarpment. Other tributaries include the Talek River, the Engare Engito and the Sand River which is the last main tributary on the Kenyan side of the basin joining the Mara River in the Serengeti plains on the Kenya-Tanzania border. On the Tanzanian side

of the basin the main tributary is the Bologonja River, the Mara River then flows through the Mosirori Swamp, into the Mara Bay, Lake Victoria at Musoma. The Mara River flows a great distance while supplying the ecosystems, human beings, livestock and game with much needed water for survival, health and other necessary natural and artificial processes and practices.

 Land use and land cover varies within the basin and is mainly characterized by the inhabitant's different activities. It consists of mainly closed and open forests, tea plantations in the upper slopes of the Mau Escarpment, agricultural land, shrublands and grasslands used for livestock and game grazing or as game reserves, savannah grasslands which comprise shrub grasslands and wetlands. Land use and land cover change resulting from destruction of the Mau Forest in the upper Mara reaches and conversion of natural land to agriculture among other human activities in other parts of the basin is a major cause of concern has been perceived to have a profound effect on the flow of the Mara River, causing increased peak flows during the rainy season and reduced low flows during the dry season (Mati et al., 2005).

 Owing to the heterogeneity of the Mara River basin and the spatial distribution of topography, soil characteristics, vegetation, land use/ land cover, rainfall and evaporation, a physically based distributed model is a suitable tool for simulation of the ongoing hydrological processes within the basin.

1.3. Knowledge Gap

A number of studies have been carried out in the Mara Basin focused on land use/ land cover change and hydrological processes. Serneels et al., (2001) applied advanced change detection techniques to coarse and fine spatial resolution remote sensing data in the detection of a regional-scale pattern of land cover changes in the buffer zones around the Masai Mara National Reserve and in another study, Serneels et al., (2001) also utilized a spatial statistical model to determine the driving factors of land use change in the land surrounding the Masai Mara National Reserve.

In regard to water availability within the basin, Hoffman, (2006) used GIS to geospatially map available water resources in terms of availability, demand and use in the Mara Basin. Gereta et al., (2002) made use of an ecohydrology model to predict the impact of deforestation, irrigation and a proposed Amala Weir water diversion project on the Serengeti ecosystem. Mati et al., (2005) studied the spatial extent of land use/land cover change between the years 1986 and 2000 in the Mara Basin based on analysis of Landsat ETM (Enhanced Thematic Mapper) imagery. Mutie et al., (2006) evaluated the effects of land use changes on the flow of the Mara River using a combination of remote sensing techniques and the USGS Geospatial Stream Flow Model.

The impact of the climate variability, specifically rainfall and air temperature is also very critical in land cover where it directly influences the plant phenological cycles and in hydrology where rainfall events and dry spells are almost directly reflected in the hydrograph of a watershed. Impacts of temporal variability of rainfall and air temperature on vegetation production in the Mara-Serengeti ecosystem are examined in a study by Ogutu et al., (2007).

Most of these studies however, have focused on the whole basin in its entirety which may be inaccurately represented owing to the spatial extent, variability in; topography, land use/ land cover, soils and climate. The tributaries of the Mara River are critical headwaters that warrant focused attention and this study seeks to determine the impact of land use changes in the upper reaches of the Mara River basin on the flow of the Mara River and the impact of rainfall and air temperature variability. This study also seeks to assess the use of radar rainfall estimates as a viable option to compliment or substitute rain gage data which, in many cases may not be reliable due to various constraints. This will be achieved by means of variable scenarios of land use/land cover and climate change within the SWAT model

1.4 Goals and Objectives

The goal of this research is to determine the effects of different land use and climate change scenarios and the soil and water management practices of the basin on the hydrology of the Mara River. The specific objectives of this study are to:

1. Map existing field scale land use practices using remote sensing and field observations.

2. Determine the impacts of land use change on water flux.

3. Determine the impacts of rainfall and air temperature variation on the hydrology of the upper Mara River

1.5. Research Questions

The research objectives above aim to answer the following research questions:

1. What are the current field scale land use activities in the different sections of the basin?

2. How have land use changes affected the water budget of the Mara River?

3. What are the upper limits of the effects of different land cover composition on the water budget of the upper Mara River?

4. What impact do projected climate change scenarios have on the water budget of the upper Mara River?

5. How does the use of satellite rainfall estimate data compare to rain gage station measurements as an input to the SWAT model?

The findings of this study have provided scenarios on the impacts of land use and climate change in the upper Mara River basin therefore adding to the existing literature and knowledge base and promoting better land use management practices in Kenya and in similar densely populated, highly agricultural watersheds all over the world.

2.0 LITERATURE REVIEW

2.1 Land Cover Classification

SWAT model requires a spatially explicit land cover / land use map that is used in the simulation of the land phase of the hydrological cycle which is the reason for the land cover mapping exercise carried out in this study. The land cover map is also important in the SWAT model to determine the rainfall-runoff characteristics of the watershed. The expected outcome is a classified land cover map of a high accuracy (about 80%) consistent with a classification scheme adapted to the existing land use/land cover classes in the Mara Basin. The land cover mapping was achieved by means of remote sensing data, a useful addition to the already existing classical methods that make use of ground surveying. In satellite based classifications using Landsat and SPOT imagery, most authors have claimed accuracy of between 60 to 90 per cent (Skidmore et al., 1988). Therefore, remote sensing was seen as a useful approach to derive spatially explicit land surface characteristics from remotely sensed data, where ground surveying is not feasible.

Components of a remote sensing project include: A problem statement, a hypothesis, a methodology which includes the choice of imagery based on spatial, spectral and temporal scale, consideration of biophysical variables, classification needs, a classification scheme and a suitable method of classification.

2.1.1 Classification scheme

The land use/land cover classification system most suitable for this study and compatible with the schema used by the hydrological model SWAT is the U.S Geological Survey

Table 2-1 Anderson Land Use/ Land Cover Classification Scheme

LEVEL I	LEVEL II
1 Urban or Built-up Land	11 Residential 12 Commercial and Services 13 Industrial 14 Transportation, Communications, and Utilities 15 Industrial and Commercial Complexes 16 Mixed Urban or Built-up Land 17 Other Urban or Built-up Land
2 Agricultural Land	21 Cropland and Pasture 22 Orchards, Groves , Vineyards, Nurseries, and Ornamental Horticultural Areas 23 Confined Feeding Operations 24 Other Agricultural Land
3 Rangeland	31 Herbaceous Rangeland 32 Shrub and Brush Rangeland 33 Mixed Rangeland
4 Forest Land	41 Deciduous Forest Land 42 Evergreen Forest Land 43 Mixed Forest Land
5 Water	51 Streams and Canals 52 Lakes 53 Reservoirs 54 Bays and Estuaries
6 Wetland	61 Forested Wetland 62 Nonforested Wetland
7 Barren Land	71 Dry Salt Flats. 72 Beaches 73 Sandy Areas other than Beaches 74 Bare Exposed Rock 75 Strip Mines Quarries, and Gravel Pits 76 Transitional Areas 77 Mixed Barren Land
8 Tundra	81 Shrub and Brush Tundra 82 Herbaceous Tundra 83 Bare Ground Tundra 84 Wet Tundra 85 Mixed Tundra
9 Perennial Snow or Ice	91 Perennial Snowfields 92 Glaciers

Source: Anderson et al., 1976

Land-Use/ Land-Cover Classification System (LULCCS) for Use with Remote Sensor Data which is primarily a resource–oriented land-cover classification system. This classification scheme is designed for US land cover types and needed to be modified to match existing land covers in the Mara basin. A classification scheme has to be taxonomically correct, which means it is exhaustive, mutually exclusive, and hierarchical in order to guarantee high classification accuracies (Jensen, 2005).

The classification scheme used for the Mara basin is a modification of the Anderson land use land cover classification scheme for remote sensor data and included only the land use/ land cover types found in the Mara basin shown in the table below:

Table 2-2 Mara Land use/land cover classification scheme

Land Cover Type
Cloud
Forest
Water
Bushland
Grassland
Agriculture
Bare soil

This scheme is implemented in the Mara Basin and was used in conjunction with the Anderson classification scheme for the classification of the different land use/ land cover types in the Mara Basin.

2.1.2 Conceptual discussion on Mara Land use/ Land cover classification

2.1.2.1 Variable consideration

The variables considered for the study were biophysical variables and hybrid variables. Hybrid variables are created by systematically analyzing more than one biophysical variable (Jensen, 2005) and are such as land use/ land cover which is obtained from the percent reflectance of the different spectral bands of the Landsat TM image. Biophysical variables consist of fundamental biological and physical information obtained by the remote sensor system. The biophysical variables utilized in the study were the Kauth Thomas Tassel Cap Transformation, the x, y location from the orthocorrected imagery and texture in terms of brightness values of the pixels in theimagery.

2.1.2.2 Data choice

Landsat TM imagery was chosen because of its high spatial resolution, regularity of acquisition and it is also readily available for multiple days and years. This imagery was downloaded from the United States Geological Survey Global Visualization (USGS-GLOVIS) website. Ground reference data from a field study carried out in June 2008 was available as a source of ancillary data and was of much importance in the classification of the satellite imagery.

2.1.2.3 Spectral information and resolution

This refers to the number and dimension of the specific wavelength intervals (referred to as bands or channels) in the electromagnetic spectrum that a scanner or sensor is sensitive (Jensen, 2005). The Landsat Thematic Mapper sensor's spectral resolution consists of 7

spectral bands. They are; the Blue, Green, Red, Near infrared, Thermal infrared and 2 Middle infrared bands. Different bands of the electromagnetic spectrum are used to detect or differentiate biophysical variables and the selection of these bands is carried out in a way to increase the contrast between them and their background, improving the probability that the desired features will be separated from the remote sensor data.

2.1.2.4 Spatial information and resolution

The spatial resolution is the measure of the smallest angular or linear separation between two objects that can be resolved by the remote sensor system. The Landsat 4 Thematic Mapper has a swath width of 185 km and a nominal spatial resolution of 30 x 30m for six of its multispectral bands. A rule of thumb is to have a nominal spatial resolution of at least less than half the size of the object you wish to detect/map (Jensen, 2005). The Landsat imagery is a suitable choice for classification of the features that are of interest in a spatial sense. However, this is still of little importance if the features cannot be separated spectrally from the background.

2.1.2.5 Temporal info and resolution

This refers to how often a remote sensor system maps a particular area. The Landsat Thematic Mapper has a temporal resolution of 16 days meaning it obtains imagery of a particular location every 16 days. Remote sensing obtains images at particular points in time and this can be used to observe trends, processes and make predictions. The temporal resolution of 16 days was sufficient for the land cover mapping because it did

not involve the detection of change or analysis of any trend occurring over a period of time.

2.1.2.6 Radiometric info and resolution

The radiometric resolution is the sensitivity of a remote sensing detector to differences in signal strength as it records the radiant flux emitted, or back scattered from the terrain. It defines the number of signal levels that can be discriminated and has a major impact on the ability to measure properties of objects in a scene. The Landsat Thematic Mapper records data in 8 bits (values from 0 to 255). High radiometric resolution increases the chances that phenomena will be remotely sensed more accurately.

The considerations of the various capabilities of the sensor and attributes of the imagery are highly dependent on the purpose of the study and the variables to be extracted. This was an important piece of the land cover mapping process and laid the groundwork for a successful land cover classification.

2.1.3 Land Cover Classification by means of an Expert System

There are different types of supervised classification that can be used in the extraction of thematic information from an image. Classification can be carried out by a number of methods and they include; Algorithms based on parametric and nonparametric statistics that use ratio and interval –scaled data, use of hard or soft (fuzzy) set of classification logic to create their respective thematic outputs, use of per-pixel or object oriented classification logic, and hybrid approaches (Jensen, 2005).

In a supervised classification, the location of the land cover and land use types are known prior to classification and uses a combination of ancillary data which may be from a field work study, interpretation of photographs, map analysis and personal experience. Well known homogenous land cover types are located from the ancillary data and their spectral characteristics are used to train the classification algorithm for the eventual classification of the rest of the image. Maximum likelihood classifiers which are a widely used form of supervised classification, assume the training data statistics for each class in each band possess a multivariate normal distribution (Gaussian) in order to compute the class variance and covariance matrices. In this case training data with bi-modal or n-modal histograms in a single band are not ideal. According to Foody (1996), it is difficult to incorporate non image, categorical data into the maximum likelihood classification.

On the other hand, supervised classification based on decision rules have proven to be successful and they are able to perform a classification based on conditions and rules stored as a knowledge base within the computer and these can be called upon when it comes to the solution of classification problems. Expert classifiers make no assumption regarding the distribution of the data and are non parametric. Non parametric clustering algorithms are able to incorporate any kind of spatially distributed data into the classification and do not require normally distributed training data (Jensen, 2005). Non parametric algorithms have a higher chance of success for this study because unlike the maximum likelihood classifiers which are based on probability and the assumption of the normalcy of the datasets, non parametric algorithms are able to successfully handle datasets that are not normally distributed with accurate results.

An expert system used in land cover classification makes use of a decision tree that predicts class membership. In the tree structure, rules and conditions are evaluated in order to test hypotheses (Jensen, 2005). These hypotheses are the classes that are to be classified. Advantages of expert systems include: It is possible to evaluate the output of an expert system and evaluate how the conclusion was reached, this method makes no assumption regarding the distribution of the data, the decision tree can reveal non-linear and hierarchical relationships among the input variables and use then to predict class membership, machine learning techniques are able to deal effectively with tasks involving highly dimensional data such as hyperspectral data. According to Pat and Mather (2001), a decision tree classifier is able to perform automatic feature selection with reduced complexity while the tree structure provides a clear and understandable picture of the predictive and generalizing ability of the data. From these and other available methods for classification, the user can be able to select whichever method that best fulfill their classification objectives.

An expert system was used in this study based on reasons and advantages that are discussed later in this chapter.

2.1.4 Land Cover Map Accuracy Assessment

Thematic information derived from satellite data may contain error and it is important that these sources of error are identified and minimized as much as possible and the user is informed by the preparer of the data how much confidence they should have in the thematic information. Remote sensing–derived thematic maps therefore, should be

subjected to a thorough accuracy assessment before being used in scientific investigations and policy decisions (Stehman and Czaplewski, 1998; Paine and Kiser, 2003)

The accuracy of a land cover map derived from remotely sensed data is a function of two basic factors: the information content of these data in relation to the classification legend desired, and the correctness of the labeling step. The first factor can be understood in terms of the uniqueness of the relationship between the land cover type to be identified and the signal recorded by the sensor. On the other hand, the second factor depends primarily on the expertise of the analyst. To understand the land cover map accuracy and its limitations, the magnitude and relative importance of these causes must be determined (Cihlar, 2003). To be able to determine the accuracy of the map, attention is focused on the error or confusion matrix.

2.1.4.1 Kappa Analysis

Kappa analysis is a discrete multivariate technique used in accuracy assessment and yields a statistic $\widehat{K}$ (K hat Coefficient of Agreement) which is an estimate of Kappa and is a measure of agreement or accuracy between the remote sensing-derived classification map and the reference data as indicated by a) the major diagonal and b) the chance agreement, which is indicated by the row and column totals (Rosenfield and Fitzpatrick-Lins, 1986)

K is computed using the formula:

$$\widehat{K} = \frac{N \sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k} (x_{i+} \times x_{+1})}{N^2 - \sum_{i=1}^{k} (x_{i+} \times x_{+1})} \tag{2-1}$$

Where k is the number of rows (e.g., land cover classes) in the matrix, x_{ii} is the number of observations in row i and column i, and x_{i+} and x_{+1} are the marginal totals for row i and column i, respectively, and N is the total number of observations. $\widehat{K}$ values >0.80 (i.e., >80%) represent strong agreement or accuracy between the classification map and the ground reference information. $\widehat{K}$ values between 0.40 and 0.80 (i.e., 40 to 80%) represent moderate agreement. $\widehat{K}$ values <0.40 (i.e., <40%) represent poor agreement (Landis and Koch, 1977)

2.2 Hydrological Modeling

2.2.1 The Hydrological Cycle, GIS and Runoff Modeling

The Earth holds a large amount of water in different spheres which is in constant motion, which is known as the hydrological cycle (Neitsch et al., 2005). This cycle shows water as it travels through different global systems or storages by means of different processes. These processes are precipitation, runoff, evaporation, infiltration, transpiration, percolation, recharge of groundwater, interflow and groundwater discharge.

Precipitation is water released from the atmosphere in forms such as rain, snow, sleet, or hail. This results from the evaporation of water from the earth into the atmosphere where it is temporarily held and accumulates saturating the atmosphere and eventually released. Precipitation is considered as the major input in watersheds models.

Evaporation occurs when water is changed from a liquid state to a vapor state and moves back in to the atmosphere. This is increased by solar radiation, increases in air and wind temperature while high moisture content in the air reduces the potential for

evaporation. *Transpiration* is the release of water by plants as a by-product of photosynthesis. Evapotranspiration is the combination of the two processes of evaporation and transpiration because of the difficulty in separation of the two processes (Ward, 2005).

Infiltration is the entry of water into the soil. This is governed by different factors and as a result of this it varies from place to place. It is dependent on soil properties such as the organic matter content, density, texture, hydraulic conductivity and porosity. The soil surface conditions also affect infiltration where compacted soil will restrict infiltration, vegetation like forests slow down the water flow and allow for more infiltration as opposed to paved areas. Topography, roughness and slope and human activities that alter the soil surface like in urban and agricultural areas affect the infiltration of water in an area.

Percolation is the downward movement of water after it enters the soil by means of gravity through the soil profile. That which moves past the plant root zone toward the underlying geologic formation is called deep percolation, is out of reach of the plant roots and goes towards replenishing the groundwater supply and this process is known as groundwater recharge.

Runoff is the portion of precipitation, snowmelt, or irrigation water that flows over and through the soils, eventually making its way into surface water systems. This component of the hydrological cycle is of a lot of importance in this study as it is the component the hydrological model SWAT using existing data will simulate for different scenarios. Contributions to it include and overland flow, interflow and groundwater flow. A large percentage of surface runoff reaches streams, where it's described as streamflow or

discharge. Overland flow can also occur when the soil is saturated (soil storage is filled). When all the voids, cracks and crevices of the soil profile are filled with water and the excess begins to flow over the soil surface.

Interflow may occur when the water's downward movement is restricted by an impenetrable layer of material which causes it to move laterally and discharge that may have been formed naturally or by human activities.

Groundwater flow occurs in the hydrological cycle and this process creates a baseflow for surface water bodies and for groundwater recharge. A large percentage of this water is used for drinking and irrigation.

The hydrological cycle is extremely important in the movement of water in any natural system. Man's activities in the Mara Basin have altered the hydrological processes and have affected the quality and quantity of water over time. The study aims to simulate the hydrological cycle in the study area in different scenarios and predict the impacts of different land use and land cover on the sediment and water output of the Mara River. This simulation or prediction is made possible by application of a hydrological runoff model in a Geographical Information System (GIS).

Geographic Information Systems (GIS) has become a key part of hydrologic studies because it has proved useful in hydrologic modeling processes such as the spatial and temporal distribution of inputs and parameters controlling surface runoff. Maps describing topography, land use, land cover soils, rainfall and meteorological variables may become model parameters or inputs in the simulation of hydrologic processes (Vieux, 2001).

GIS is beneficial in hydrological modeling because it is able to provide a visual based simulation environment and scenario management and analysis capabilities. It is also much easier and practical to display and assess the hydrological, spatial and seasonal variability of the parameters involved in the modeling process. GIS provides well developed algorithms to deal with geographic data of high spatial detail and information content. However, depending on the watershed, a lot of this data requires preprocessing and conversion before it can be used for distributed hydrological modeling.

Hydrological models usually require a surface representation of a parameter measured at points, several methods for generating a two dimensional surface from point data have been developed and include; Kriging, Moving average, Splines, Local regression, and Linear interpolation (Vieux, 2001). Values can be interpolated across distinct zones. GIS is also valuable in determining the spatial resolution necessary for hydrological process simulation. Delineation of drainage networks from Digital Elevation Models (DEM) is another important function of GIS in hydrological modeling.

GIS is also useful in the analysis of land use and land cover patterns in terms of spatial and temporal variability, soil types, population distribution

2.2.2 Methods of simulation/prediction of river flow and sediment concentration

The primary aim of the traditional hydrological model is to predict the amount of the discharge from a drainage basin. There are two fundamental or classical types of hydrological models; deterministic and stochastic models and these can be further described whether the description of hydrological process is empirical or physically based (Olsson and Pilesjo, 2002).

There are three major types of deterministic models; physically based models, empirical lumped models, empirical distributed models, and physically based distributed models. Empirical models are based on regression and correlation results from statistical analyses from time series data. The equations derived are based on measurement knowledge or observed phenomena without demands on understanding of the underlying process and are often referred to as black box models. Truly physical models are based on formulas of physical relations and are referred to as white box models (Kirby et al. 1993) because every part of the process is understood.

Physically based models are most suited for studying catchment change scenarios. Prediction of discharge from catchments and monitoring of pollutant and sediment dispersal are well suited for physical models (Abbot and Refsgaard 1996). It is important to note that the whole conceptual understanding of a hydrological system cannot be expressed in mathematical terms thus there will always be a systematic error introduced based on the excluded or unknown relationship. This is a source of error in many physical modeling processes which gives rise to the need for calibration of the model to time series data (Olsson and Pilesjo, 2002).

In a lumped model, the model uses parameter band variables that represent average values for the entire catchment. The averages can be derived either physically or empirically which can give the model a semi-empirical appearance. These lumped models are mainly used in rainfall-runoff modeling.

Distributed hydrological models are supposed to describe flow processes in each and every point inside a catchment. Difficulties in the general and conceptual framework coupled with time and memory consuming programs make these models practically

impossible to use (Olsson and Pilesjo, 2002). Simpler models instead try to estimate the different flow patterns discretisised into nodes with orthographic spacing and these nodes can be seen as center points in square shaped areas known as pixels or cells. Models based on this type of cell structure are directly compatible with remotely sensed and grid (raster) GIS data. In terms of vertical extent, each cell may be given a depth, or be discretisised into a number of overlaying cells (a column). For each cell the water discharge to neighboring cells is calculated according to the active hydrological processes. The flow distribution inside the catchment is mapped and even if the processes are estimated as a continuum, the stored results are discretisised into cells (Abbot and Refsgaard 1996).

Distributed hydrological modeling is advantageous in terms of studying effects of land use changes because its distributed nature enables the simulation and estimation of spatial variations, characteristics and changes inside a catchment. It not only provides a single outlet discharge, but multiple outputs on a temporally and spatially distributed basis. The disadvantages with the distributed type of modeling are the large amount of data and the heavy computational requirements also a large number of parameters and variables that have to be evaluated. The effect of scale choice (cell size) is also an uncertainty (Beven and Moore 1993)

A stochastic model makes use of random elements drawn from statistically possible distributions meaning the simulations will yield different results when repeated with the same input data. With most stochastic models, the common approach is to conduct several simulations (the Monte Carlo technique) and produce average estimates with specified confidence intervals.

Lastly, with the incorporation of computers and high quality spatial data, interest has shifted from lumped models toward spatially distributed models, where water movement within a drainage basin can be simulated. Spatially distributed hydrological modeling can be applied to movement of pollutants, simulation of nutrient leakage in agricultural lands, impact of vegetation and land use change on hydrological regimes and lastly, the impact of land surface (e.g. agriculture and forestry) management practices on hydrological regimes (Olsson and Pilesjo,2002).

2.2.3 Support for the simulation/ prediction methods utilized in this study

The SWAT model is a long term yield model and is able to simulate long term impacts and this is useful for management and decision making for a river basin. In the case of this study, the predicted impacts will help in making decisions that will ensure the maintenance of adequate environmental flows in the basin.

The SWAT model is best suited this study because of its ability to simulate impacts over a long time and because it is a semi-distributed, physically based model requiring inputs of weather, soil properties, topography, vegetation and land management practices in the watershed. It is also computationally efficient and can be used to simulate processes over very large areas like the Mara Basin without excessive data and investment of time and resources. The Mara Basin covers an area of 13,750 km^2 (Mati et al., 2005) and the model with sufficient data inputs can accurately simulate the impacts of different practices within the basin.

The model is suitable for this study because watersheds with no previous gauge measurement data can be modeled. The Mara Basin is a gauged river basin with four

gauging stations but there is a substantial amount of data missing for the gauging stations for a number of years. Moreover, in the process of simulation, gauging stations data is required at the outlets of the sub basins or sub watersheds in simulations in order to calibrate the model by fitting the simulated or predicted values to the observed or measured values. Without few of these gauging stations available in the basin, the SWAT model is a good choice for the simulation of water output for this study.

Setegn (2008) applied SWAT2005 model to Lake Tana Basin to model the hydrological water balance with an objective of testing the performance and feasibility of the model in streamflow prediction in the basin. The different algorithms used were effective in minimizing the differences between observed and simulated flow in the Lake Tana Basin. Despite different uncertainties, the SWAT model produced good simulation results for daily and monthly time steps. Jayakrishnan (2005) used SWAT to model the hydrology of Sondu River basin in Western Kenya. This study demonstrated that the application of detailed hydrologic models, developed and studied widely in the United States, to African river basins is possible given proper and adequate input data collection to improve model parameter calibration and simulation results.

2.2.4 Description of the Model-Soil and Water Assessment Tool (SWAT)

The Soil and Water Assessment Tool (SWAT) is a hydrological model that can be applied at the river basin, or watershed scale. Developed for the United States Department of Agriculture (USDA) Agricultural research Service (ARS) for the purpose of simulation of impact of land management practices on water, sediment and

agrochemical yields in large watersheds with varying soils, land use and agricultural conditions over extended time periods (Neitsch et al., 2005).

The SWAT model (Soil and Water Assessment Tool) is used to simulate processes affecting water quantity, sediment and nutrient loads in a catchment (Abbaspour et al., 2007). Arnold et al., (1998) defines SWAT as a semi-distributed, time continuous simulator operating on a daily time step. It is developed for assessment of the impact of management and climate on water supplies, sediment, and agricultural chemical yields in sub-basins and larger basins.

It allows simulation of a high level of spatial detail by dividing the watershed into a large number of sub-watersheds which are characterized by one or more hydrological response units (HRUs). Each HRU corresponds to a particular combination of soil and land use within the sub-basin, only soil types and land use classes exceeding the user-defined threshold area are considered to set the overlay combination. The program is provided with an interface in Arc View GIS (Di Luzio et al., 2002) for the definition of watershed hydrologic features and storage, as well as the organization and manipulation of the related spatial and tabular data.

2.2.4.1 Hydrological Component of SWAT

The simulation of the hydrology of a water-shed is done in two separate divisions. One is the land phase of the hydrological cycle that controls the amount of water, sediment, nutrient and pesticide loadings to the main channel in each sub basin. Hydrological components simulated in land phase of the hydrological cycle are canopy storage, infiltration, redistribution, evapotranspiration, lateral subsurface flow, surface runoff,

ponds, tributary channels and return flow. The second division is routing phase of the

hydrologic cycle that can be defined as the movement of water, sediments, nutrients and

organic chemicals through the channel network of the watershed to the outlet. In the land

phase of hydrological cycle, SWAT simulates the hydrological cycle based on the water

balance equation.

$$SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{qw}) \qquad (2\text{-}2)$$

Where; SW_t is the final soil water content (mm), SW_o is the initial soil water content on

day i (mm), t is the time (days), R_{day} is the amount of precipitation on day i (mm), Q_{surf} is

the amount of surface runoff on day i (mm), E_a is the amount of evapotranspiration on

day i (mm), W_{seep} is the amount of water entering the vadose zone from the soil profile on

day i (mm), and Q_{gw} is the amount of return flow on day i (mm).

 More detailed descriptions of the different model components are listed in (Arnold et al.,

1998, Neitsch et al., 2005).

Surface runoff or overland flow generally occurs on a slope whenever the rate of

precipitation exceeds the rate of infiltration. SWAT uses two methods for estimating the

surface runoff: the SCS curve number method (USDA-SCS, 1972) and the Green &

Ampt infiltration method (Green and Ampt, 1911). The Green & Ampt infiltration

method makes use of sub-daily precipitation and calculates the infiltration as a function

of the wetting front matric potential and effective hydraulic conductivity. Using daily or sub daily rainfall data, the SWAT model simulates surface runoff volumes peak runoff rates for each HRU. In this study, the SCS curve number method is used to estimate surface runoff because of the unavailability of sub-daily rainfall data for Green & Ampt method.

The SCS curve number equation is (USDA-SCS, 1972):

$$Q_{surf} = \frac{(R_{day}-I_a)^2}{(R_{day}-I_a+S)} \qquad (2\text{-}3)$$

In which, Q_{surf} is the accumulated runoff or rainfall excess (mm), R_{day} is the rainfall depth for the day (mm), S is the retention parameter (mm) which varies spatially due to changes in soils, land use, management and slope and varies temporally due to changes in soil moisture content. The retention parameter is defined as:

$$S = 25.4 \left(\frac{1000}{CN} - 10\right) \qquad (2\text{-}4)$$

Where CN is the curve number for the day. The initial abstractions I_a, is commonly approximated as *0.2S* and equation 3 becomes:

$$Q_{surf} = \frac{(R_{day}-0.2S)^2}{(R_{day}+0.8S)} \qquad (2\text{-}5)$$

The SWAT2005 version includes two methods for calculating the retention parameter; the first one is retention parameter varies with soil profile water content and the second method is the retention parameter varies with accumulated plant evapotranspiration. The soil moisture method (equation 2-5) over-estimates runoff in shallow soils. But calculating daily CN as a function of plant evapotranspiration, the value is less dependent on soil storage and more dependent on antecedent climate (Setegn, 2008).

$$S = S_{max}\left(1 - \frac{SW}{[SW + \exp(w_1 - w_2.SW)]}\right)$$
(2-6)

In which S is the retention parameter for a given day (mm), S_{max} is the maximum value that the retention parameter can have on any given day (mm), SW is the soil water content of the entire profile excluding the amount of water held in the profile at wilting point (mm), and w_1 and w_2 are shape coefficients. The maximum retention parameter value, S_{max}, is calculated by solving equation 3. Using CN_1.

$$S_{max} = 25.4\left(\frac{100}{CN_1}\right) - 10$$
(2-7)

When the retention parameter varies with plant evapotranspiration, the following equation is used to update the retention parameter at the end of every day:

$$S = S_{prev} + E_0 * exp = \left(\frac{-\text{cncoef}- S_{prev}}{S_{max}}\right) - R_{day} - Q_{surf} \qquad (2\text{-}8)$$

In which S_{prev} is the retention parameter for the previous day (mm), E_o is the potential evapotranspiration for the day (mm/day), *cncoef* is the weighting coefficient used to calculate the retention coefficient for daily curve number calculations dependent on plant evapotranspiration, S_{max} is the maxi-mum value the retention parameter can achieve on any given day (mm), R_{day} is the rainfall depth for the day (mm), and Q_{surf} is the surface runoff (mm). The initial value of the retention parameter is defined as $S=0.9*S_{max}$.

The SCS curve number is a function of the soil's permeability, land use and antecedent soil water conditions. SCS defines three antecedent moisture conditions: I – dry (wilting point), II – average moisture, and III – wet (field capacity). The moisture condition I curve number is the lowest value the daily curve number can assume in dry conditions. The curve numbers for moisture conditions I and III are calculated with equations 2-9 and 2-10.

$$CN_1 = CN_2 - \left(\frac{20.(100-CN_2)}{(100-CN_2 + \exp[2.533-0.0636.(100-CN_2)])}\right) \qquad (2\text{-}9)$$

$$CN_3 = CN_2 \cdot exp[0.00673 \cdot (100 - CN_2)] \qquad (2\text{-}10)$$

Typical curve numbers for moisture condition II are listed in various tables (Neitsch et al., 2005). The values are appropriate for a 5 % slope. Williams (1995) developed an equation to adjust the curve number to a different slope:

$$CN_{2S} = \frac{(CN_3 - CN_2)}{3} . [1 - 2 . \exp(-13.86 . slp)] + CN_2 \qquad \text{(2-11)}$$

In which CN_1 is the moisture condition I curve number, CN_2 is the moisture condition II curve number for the default 5 % slope, CN_3 is the moisture condition III curve number for the default 5 % slope, CN_{2S} is the moisture condition II curve number adjusted for slope and *slp* is the average percent slope of the sub basin.

SWAT calculates the peak runoff rate with a modified rational method. There are many methods that are developed to estimate potential evapotranspiration (PET). Three methods are incorporated into SWAT: the Penman-Monteith method (Monteith, 1965), the Priestley-Taylor method (Priestley and Taylor, 1972) and the Hargreaves method (Hargreaves et al., 1985). For this study the Hargreaves method was used due to lack of weather data such as wind speed, humidity and sunshine hours.

2.2.4.2 Routing Phase of the hydrological cycle
Once SWAT determines the loadings of water, sediment, nutrients and pesticides to the main channel, the loadings are routed through the stream network of the watershed. Routing in the channel is divided into four components: water, sediment, nutrients and

organic chemicals.

In flood routing, flow is routed through the channel and possible losses are taken into consideration such as evaporation and transmission through the channel bed. Another potential loss is the abstraction or removal of water from the channel for human or agricultural use. Flow may be supplemented by rain falling directly on the channel or addition of water from a point source discharge. Flow through the channel is routed using a variable storage coefficient method developed by Williams (1969) or the Muskingum routing method.

2.3 Effects of land management practices on hydrology and soil erosion

Population growth-induced agricultural intensification is taking place at a high rate in the Mara Basin. This has had an influence on the land use and land management practices in the area especially in the Mau forest area which is the catchment of the Mara basin. Management of these agricultural lands is crucial to the hydrology of an area. Deforestation and change and loss of species are constantly taking place.

Different studies on these effects of management practices have been carried out with a focus on the effects on the hydrology and resultant soil erosion of the study area. Juckem et al. (2008) studied the effects of climate and land management change on streamflow in the driftless area of Wisconsin. Land management factors were evaluated by measuring infiltration rates in three landscape settings: ridges, hillslopes, and valleys. Effects of land management practices were evaluated by categorizing the infiltration measurement site according to the relative intensity of the current land use. Examples of relatively intense land use included: cultivated agricultural fields, pastures, and gullies;

33

relatively less intense land use included: fallow fields, non-grazed woodlands, and grassy agricultural waterways. Changes in infiltration were assumed to correspond with gradual increases in relatively less intensive land uses (e.g., forested acreage) and decreases in relatively more intensive uses (e.g., pasture and cropland) in the Kickapoo and Coon Creek Watersheds. At the site-scale, higher infiltration rates were recorded under less intensive land use, For example, at one hillslope site, infiltration rates were more than twice as high for a forested plot as compared with a pastured plot less than a distance of 15 m away and at nearly the same elevation.

The study revealed that land management practices can influence how precipitation is partitioned into runoff or recharge. The results suggested the potential for changes in land management practices, specifically a transition to generally less intensive agricultural practices and cessation of grazing on hillslopes, to preferentially increase total infiltration and decrease runoff for a given precipitation event. As suggested by Potter (1991), Gebert and Krug (1996), and Krug (1996), among others, improved agricultural land management practices are likely an important factor responsible for decreased storm flows and increased baseflows.

Haddleland (2006) studied the hydrologic effects of land use changes, dams and irrigation in North America and Asia over a period of 300 years using a macroscale hydrologic model. The simulation results showed that the expansion of croplands over this period of time has resulted in the increase of runoff volume in North America and Asia by 2.5 % and 6 % respectively. Reservoir operations and irrigation practices in the western part of USA and Mexico resulted in a 25 % decrease in streamflow in June and a 9 % decrease in annual runoff volumes reaching the Pacific Ocean. In the part of

South East Asia draining into the Pacific Ocean, land use changes have caused increase in runoff volumes throughout the year, with an average annual increase in runoff of 12%. Irrigation increases evapotranspiration and decreases runoff, and in North America irrigation water use to some extent cancels out the simulated increase in runoff caused by land use changes over the past 300 years. In the Asian region, simulated irrigation water requirements for current conditions are slightly lower than the simulated increase in runoff caused by land use changes. However, simulated irrigation water use is only about half the amount of irrigation water requirements.

These aforementioned studies provide evidence of the different effects of land management on the hydrology in the respective study areas. In the Mara Basin, the land management practices-or lack of them have an effect on the hydrology of the Mara Basin and the application of the SWAT model is expected to give an idea of the direct and indirect effects of these practices.

2.4 Effect of Land use and Land cover and Climate Change dynamics on hydrology

Climate change can be defined as any change in climate over time, whether man made or natural. Signs of climate change observed include sea level rise, changes in precipitation, rise in surface temperature and decreased snow cover (IPPC, 2001). According to Phoon et al. (2004), the climate system affects all aspects of the hydrological cycle including water supply and demand and ecosystems. With regards to hydrology, climate change will directly affect precipitation and evaporation. Higher temperature accelerates the hydrological cycle which alters precipitation amounts, magnitude and timing of runoff including intensity and frequency of floods and droughts.

By increasing the energy available for evaporation, higher temperatures would alter soil moisture and infiltration rates. With all those changes above, river flows and groundwater recharge would be altered, which then affect the catchment water balance. The assessment of the likely impacts of future global warming on water resources system is crucial for the sustainable management of water resources. The study of the impacts of climate change also helps to provide scientific information to enhance public awareness and form a sound base for political decisions to be made on the mitigation of the impacts global warming.

Climate change caused by the increase in the concentration of greenhouse gases poses global challenges for water planners as it affects water variability and availability worldwide. In view of the importance of water in that region of the Mara and the greater Lake Victoria basin, assessing the potential climate change impacts is very vital. For a better prediction in the climate change impact assessment, factors such as downscaling and variation between climate scenarios should be taken into consideration. Downscaling of climate scenarios generated from the general circulation models is necessary for hydrological studies because of the mismatch in spatial and temporal resolutions (e.g. Arnell, 1996; Russo & Zack, 1997; Robock et al., 1993).

Dynamics of land cover has been proven to have significant effects on the hydrology of a river basin. Climate change is also found to be a significant factor in such studies. Bewket and Sterk (2005) studied the impacts of land cover dynamics on the hydrology of the Chemoga watershed in the Blue Nile basin in Ethiopia. The land cover dynamics between 1957 and 1998. Historical stream flow patterns were analyzed with reference to historical dynamics in the land cover of the study watershed. The results showed in the

Chemoga watershed, land cover/land use changes have occurred since the 1950s. The major change that occurred was the increase of the cultivated area at the expense of the open grazing area. The forest cover also showed a slight increase, which was mainly the result of the increase of areas under eucalypt plantations. These dynamics in land cover/land use appear to have affected the stream flow of the watershed. Over the period between 1960 and 1999, the total stream flow decreased at a rate of 1.7 mm per annum, whereas rainfall decreased only at a rate of 0.29 mm per annum. The decrease in the stream flow was caused by the slight decrease in the rainfall, increased transpiration losses because of the increased tree cover and a decreased contribution from the base flow, as revealed by the analysis of extreme low flows. Both the increased transpiration losses and the decline in base flow were associated with changes in the land cover of the watershed and/or watershed degradation. The other factor contributing to the decreased stream flow, particularly during dry season, was the increased water abstraction expected to result from the increased human and livestock populations in the watershed. Climate change is also a factor of fundamental importance, with reduced rainfall there is almost a guarantee there will be an impact on the hydrology because of a reduced amount of water input into the system.

2.5 Related Studies

There are a number of other studies carried out in this area related to land use and land cover change and the hydrology of the Mara River Basin. A study was carried out by Mutie et al. (2006) to evaluate effects of land use change within the basin on the stream flow of the Mara River using the United States Geological Survey geospatial stream flow

37

hydrological model. This study revealed significant land use and land cover changes within the basin attributed to forest clearing for timber and settlement, and also due to conversion of natural land to agriculture. These changes in land use were impacting the hydrology of the Mara River causing erratic flows the simulation results from the streamflow model indicated that compared to the 1973 land cover dataset, the 2000 land cover dataset produced more streamflow even for minimal amounts of rainfall that did not produce streamflow in the 1973 dataset. The 2000 hydrograph peaked higher, faster and receded more sharply than the hydrograph obtained from the 1973 dataset a possible implication that less surface runoff is going to the baseflow.

Setegn (2008) applied SWAT2005 model to Lake Tana Basin to model the hydrological water balance with an objective of testing the performance and feasibility of the model in streamflow prediction in the basin. The different algorithms used were effective in minimizing the differences between observed and simulated flow in the Lake Tana Basin. Jayakrishnan et al. (2005) used SWAT to model the hydrology of Sondu River basin in Western Kenya. This study demonstrated that the application of detailed hydrologic models, developed and studied widely in the United States, to African river basins is possible given proper and adequate input data collection to improve model parameter calibration and simulation results.

Jacobs et al. (2007) conducted a study in the Upper Tana River Basin using the SWAT model to predict impacts of land use and to evaluate reforestation scenarios in an area where unregulated deforestation and expansion of cultivation practices onto marginal soils has resulted in significant siltation of the Masinga dam reservoir, reduced ecosystem function, and more erratic downstream flows. The SWAT hydrological model was paired

with an economic model which was set up to determine the opportunity costs associated with the reforestation and economic incentives or green payments used to induce the users of the upper catchment to reforest their land and implement best management practices. The analysis found that reforestation would decrease sediment loading in the Masinga Reservoir by 7 percent which translates to about 250,000 tons of sediment per year with the economic model estimating a compensation of about $33 for each ton of sediment retained in the fields of the upper catchment users. The increased operational potential of the Masinga Dam would enable it to better supply services such as hydroelectric power, water for irrigation and urban consumers and this would be a result of restoration activities in the upper catchment areas. Additional benefits of reforestation would include improved water routing and channeling following large rainfall events and hence a reduced likelihood of flooding. The findings of this research can assist environmental policy implementation by the Kenyan government that will foster improved environmental results and enable schemes such as the payment of environmental services to be integrated into environmental management.

Van Loon and Droogers (2007) used SWAT in Kitui, Kenya to demonstrate how the physical based component of Integrated Water Management Support Methodologies (the SWAT tool), can be used to support water managers and policy makers on relatively small reservoirs in a developing country. The Kitui area was selected as it presents a typical case for a developing country with alternating wet and dry periods within one year and small-scale local human interventions on the water resource.

One of the conclusions from this demonstration case is that the strength of the SWAT model lies in its completeness and the high physical detail of the model. All components

of the water balance are modeled in detail and the results can be analyzed at all temporal and spatial scales from day to year and from Hydrological Response Unit (HRU) to the entire basin. This proves very important because knowledge on the hydrological system, for example water storage in different components of the system, the fluxes between these components and the available water for human use, are the basis of proper water management. Overall, it can be concluded that SWAT is a very strong tool to support water managers and policy makers as physical processes and human interventions can be analyzed in great detail. The only requirement to an actual implementation of SWAT for Kitui (and other areas) is sufficient data and time to setup the model. If this can be achieved SWAT is a key component of Integrated Water Management Support Methodologies.

2.6 Land use/ land cover Change and hydrology of the Mara Basin

2.6.1 Land use/ land cover change in the Mara Basin

The Mara basin is characterized by a lot of different land uses which result from the different activities carried out within the basin. The Mara Basin consists of mainly closed and open forests, tea plantations in the upper slopes of the Mau Escarpment, agricultural land, shrublands and grasslands used for livestock and game grazing or as game reserves, savannah grasslands which comprise shrub grasslands and wetlands.

According to a study carried out by Mati et al., (2005) based on analysis of Landsat imagery, the different land cover in the Mara River Basin between the year 1986-2000 changed significantly in terms of spatial extent. The Mara Basin is mostly a rangeland and in 1986, 69% of the basin area consisting mostly of savannah grassland and

shrubland. By 2000 however, the rangelands had been reduced significantly because of encroachment by agriculture, which on the other hand had increased by 55%. The closed forest area had also reduced by 23% as a result of forest clearance for timber and tea plantations, resulting in an increase in open land by 82%. The wetlands showed a significant increase attributed to sediment build up in the mouth of the river resulting from erosion in the upstream and erratic river flows which have been caused by change in the vegetation cover in terms of deforestation, conversion of rangelands to agriculture and poor soil and water conservation practices within the basin.

According to a study done by Serneels et al., (2001), since the early 1970s, the land surrounding the Masai Mara National Reserve has been steadily converted into agricultural land with large scale wheat farming being a major part of it. From 1975 to 1995, wheat farming in the Loita Plains in the lower Mara Basin has increased by an area of 44,000 ha. The increase in mechanized agriculture and rangelands modification is said to be driven by factors such as land suitability, and economic factors while smallholder agriculture is driven by factors such as changes in demography caused by in and out migration and population growth within the basin (Entwistle et al., 1998). The land use land cover change in the upper part of the basin is brought about by smallholder agriculture while that of the lower basin is brought about mainly by mechanized agriculture in form of wheat and maize farming.

The ability to forecast land use and land cover change and, ultimately, to predict the consequences of change, will depend on our ability to understand the past, current, and future drivers of land-use and land-cover change. In the Mara Basin, these factors as well as other emerging social and political factors may have significant effects on future land

use and cover. Patterns of land use, land-cover change, and land management are shaped by the interaction of economic, environmental, social, political, and technological forces on local to global scales.

2.6.2 River flow and sediment output in the Mara River Basin

Water output or discharge in the Mara Basin has been measured by means of three gauging stations along the river since the year 1963, two of them on the Nyangores and Amala rivers, tributaries of the Mara River and one on the Mara River at Mara Mining site. Sediment data have not been very consistently measured and only recently was this parameter included in the regular measurement regime of the Mara River. The discharge data from the Kenyan side of the basin is managed and collected by the Lake Victoria South Water Resource Management Authority which is the governing body responsible for water resource management in the Mara basin.

Existing river flow data show that recently there has been erratic flow of the river and this has been experienced within the last two decades which have seen sharp peaks and dips in the hydrograph indicating short lived excess flow in the rainy seasons and very low flows in the dry seasons. This has significant implications with regard to the amount and availability of water in the Mara River Basin.

According to Mutie (2005), land use changes have brought about changes in the hydrological regime with sharp increases in flood peaks, attenuation of river hydrographs and a reduced baseflow. Increased sedimentation has had far reaching effects downstream and is choking the wetlands and river channel causing increased flooding. Land use changes in terms of increase in agricultural land, modification and destruction of forests,

have led to increased soil erosion and increased sediment loads in the river which poses a significant threat to the water quality of the river.

In a baseline water quality study of the Mara Basin carried out by Singler and McClain, (2006), sediment levels and water turbidity ranged from 7.1 NTU at the river mouth to 1999.0 NTU at the New Mara site located within the Masai Mara National Reserve .Total suspended sediments (TSS) ranged from 0.02 g/L at the river mouth to 2.79 g/L at the New Mara Bridge.

Sediment loads increase in aquatic systems through erosion from poor agricultural practices, grazing and deforestation, mining activities, construction and dredging. In the Mara Basin there is a lot of small scale farming activity in the upstream which is perceived to be a major contributor of sediment in the Mara River this has had adverse effects downstream and upstream alike with the sedimentation of dams and the increased flooding down in the Mosirori Swamp. Increased sediment load also negatively impacts on aquatic biota by reducing light penetration, reducing suitable habitat, smothering fish fry, clogging their gills and ultimately altering the biodiversity of the system (Singler and McClain, 2006).

3.0 METHODOLOGY

3.1 Land Cover mapping from Satellite Imagery

3.1.1 Ground truthing/ referencing exercise

Data collection for this study included a 5-week field survey in the Mara River Basin, and also extensive research and data analysis outside of this period. The field study was conducted from mid May to late July 2008.

The field study was aimed at collecting land cover/land use information for use in a land cover classification of remote sensor data and a survey of on-farm soil and water management practices in areas adjacent to the river. A survey of man-made reservoirs was also carried out and their locations recorded.

Land use/land cover information was obtained during the field visit as part of a ground truthing /ground referencing exercise and it included observation of different land cover within the Mara River basin with reference to previously obtained remote sensor data which included Landsat TM imagery and aerial photography of the study area. Using a Global Positioning System (GPS) unit, the geographic positions of these features were marked by their coordinates and input into ArcGIS software to generate point GIS layer that were overlaid on the previously obtained land cover data and was used to verify the observations in the field.

A land use/land cover classification scheme was formulated that would accurately and adequately represent the land cover/land use within the Mara River basin. This scheme however follows the basic principles of the USGS Land use/land cover classification system (LULCCS) for use with remote sensor data level classification (Anderson et al., 1976).

The land use/ land cover types present in the basin were first classified according to this scheme shown in the previous chapter and then were reclassified to the modified Land use/land cover classification scheme for the Mara basin which is used for classification of the land cover dataset for input into the SWAT model.

3.1.2 Classification Scheme, Scale and Choice of Imagery

The imagery chosen for the land cover detection exercise was satellite imagery from the Landsat 4/5 Thematic Mapper (TM) sensor which has been in use since 1982. This imagery was suitable for the exercise because Landsat was built for earth observation purposes. Both its spatial resolution of 30m pixel and 7 band radiometric resolution make it suitable for land cover classification and its instantaneous field of view of 15 degrees gives it a swath width of 185 km (Van der Meer et al., 2002) covering a large area in the process, capturing many features within one scene. Landsat imagery has many applications in land use mapping and this includes; classifying land uses, tracking socio-economic impacts on land use, categorizing land capabilities among other uses.

The Landsat imagery that was required was identified and requested from the United States Geological Survey (USGS) via their Global Visualization (GLOVIS) website and downloaded for use. Landsat imagery is obtained using a system of paths and rows across the globe where the Landsat sensor passes over. The imagery used for the land cover classification was 2 images of Path 169, Row 61 and Path 169, Row 60 from the 5th of September 2008. These images were obtained 3 months after the ground truthing exercise in the study area which was the most viable option due to constraints in the availability of

cloud free Landsat imagery obtained by the TM sensor. The process of land use/land cover classification is lengthy and involves a number of steps.

3.1.3 Data preparation

This is a very important stage when dealing with remote sensor data because it comes in a very raw form which had to be prepared and processed for the specific needs and requirements of the study. Failure to sort and prepare the data for the procedures to be undertaken normally results in unnecessary mistakes that may be time consuming and may also significantly reduce the quality or damage the data to be processed.

3.1.3.1 Image Stacking

The Landsat images once downloaded were unzipped and they consisted of the constituent spectral bands that made up the satellite image. These constituent bands are stored in a Tagged Image File Format (TIFF) format which are not as useful individually as when stacked together to form a single image. The necessary bands were identified, which were band 1 to band 5 and band 7. Band 6 which is the thermal band, was excluded from the image when the stacking is done because it is the thermal infrared band which ranges from 10.4-12.5μm on EMR scale and useful in measuring the amount of infrared radiant flux (heat) emitted from surfaces which enables the detection of geothermal activity, vegetation stress and measurement of soil moisture which is not necessary for this mapping exercise. The 6 bands (1-5 and 7) were then stacked together to form the image and it was further prepared for processing.

3.1.3.2 Image Reprojection

The images were then reprojected using the reproject tool in ERDAS IMAGINE (ERDAS/Leica Geosystems, 2002) to the Universal Transverse Mercartor Zone 37 South projection which is the projected coordinate system for the study area. This was necessary to have all the images in a uniform projection which would prevent problems of pixel shifting and misplaced features. This was also essential for use with the ground referencing data which was a GPS point GIS layer that required precise positioning in order to make accurate references on the land cover types found in that particular area of the image to be classified.

3.1.3.3 Image Subsetting

Image subsetting was carried out and is the process of reducing the whole image to a smaller more workable area of interest (AOI) to reduce the computation time of the processes and eliminate the production of excess data that unnecessarily lengthens the classification process. This was carried out using the subset tool found in ERDAS IMAGINE which enables to choose an area to subset and in this case the subset area was the Mara River basin GIS layer delineated from the SRTM digital elevation model (DEM) of the study area. This was further subset into the Amala and Nyangores watersheds that were used in the modeling exercise.
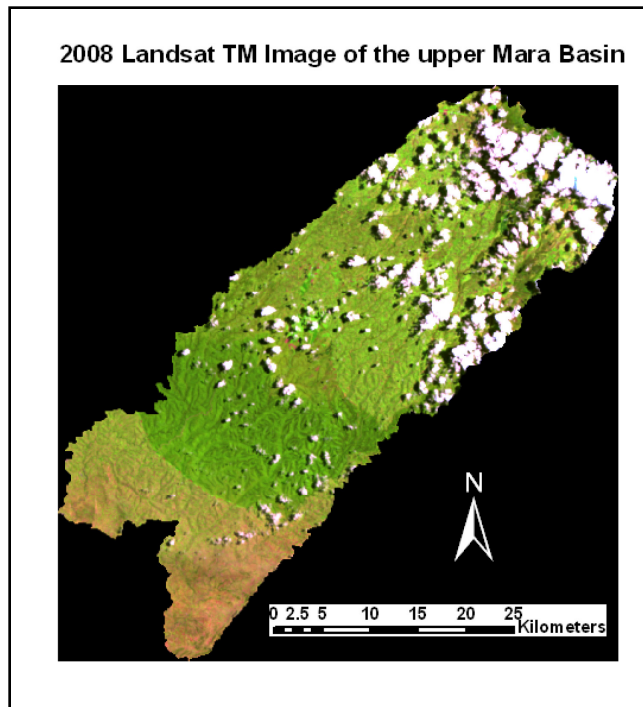
Figure 3-1. 2008 Landsat TM Image of the Upper Mara Basin

3.1.4 Atmospheric Correction

Atmospheric correction was deemed necessary for this image and ATCOR 2 was used for the atmospheric correction of the Landsat imagery The ATmospheric CORrection (ATCOR) module in ERDAS IMAGINE consists of ATCOR 2 and ATCOR 3 (Jensen, 2005) used for flat and rugged terrain, respectively. The atmospheric correction algorithm employs the MODTRAN 4+ (Alder-Golden et al., 1999) radiative transfer code to calculate look-up tables (LUT) of the atmospheric correction functions (path radiance, atmospheric transmittance, direct and diffuse solar flux) that depend on scan angle, relative azimuth angle between scan line and solar azimuth, and terrain elevation.

Atmospheric characteristic information was provided and the correction algorithm was used to compute the required parameters essential for characterizing the scattering and

absorption characteristics of the atmosphere on a specific date. The image radiance was then inverted to scaled surface reflectance using the atmospheric characteristics.

The atmospheric correction process significantly reduced the cloud and haze from the image. To reduce the anomalies in the brightness values in the image that may have resulted from the process, it was filtered by means of a spatial filter.

3.1.4.1 Spatial Filtering

Spatial frequency is a parameter which is characteristic of remotely sensed imagery and is defined as the number of changes in brightness value per unit distance for any particular part of an image (Jensen, 2005). If there are few changes in the brightness values, this is known as a low frequency area and if the brightness values change drastically over short distances, it's known as an area of high frequency detail. Since spatial frequency describes the brightness over a spatial region, extraction of quantitative spatial information is done by considering the local (neighboring) pixel brightness values than just a single pixel value (Jensen, 2005). Spatial frequency may be enhanced or subdued by spatial convolution which makes use of convolution masks.

Low frequency or low-pass filters are used to de-emphasize or reduce the high frequency detail where a particular input pixel value BV_{in}, is evaluated together with the pixels surrounding the input pixel and outputs a new brightness value, BV_{out}, that is the new mean of this convolution (Jensen, 2005). The atmospherically corrected imagery was smoothed by means of a low-pass 3 by 3 convolution filter to smoothen out the brightness values in the image and reduce gross differences in the brightness values of the imagery. This was necessary for the next process which involved selection and extraction

of brightness values from pixels of pseudo-invariant features present in both of the overlapping images and testing them for reflectance differences.

3.1.4.2 Test for Reflectance Differences

When atmospheric correction is carried out on two overlapping satellite images, pixels are selected of pseudo-invariant objects whose reflectance values do not change over time and the pixel reflectance values from the two atmospherically corrected overlapping images are compared to see whether there is a significant statistical difference between them. If there is no significant statistical difference between the reflectance values of the pseudo invariant objects this shows that the atmospheric correction was successful. On the other hand, if the statistical tests determine that there is a significant statistical difference between the two images, this means that the atmospheric correction was unsuccessful and has to be repeated and adjusted accordingly until the desired result is obtained where haze and aerosols are significantly reduced without changing the reflectance values of the pseudo-invariant features.

To test for these differences, a paired t-test and a paired samples Wilcoxon test were carried out. The paired t-test is used with the assumption that the data in question follows a normal distribution and tests the null hypothesis that the true difference in the means is equal to zero. The paired samples Wilcoxon test (Wilcoxon, 1945) is a non-parametric statistical hypothesis test for the case of repeated measurements on a single sample and can be used as an alternative to the paired student's t-test, when the population cannot be assumed to be normally distributed. It tests the null hypothesis that the difference between the means is zero or in other words, there is no difference between the means of

the reflectance values of the pixels of pseudo-invariant objects. The tests carried out give a p-value and using this you either reject the null hypothesis or fail to reject it, enabling you to make a decision on whether or not the reflectance values of the different spectral bands you are testing differ from each other.

If they do not differ this suggests that the atmospheric correction worked successfully and the image is ready for spectral enhancement and classification. The pseudo-invariant features were selected from the corrected and uncorrected images and their pixel values were tested for reflectance differences. A result from the t-test and the paired samples Wilcoxon test indicating no significant change between the pixel values allowed continuation to the next step which involved spectral enhancement of the image.

3.1.5 Spectral Enhancement

Spectral enhancement was carried out by the means of two processes which included performing a Kauth-Thomas Tasseled Cap Transformation (Kauth and Thomas, 1976) and production of texture images. According to Jensen (2005), remote sensor data has been used to model and extract vegetation biophysical since the 1960s and this has largely involved the use of vegetation indices which are dimensionless, radiometric measures that provide indication of relative abundance and activity of green vegetation.

The Kauth Thomas Tasseled Cap Transformation is a transformation produced by Kauth and Thomas (1976) for the Landsat Multispectral Scanner which created a four dimensional feature space resulting in four new axes: the soil brightness index (B), greenness vegetation index (G), yellow stuff index (Y), and non-such (N). Crist et al.(1986) derived the visible, near-infrared and middle-infrared coefficients for

transforming Landsat Thematic Mapper imagery into brightness greenness, and wetness variables: The tasseled cap transformation theoretically is a global index and can be used worldwide to disaggregate the amount of soil brightness, vegetation and moisture content in a Landsat Thematic Mapper image (Jensen, 2005).

The derived visible, near-infrared, and middle infrared coefficients for the transformation of Landsat Thematic Mapper imagery into Brightness (B), Greenness (G) and Wetness (W) variables are:

$$B = 0.2909TM1 + 0.2493TM2 + 0.4806TM3 + 0.5568TM4 + 0.4438TM5 + 0.1706TM7 \qquad (3\text{-}1)$$

$$G = -0.2728TM1 - 0.2174TM2 - 0.5508TM3 + 0.7221TM4 + 0.0733TM5 - 0.1648TM7 \qquad (3\text{-}2)$$

$$W = O.1446TM1 + 0.1761TM2 + 0.3322TM3 + 0.3396TM4 - 0.6210TM5 - 0.4186TM7 \qquad (3\text{-}3)$$

The Kauth Thomas transformation was carried out on the Landsat TM images and was used to transform them into the brightness, greenness and wetness variables this was used together with the texture images that were generated for each of the bands in the image and were stacked together with the image as one in order to enhance the spectral and visual quality of the image in the process of signature extraction in the classification process. This resulted in a stacked image containing 15 bands which greatly enhanced the distinction of different features thus speeding up the process of generation of an Area of Interest (AOI) of pixels belonging to specific classes that were exported into a table as training data for classification by means of an expert classifier.

3.1.6 Classification by Means of an Expert Classifier

Liebowitz (1988), described an expert system as the ability to solve problems and efficiently and effectively in a narrow problem area and to perform at the level of an expert. These systems represent the expert's knowledge base as data and rules in the computer. The expert's knowledge base may result from different sources and this includes field experience with natural and man-made features, images, books journals, manuals and articles. This knowledge is processed using different approaches; algorithms using conventional computer programs that contain little more than the basic algorithm for problem solving with all the knowledge usually contained in the program code, heuristic knowledge-based expert systems that incorporate human knowledge and experience to create a knowledge base and uses the most appropriate knowledge to reason through a problem and artificial neural networks that arrive at decisions in an unstructured and non-algorithmic fashion.

According to Huang and Jensen (1997), the knowledge process presents a well-known problem when creating expert systems that is often referred to as the "knowledge acquisition bottleneck." This arises because the process is time consuming requiring a lot of involvement by the expert and also the inability of the experts to explicitly formulate their knowledge in a form that is sufficiently systematic, correct and complete to form a computer application even though they are able to use their knowledge in decision making (Bratko et al.,1989). It is therefore quite difficult to build an expert classifier based on the expert knowledge and because of this effort was directed to machine learning in order to automate the knowledge building process.

Machine learning can be described as the science that involves computer modeling of learning processes and enables a computer to acquire knowledge from existing data or theories using certain inference strategies such as induction or deduction (Huang and Jensen, 1997). In machine learning, the process of inductive learning can be viewed as a heuristic search through a space of symbolic descriptions for plausible general descriptions, or concepts, that explain the given input training data and are useful for predicting new data (Dietterich and Michalski, 1983)

Inductive learning can be explained as an ability to make accurate generalizations from scattered facts and information from the environment using inductive references. An advantage of inductive learning is that it functions well with a few good examples to function as training data which is much easier than having the expert in the particular domain provide complete theories. The procedure of applying this technique to build a knowledge base for a remote sensing image analysis expert system that involves training, decision tree generation and creation of production rules that compose the knowledge base and are used by the expert system for land cover classification.

The training process provides the examples of the concepts or relationship to be learned. In the case of remote sensing image classification, examples are set of training objects, each of which Huang and Jensen (1997) described as;

$$[\text{ attribute } 1, \ldots, \text{ attribute } \_ \, n, \text{ class } \_ \, i \,]$$

The training data is preclassified by the expert based on their expertise and ground reference information.

The learning algorithm generates a decision trees from the training data and are transformed into production rules (Jensen, 2005). A decision tree can be likened to a

54

classifier composed of leaves that correspond to classes, decision nodes that correspond to the attributes of the data being classified, and arcs corresponding to alternative values of these attributes.

A decision tree is a type of multistage classifier that can be applied to a single image or a stack of images and according to Jensen (2005) is one of the best ways to conceptualize the expert classifier. It is made up of a series of binary decisions that are used to determine the correct category for each pixel. The decisions can be based on any available characteristic of the dataset. For example, you may have an elevation image and two different multispectral images collected at different times, and any of those images can contribute to decisions within the same tree. No single decision in the tree performs the complete segmentation of the image into classes. Instead, each decision divides the data into one of two possible classes or groups of classes. A decision tree takes as input an object or situation described by a set of properties, and outputs a yes/no decision. Decision trees therefore represent Boolean functions.

A decision tree evaluates rules and conditions in order to test a hypothesis. It is built by the process of asking a series of questions that are answered by either true (positive) or false (negative) and reaches its decision by performing a number of tests. A manually constructed decision tree would be based on the expert's knowledge of the study area/ features and would result in a decision tree that is "over-fit' for the data and not useful in outside predictions say, another image because it is based on small correlations within the training data that would not necessarily exist in another image.

Decision trees make use of recursive partitioning to predict class memberships of a particular dataset. Recursive partitioning is able to establish underlying relationships

between the feature classes of data that may as well be impossible to do so manually. This can be done based on a given data matrix that may have hundreds or even thousands of rows and columns. The concept behind recursive partitioning is the splitting (partitioning) of the training dataset. Taking the example of a numerical dataset, the features will be called F_n and the possible values for those features will be called C_n. As a result, every question that can be asked can take the form, "Is F_n less than or equal to C_n?". The resulting answer will direct us down the appropriate path, for example, if F_n is less than or equal to C_n, then go left, otherwise, go right. Once these new nodes (children nodes) are created linked to a previous node (the parent node), the process can be repeated for each child node independently using only the observations present in that node, thus the recursive step.

A purity measure is developed to decide which is the best possible split from our choices. One of the measures is that of absolute purity and is the absolute purity of classes represented as a percent purity that we can achieve in the nodes, answering the question, "Based on a particular split, how good a job did we do of separating the two classes away from each other?" This purity measure is calculated for every possible split and the one that gives the highest possible value is chosen.

Stopping criteria is also very important in the process and this involves stopping the recursive partitioning process before it continues until each leaf has only one observation. This may seem as a perfect tree but in actual sense it is over fit for the training data and cannot be used for prediction purposes and is the same case with many of the decision trees constructed by an expert. Stopping criteria therefore may come in different forms such as; a maximum number of nodes in the tree which when reached the process is

halted, a minimum number of observations in a particular node can be set such that if the number of observations in a node are less than or equal to that minimum value, partitioning of that node is not attempted and it becomes a leaf, a threshold for the purity measure can be set so that if the purity value of a node exceeds that of the set threshold partitioning is not attempted. Once the procedure is stopped, we have effectively reached a leaf in our tree and based on the observations that have made it through the tree to that leaf we can assign it a class value.

Decision trees resulting from the classification of large data sets are usually large and too complex to understand, update and maintain. Decision trees are therefore transformed to a form of knowledge representation known as production rules that Jackson (1990) expressed as;

$$P_1, \dots P_m \longrightarrow Q_1, \dots, Q_n$$

Where:

If premises (or conditions) P_1 and … and P_m are true, then perform actions Q_1 and … and Q_n

Each path from the root to a leaf in a decision tree can be translated to a production rule.

$$(\text{Band } 1 >= 18), (\text{Band } 2 < 2.5) \longrightarrow (\text{class} = \text{Shadow}) \tag{3-4}$$

To increase accuracy of the classification process, a number of different works have proposed the use of multiple classifiers (boosting) that give a better result than the use of a single classifier (Pal and Mather, 2001). Boosting uses the same learning algorithm which constantly generates multiple classifiers in an iterative manner and this can be used

to reduce the error in weak learning algorithms that generate classifiers on various distributions over the training data and combines the classifications resulting from the weak learner into a single composite classifier that is significantly more accurate than any single classifier.

This is achieved by performing a cross validation also known as Leave One Out Cross Validation or K-fold cross validation where the original sample or in our case, the training data is partitioned into K subsamples and a single subsample is retained to validate the model and the remaining K-1 subsamples are used as training data. The process is then repeated K times (same as the folds) with each of the K subsamples used only once as the validation data. The K results from the folds or iterations can be combined or averaged to produce a single estimation. An advantage of this method is that all observations are used for both training and validation and each observation for validation only once (McLachlan, 2004). Cross-validation is designed to obtain relatively realistic accuracy estimates using a limited number of reference data samples for both training and accuracy assessment. The idea is to hide these polygons or classes one at a time and use the rest to estimate their values. The difference between the estimated values and the hidden actual data can be summarized using histogram or statistics (Liu X. H., & Liu Y., 2008). This can be carried out iteratively in order to get better results together with progressive refinement of the training data. The resulting classification is then subjected to an accuracy assessment to determine the accuracy of the map.

The pixel spectral and spatial locations were then saved in a table as a .csv file which contained the values from each and every band stacked to form the image. This table was then loaded into the statistical package R (RDCT, 2009) that performs the classification

of the data by means of recursive partitioning. The program R ran a script that performed classification by recursive partitioning and then performed a cross validation which involves hiding these classes obtained one at a time and use the other resultant classes to predict their values statistically (Liu et al., 2008). The iterations of cross validations to be performed can be set and in this case they were set at 5 and 10 times respectively. This was done to increase the accuracy and produced a decision tree which was used to build an expert classifier would use to perform a classification of the image.

The expert classifier was constructed in the program ERDAS IMAGINE 9.3 (ERDAS, 2006) using the Knowledge Engineer Tool. This involved identification of the hypotheses which are the classes identified in the study area; Cloud, Bushland, Cropland, Grassland, Bare soil, Shadow, Water and Forest. Specification of the expert system rules (variables) and conditions was the next step in this study and were based on remote sensing multispectral reflectance characteristics and derivatives including the Kauth Thomas Tasseled Cap transformation and texture bands. This process was made significantly easier by the recursive partitioning process carried out beforehand resulting in the decision tree that guided the building of the expert classifier that was used to classify the image and produce a classified image.

3.1.7 Land Cover Map Accuracy Assessment

Accuracy assessment of the land cover map was important to identify the sources of error and to know the confidence level of the resultant thematic information. The error matrix produced from the recursive partitioning in R provided a summary of the classification accuracy. A second accuracy assessment was carried out and this involved

ground referencing of the resultant land cover map by means of aerial photography that was available for a stretch of the study area along the Mara River and from this ground referencing exercise, an error matrix was generated which was evaluated to determine the accuracy of the map produced.

The error matrix was evaluated by means of Kappa analysis, a discrete multivariate technique that yielded the $\widehat{K}$ statistic (K_{hat} Coefficient of Agreement) estimate to measure the agreement or between the remote sensing-derived classification map and the reference data.

$$\widehat{K} = \frac{N \sum_{i=1}^{k} x_{ii} - \sum_{i=1}^{k} (x_{i+} \times x_{+1})}{N^2 - \sum_{i=1}^{k} (x_{i+} \times x_{+1})} \tag{3-5}$$

Where k is the number of rows (e.g., land cover classes) in the matrix, x_{ii} is the number of observations in row i and column i, and x_{i+} and x_{+1} are the marginal totals for row i and column i, respectively, and N is the total number of observations.

The resultant $\widehat{K}$ value was then evaluated and was found to represent a good agreement between the remote sensing derived map and the reference data. The map was then aggregated and linked to a land use look-up table for use as an input in the SWAT hydrological model.

3.2 Hydrological Modeling

The SWAT (Soil and Water Assessment Tool) model was the ideal choice for use in this study because of various reasons; it is a physically based model that requires specific information about weather, soil properties, topography, vegetation and land management practices which it uses as inputs to simulate the physical processes associated with water movement, nutrient transport, crop growth and sediment movement which is important to achieve the objectives of this study. This enables it to model ungaged watersheds and more importantly, quantify the impact of alternative input data such as changes in land use, land management practices and climate on water quality and quantity. Secondly, it uses readily available data, while more inputs can be used to simulate more specialized processes it is still able to operate on minimum data which is an advantage especially when working in areas with insufficient or unreliable data. Third, the SWAT model is computationally efficient, able to run simulations of very large basins or management practices without consuming large amounts of time and expenses. Lastly, it is a continuous time or a long-term yield model able to simulate long term impacts of land use, land management practices and build up of pollutants (Neitsch et al, 2005). These qualities of the SWAT model will enable the quantification of long term impacts of land use changes, variations in rainfall and air temperature on the hydrology of the Mara Basin.

3.2.1 Data Availability/ Model Input

3.2.1.1 Digital Elevation Model
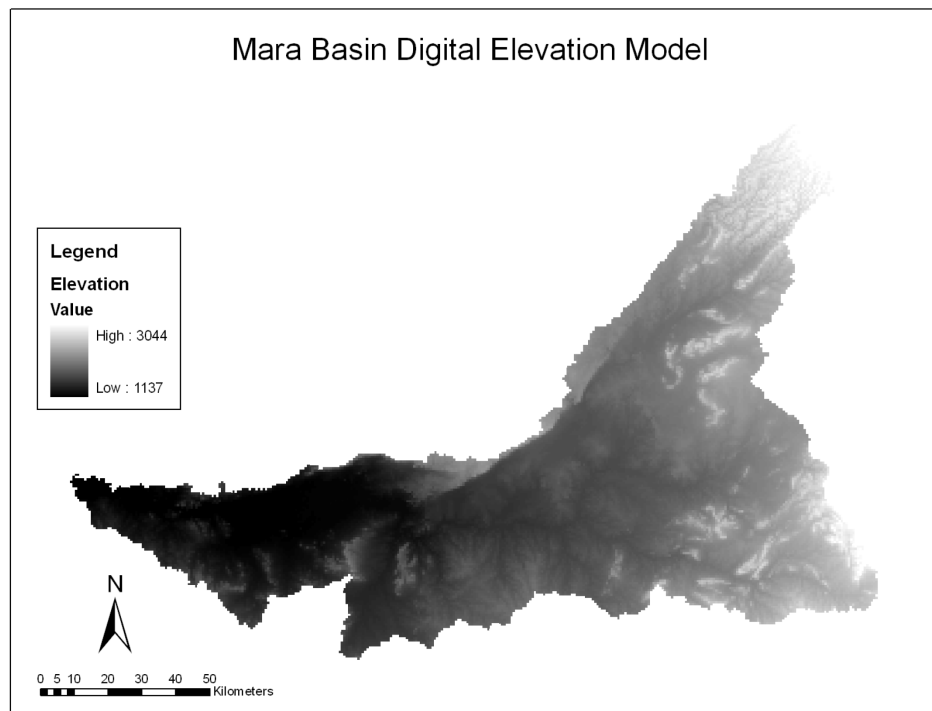
Figure 3-2. Mara Basin Digital Elevation Model

The digital elevation model (DEM) of 90m by 90m resolution for the study area obtained from the Shuttle Radar Topography Mission (SRTM) was used. The DEM gives the elevation of a particular point at a particular spatial resolution and was used in the delineation of the watershed and analysis of the land surface characteristics and drainage patterns.

Terrain preprocessing was performed with ArcSWAT in ArcMap 9.2 (ESRI, 2006) using the digital elevation model to delineate the catchments and drainage line. The steps involved were:

- Filling of sinks
- Flow direction

62

- Flow accumulation

- Stream definition

- Stream segmentation

- Catchment grid delineation
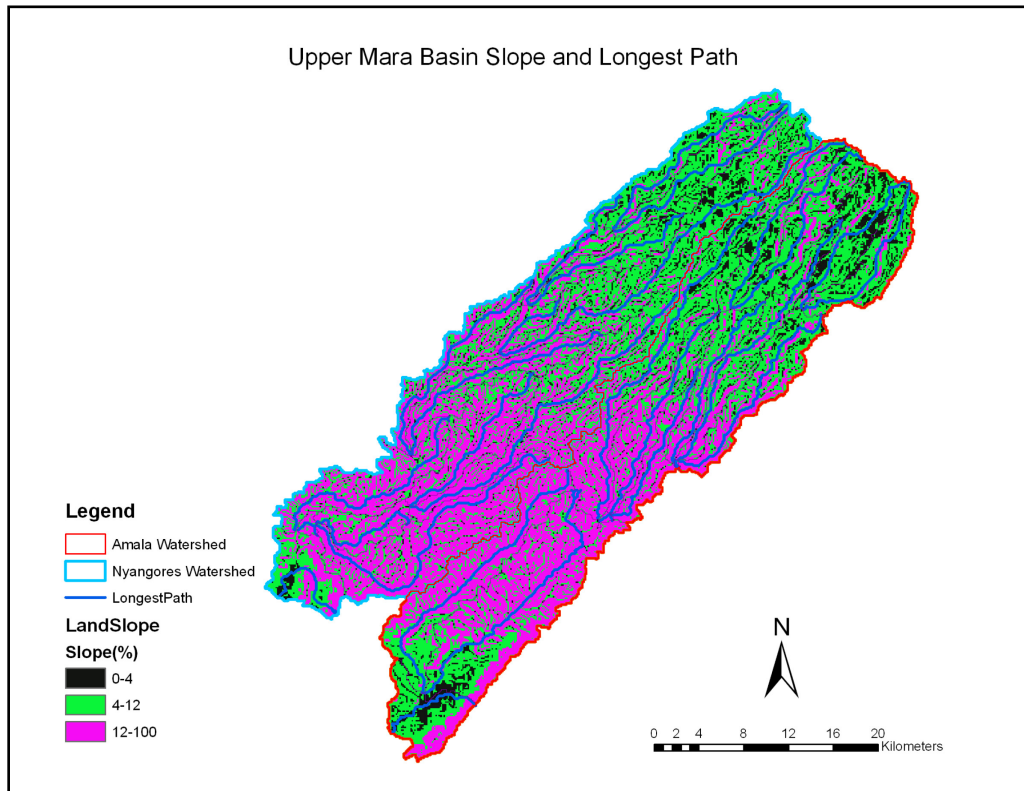
Figure 3-3. Upper Mara Basin Slope and Longest Path

3.2.1.2 Soil Data

Soil data was obtained from the Soil Terrain Database of East Africa (SOTER). GIS layers were obtained and used in the hydrological model as one of the main inputs to the SWAT model which requires soil property data such as the texture, chemical

composition, physical properties, available moisture content, hydraulic conductivity, bulk density and organic carbon content for the different layers of each soil type (Setegn, 2008). A user table specific for the Mara River basin soil layer was appended to the soil table in the SWAT database by using Arc toolbox in ArcGIS since the soil types found in the study area are not included in the US soils database.
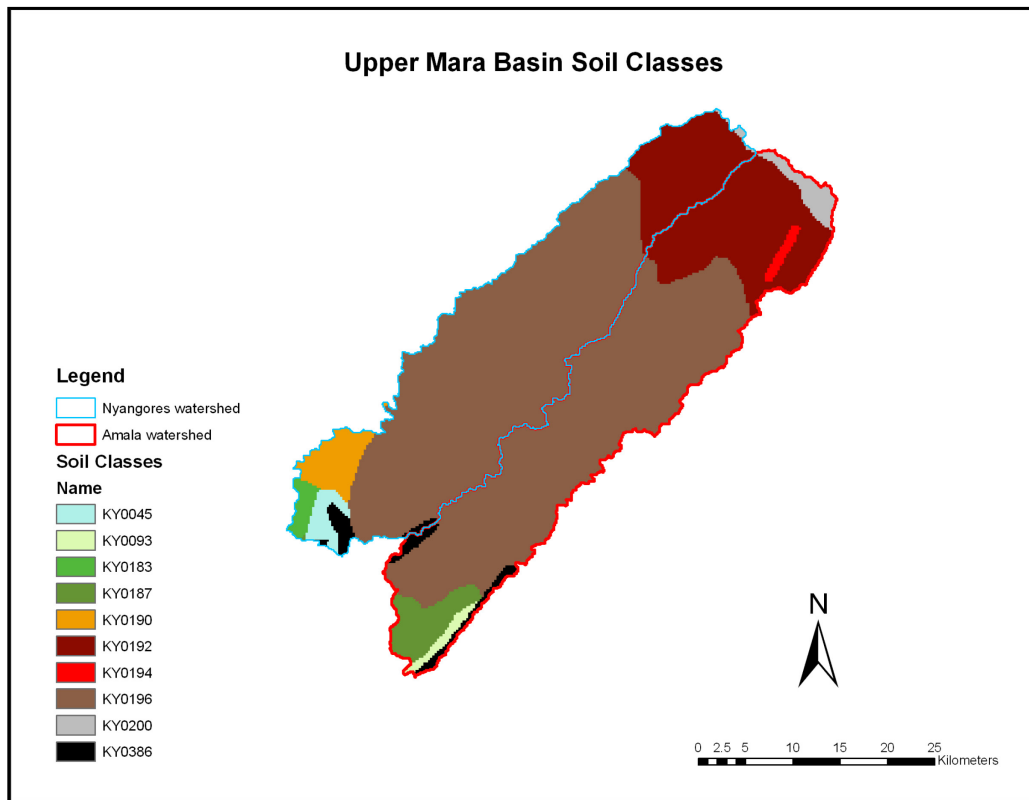
Figure 3-4. Upper Mara Basin Soil Classes

3.2.1.3 Land Use

Land use data for the year 2008 was obtained by analysis of Landsat TM imagery in the process described previously in this chapter resulting in land cover maps of the study

area. These land cover maps were then converted to shapefiles and aggregated to make them easier to input into the model for use in the hydrological modeling exercise. An existing land use dataset for the year 2002 were also used and this was obtained from the FAO-Africover project (FAO, 2003).
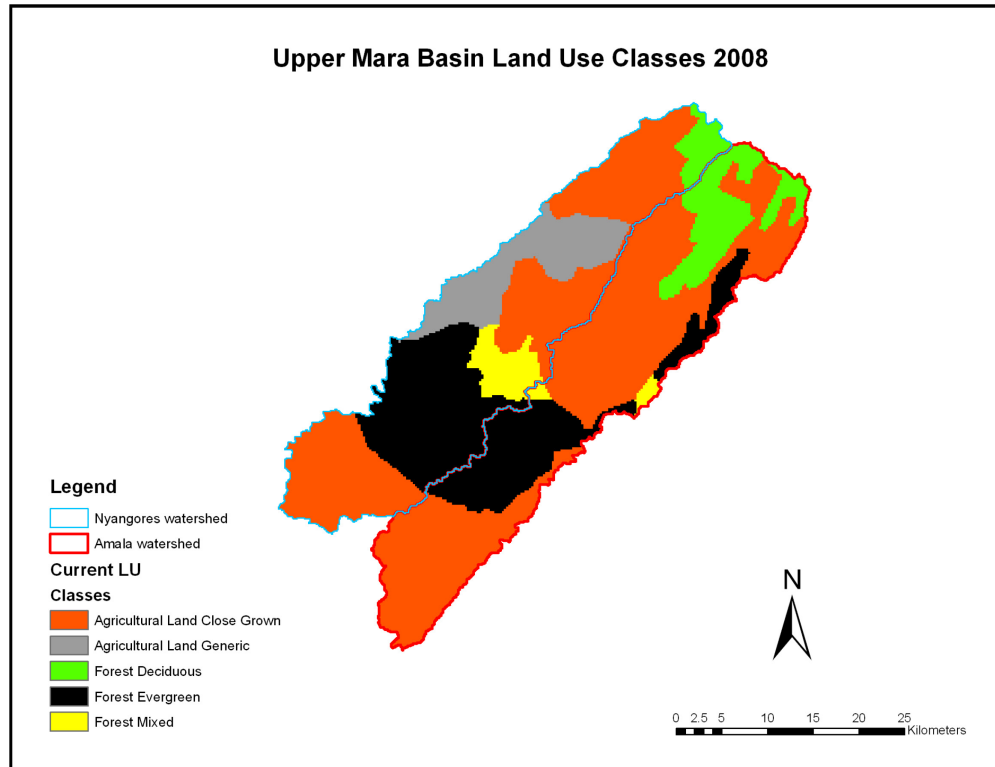
Figure 3-5. Upper Mara Basin Land Use

Land use and management is an important factor affecting different processes in the watershed such as surface runoff, erosion and evapotranspiration. Reclassification of the land use map is done in order to present them in a form that is acceptable in the model and this is the USGS Land use/ Land cover classification scheme for Use with Remote Sensor data level classification (Anderson et al., 1976).

3.2.1.4 Climate Data

Weather data used in the SWAT model consists of daily rainfall, temperature, wind speed, humidity and evapotranspiration data. The data can be observed or measured data or can be generated by a weather generator model given the appropriate descriptive statistics. In this study, the weather variables used were the daily precipitation values obtained from the Bomet Water Supply Office Station located at Bomet Town and Kiptunga Forest Station located in Elburgon District, minimum and maximum air temperature values for the period of 1996-2003 obtained from the Kericho Hail Research and Narok Meteorological weather stations. These data were obtained from the Ministry of Water Resources of Kenya and the Lake Victoria South Water Resource Management Authority in Kenya. The available weather data is of high importance as it determines what methods and equations the SWAT model uses in the calculation of different climatic and hydrological parameters in the simulation.
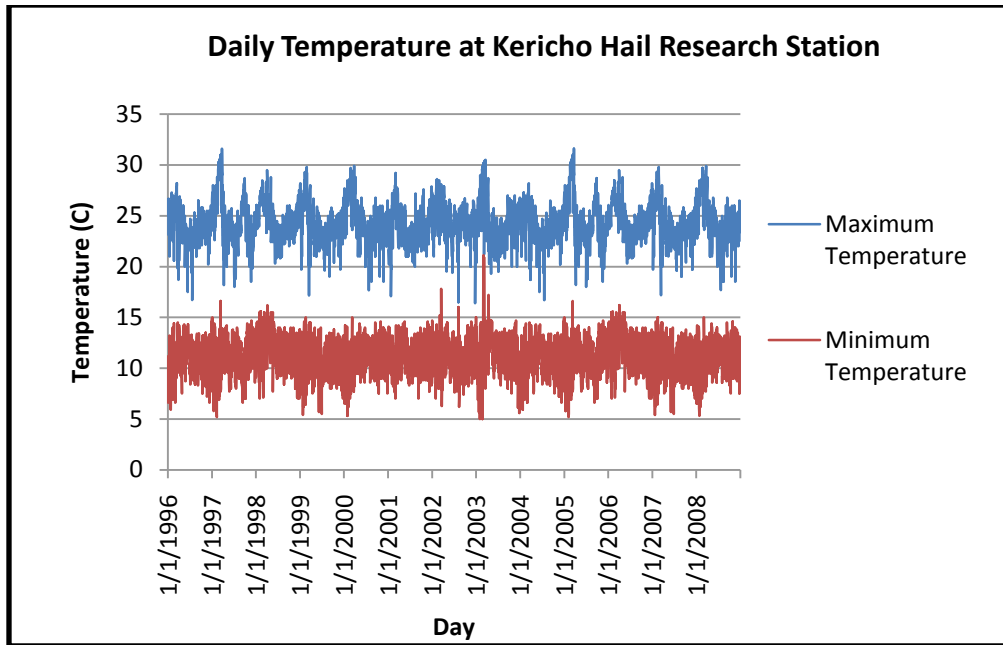
Figure 3-6. Daily temperature at the Kericho Hail Research Station
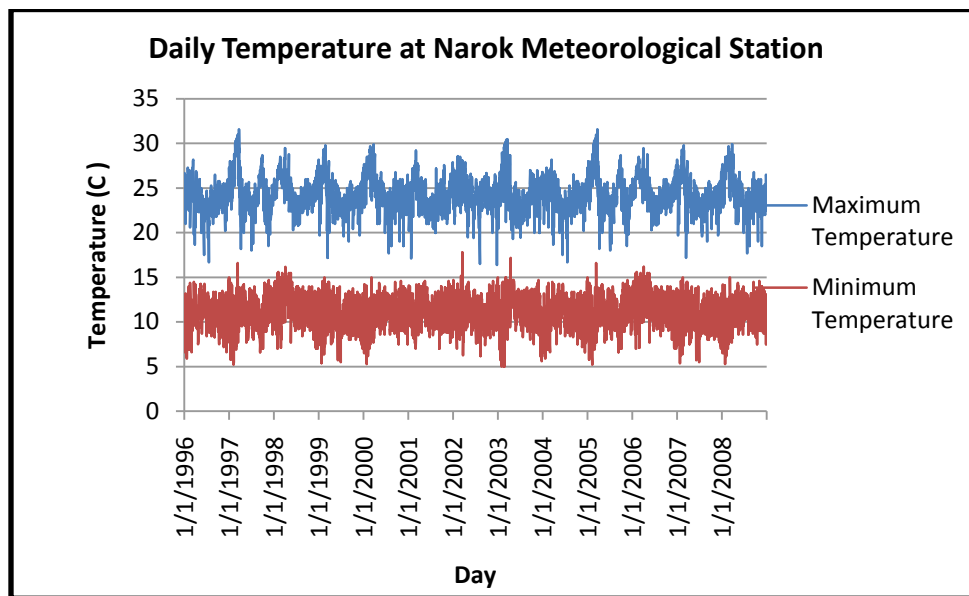
Figure 3-7. Daily temperature at the Narok Meteorological Station
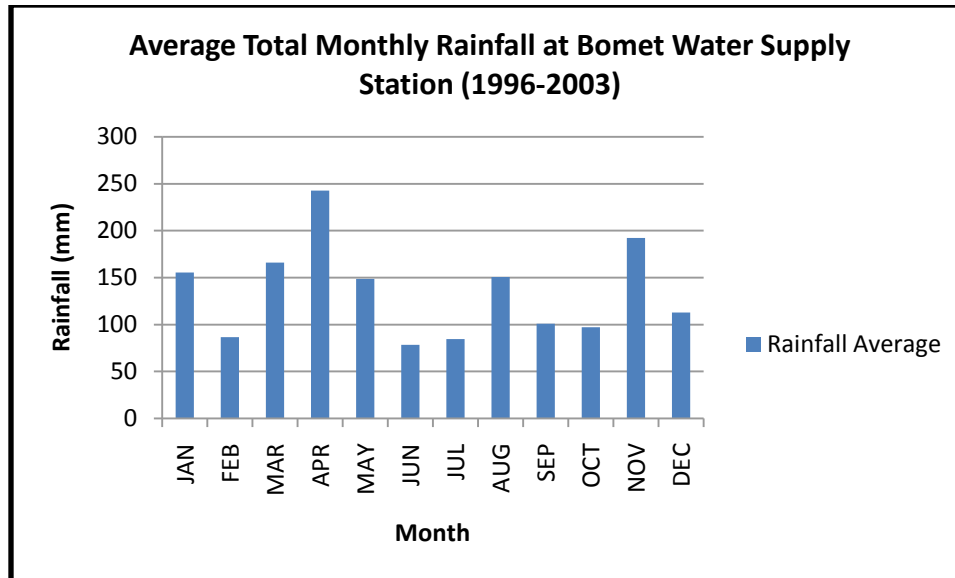
Figure 3-8. Average total monthly rainfall at the Bomet Water Supply Station
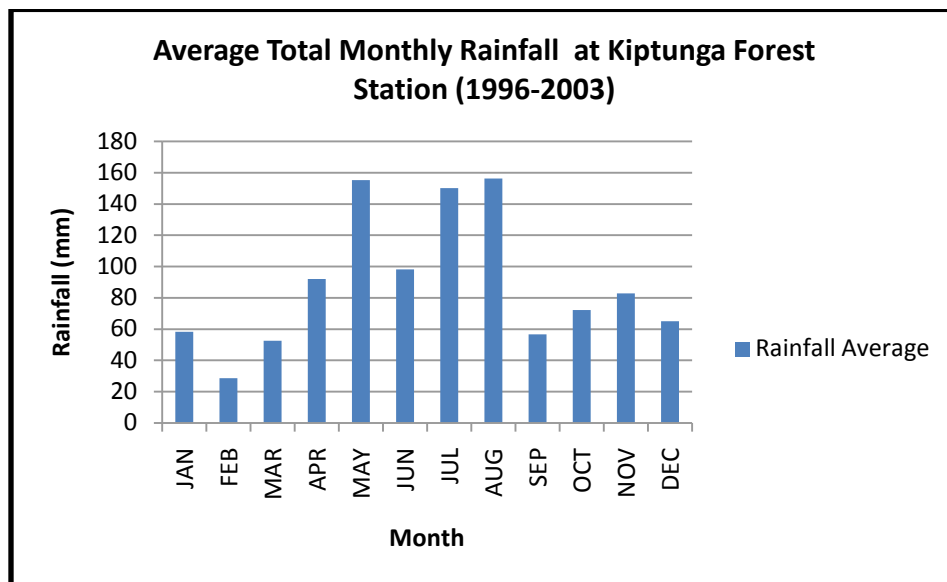
Figure 3-9. Average total monthly rainfall at the Kiptunga Forest Station
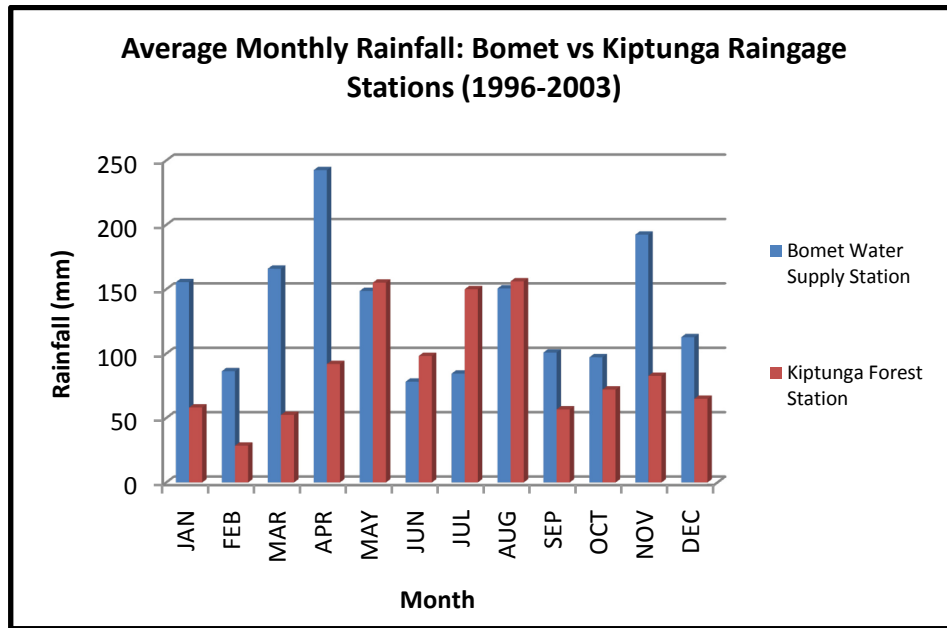
Figure 3-10. Average monthly Rainfall at the Bomet and Kiptunga Rain gauge

Stations
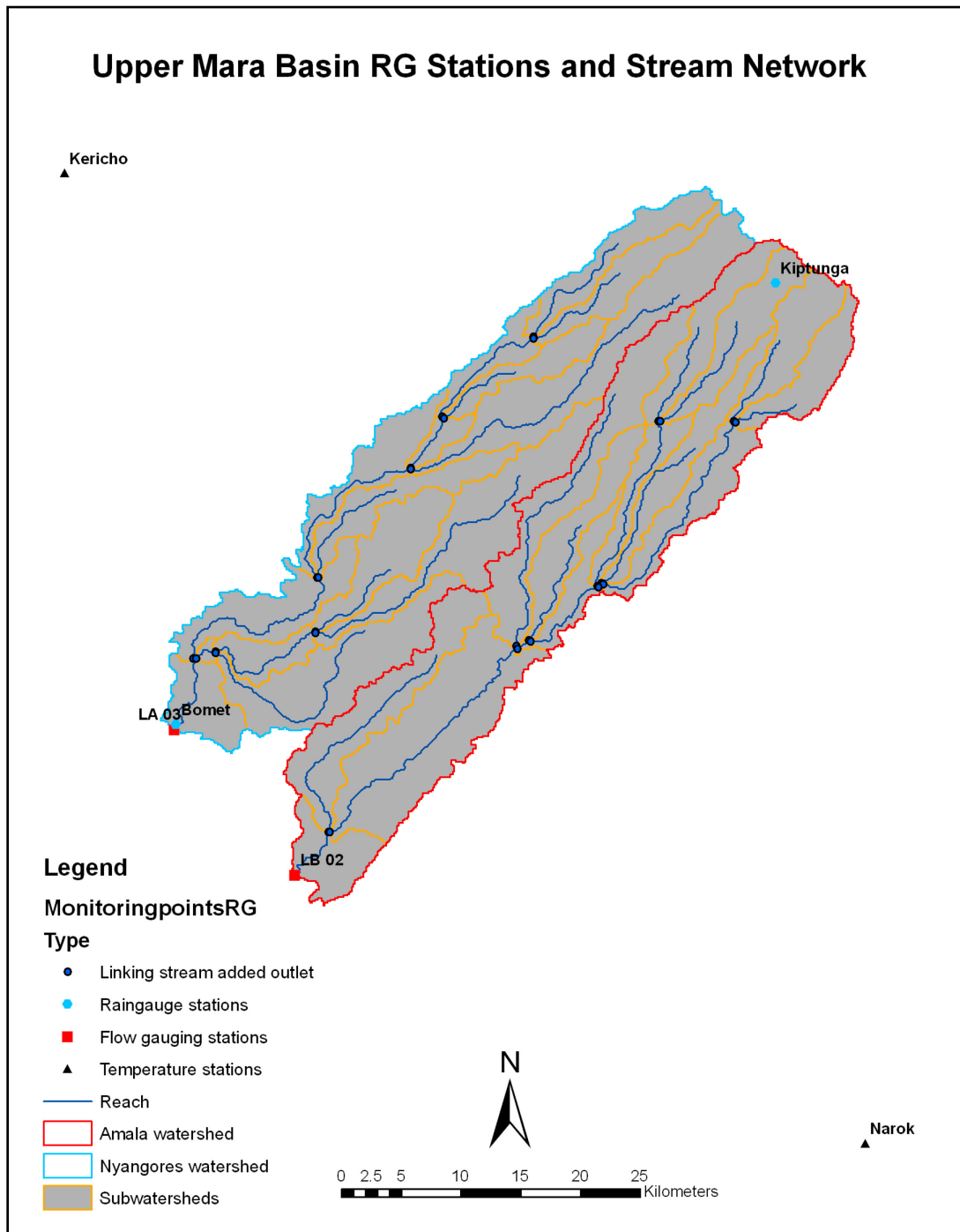
Figure 3-11. Upper Mara Rain gauge Stations with Outlets and Stream Network

Another source of rainfall data for the hydrological modeling was obtained from the Famine Early Warning System (FEWS) Rainfall Estimation (RFE) imagery. This is a computer generated product that uses Meteosat infrared data at a horizontal resolution of 10km (Xie and Arkin, 1996). The rainfall is obtained by means of a python script developed by Gann (2008) which runs in an ArcGIS environment and extracts RFE statistics from daily rasters for user defined regions such as watersheds or sub-watersheds. Output is formatted to be compatible with input file format of ArcSWAT, in this case daily time series data tables in the ArcSWAT 2005 model input format. This process resulted in the creation of 30 artificial rain gages as the centroids of the 30 sub-watersheds making up the Amala and Nyangores watersheds (Figures 3.10 and 3.11). Both the Amala and Nyangores watersheds were assigned 15 RFE Rain gauges each for use in the hydrological modeling process. The RFE data was able to provide continuous and complete data ranging from the years 2002 to 2008 which was used in the model simulations.
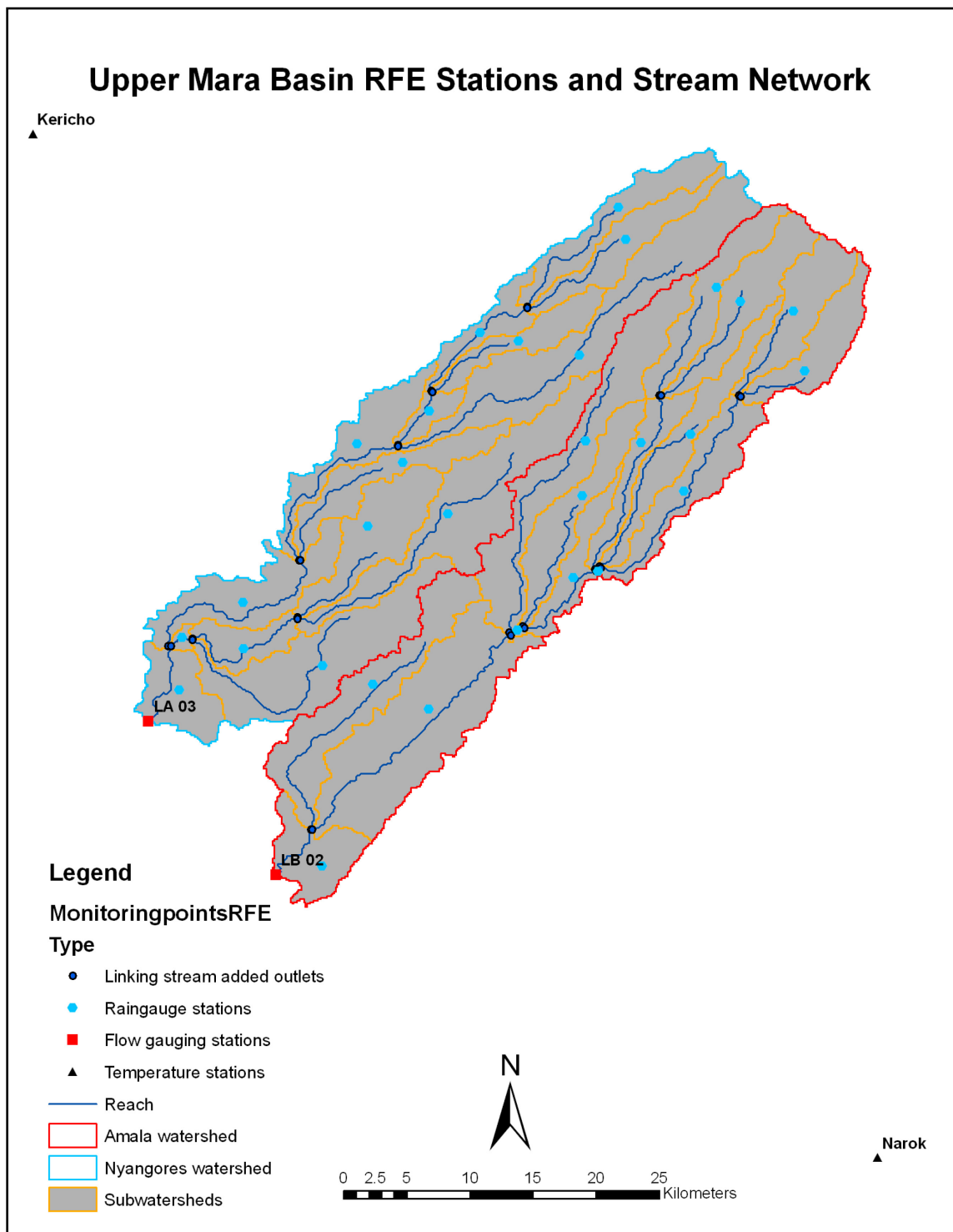
Figure 3-12. Upper Mara RFE Stations with Outlets and Stream Network

3.2.1.5 River Discharge

Daily river discharge data was obtained for the rivers Amala and Nyangores from the gauging stations located at the outlets of the basins. The discharge values for the two tributaries of the Mara; the Amala and Nyangores Rivers were used for calibration and validation of the model.

In the Nyangores watershed, the available discharge data ran from the year 1996 to the year 2008. For the rain gauge data model, out of that the 8 years of complete time series datasets 4 years were used for calibration and the remaining 4 years were used for validation. For the RFE model, 4 years were used for calibration and 3 for validation. In the Amala watershed, observed discharge data spanned from the year 2000 to 2006 and for the rain gauge model, 2 years were used for calibration and 2 years were used for validating the model. For the RFE model 3 years were used for the calibration and 2 years for validation of the model. The length of the simulations was determined by the availability and length of time series data for discharge, air temperature and rainfall which are key pieces in the model simulation.
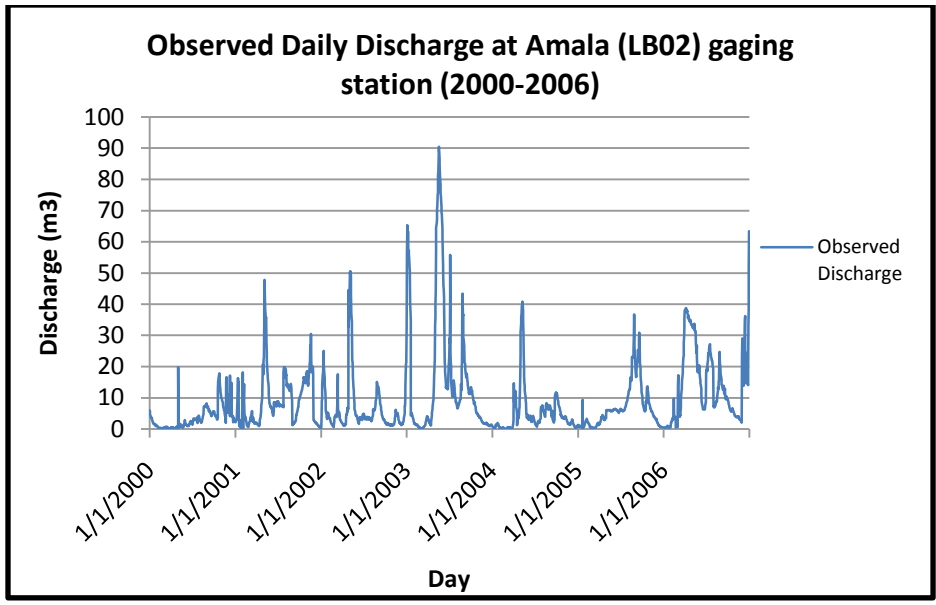
Figure 3-13. Observed Daily Discharge at Amala (LB02) gaging station
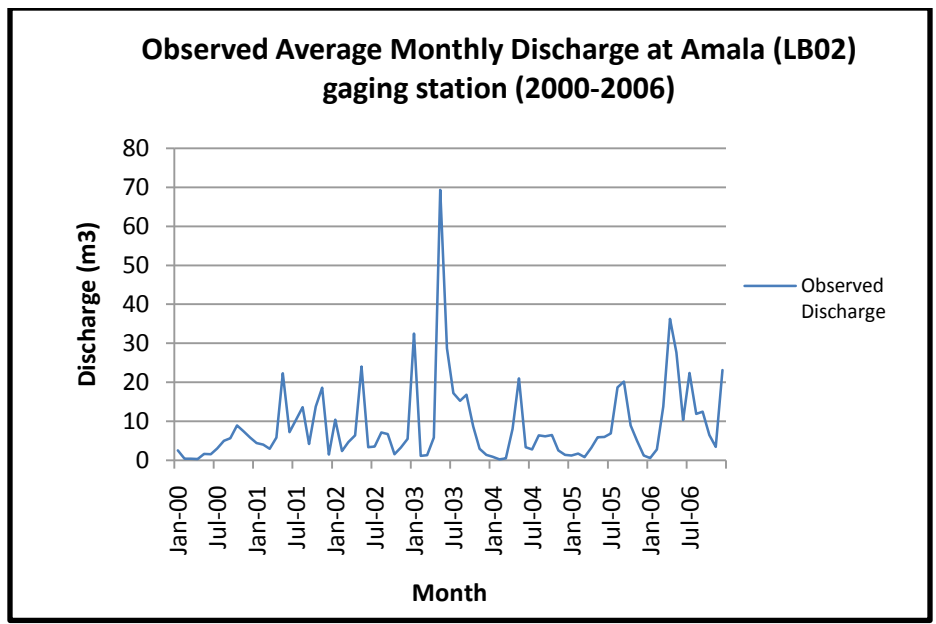
Figure 3-14. Observed Average Monthly Discharge at Amala (LB02) gaging station
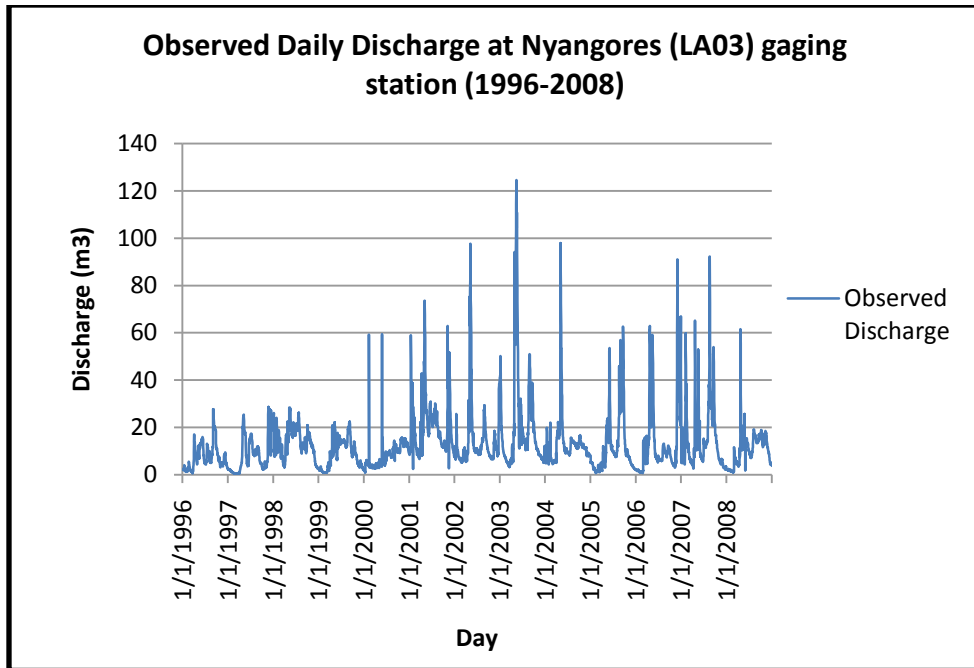
Figure 3-15. Observed Daily Discharge at Nyangores (LA03) gaging station
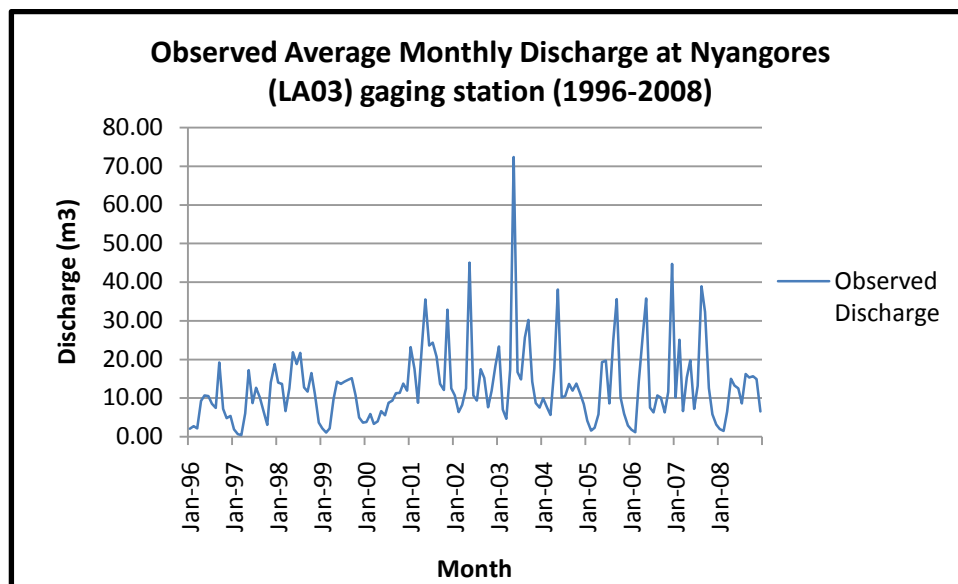
Figure 3-16. Observed Average Monthly Discharge at Nyangores (LA03) gaging station

3.2.2 Model Run

To set up a hydrological SWAT model basic data are required: topography, soil, land use and climatic data (Schuol et al., 2006). The model setup involved five steps: (1) data preparation, (2) sub-basin discretization, (3) HRU definition, (4) parameter sensitivity analysis, (5) calibration and uncertainty analysis.

The DEM was projected to the required projection parameter which is UTM Zone 37 South. A mask was used to reduce the area for stream delineation and analysis of terrain drainage patterns of the land surface. The streams were delineated from the DEM which accurately delineated the location of the streams. The land use land cover layer was reclassified into the SWAT /USGS land use code as per required by the model and linked to a user table with the land use code Since the SWAT database includes only soils from the United States and not from the study area, the soil types and their qualities from the study area had to be entered manually and thereafter the soil layer was added and linked to its various soil types in the database via a look up table.

Watershed and sub-watershed delineation was carried out using the DEM and has various steps including: DEM setup, stream definition, outlet and inlet definition, watershed outlets selection and definition and calculation of sub basin parameters. For the stream definition, the threshold based stream definition option in the Graphic User Interface was used to define the minimum size of the sub basin this defines the minimum drainage are required to form the origin of a stream (Setegn, 2008). By choosing a threshold area of 2000 ha much less than the default value of 15,000 ha, the stream definition is more enhanced and much more likely to delineate all the existing streams in the watershed of interest.

The sub-watersheds are then divided into units based on their unique combination of land use, soils and slope combinations and these units are known as HRUs (hydrologic response units). The model was then run on a default simulation of 8 years from 1996 to 2003 for the Rain gauge data and from 2002 to 2003 a period of two years for the RFE data.

3.2.3 Sensitivity Analysis

Sensitivity analysis and calibration techniques are generally referred to as either manual or automated, and can be evaluated with a wide range of graphical and/or statistical procedures. Uncertainty can be defined as the estimated amount by which an observed or calculated value may depart from the true value (Shirmohammadi et al., 2006). Parameter reduction is very important in distributed watershed models because they involve a lot of parameters. All these parameters are cumbersome to deal with therefore a sensitivity analysis is carried out in order to determine which of these parameters are influential for an efficient calibration.

3.2.3.1 Latin Hypercube One factor At a Time (LH-OAT) analysis

Latin Hypercube One factor At a Time analysis combines Latin Hypercube (LH) sampling and One factor At a Time (OAT) design for simulation. The concept of Latin Hypercube simulation (McKay et al., 1979; McKay, 1988) is based on the Monte Carlo simulation but uses a stratified sampling instead of random sampling. It subdivides the distribution of each parameter into n ranges, each with a probability of occurrence equal to 1/n. Random values of the parameters are generated such that each range is sampled

only once. Then, the model is run n times with the random combinations of the parameters. The exact steps involved are given below.

Let $Y = f(x_1, x_2,.......x_k)$, with x_1,x_2 --- are the independent variables

- Range of each variable is divided into n non-overlapping interval and one value from each interval is chosen.

- Then, n values of x_1 is randomly paired to n values of x_2

- Next, n pairs of x_1 and x_2 are randomly combined to n values of x_3 to form n triplets.

- The procedure is continued to form n k-tuplets

- This n k-tuplet is called LH Sample.

- From this LHS, sensitivity analysis is carried out by following OAT scheme.

Sensitivity of individual parameters of the model is determined by changing one factor or parameter at a time. So, the change in the output in each model run can be unambiguously attributed to the input parameter changed without the assumptions of relatively few inputs having important effects. The measure of sensitivity can be computed using sensitivity index (I) as

$$I = \frac{(y_2 - y_1)/y_1}{(x_{i2} - x_{i1})/x_{i1}} \tag{3-6}$$

Where y_1, y_2 are the output corresponding to the parameter inputs x_{i1} and x_{i2} respectively.

The sensitivity analysis was worked out on the rate of change of sum of squares of deviations of measured and simulated model output (flow) of interest and not on the model output directly. Sensitivity ranking prepared by this criterion is especially helpful for calibration of parameters.

Table 3-1 SWAT Parameters pertinent to water flow

Sl No	Name	Description
1	ALPHA_BF	Baseflow alpha factor (days)
2	GWQMN	Threshold water depth in the shallow aquifer for flow (mm)
3	ESCO	Soil evaporation compensation factor
4	SLOPE	Average slope steepness (m / m)
5	SLSUBBSN	Average slope length (m)
6	CH_K2	Channel effective hydraulic conductivity (mm / h)
7	CN2	Initial curve number (II) value
8	SOL_AWC	Available water capacity (mm water / mm soil)
9	SURLAG	Surface runoff lag time (days)
10	GW_DELAY	Groundwater Delay (days)
11	RCHRG_DP	Deep aquifer percolation fraction
12	CANMAX	Maximum canopy storage (mm)
13	SOL_K	Saturated hydraulic conductivity (mm/h)
14	SOL_Z	Soil depth (mm)
15	EPCO	Plant uptake compensation factor
16	CH_N	Manning's n value for main channel
17	BLAI	Maximum potential leaf area index
18	BIOMIX	Biological mixing efficiency

Source: SWAT 2005 Manual

3.2.4 Model Calibration

3.2.4.1 Auto Calibration and Uncertainty Analysis

For the Nyangores watershed, discharge data from the year 1996 to 1999 were used for calibration and 2000-2003 were used for validation for the Rain gauge data model and 2002 to 2008 for the RFE data model. For the Amala watershed, discharge data from the year 2000 to 2001 were used for calibration and 2002 to 2003 were used for validation for the Rain gauge model. Autocalibration was carried out for each of the watersheds and this resulted in the best parameter values for the sensitive parameters; in this particular case, 10 most sensitive parameters.

For this study, the algorithm used was the Parameter Solutions or Parasol method which involves optimization and uncertainty analysis in a single run. Different algorithms exist within the model for optimization, in this study the Shuffled complex evolution algorithm (SCE-UA) was chosen for optimization. According to Sorooshian et al. (1983), this algorithm selects an initial 'population' by means of random sampling throughout the feasible parameters space for p parameters to be optimized (delineated by given parameter ranges). The portion is portioned into several 'complexes' that consist of 2p+1 points. Each complex evolves independently using the simplex algorithm. The complexes are periodically shuffled to form new complexes in order to share the gained information. It searches over the whole parameter space and finds the global optimum with a success rate of 100%. This algorithm has been found to be robust, effective and efficient and is widely used in watershed model calibration, remote sensing and land surface modeling (Duan et al., 2003). According to van Griensven et al. (2006), it has also been applied with success on SWAT for the hydrologic and water quality parameters.

The objective function chosen for use is the Sum of Squared Residuals (SSQ) which aims at matching a simulated series to a measured time series.

$$SSQ = \sum_{i=1,n} \left[x_{i,measured} - x_{i,simulated} \right]^2 \qquad \text{(3-7)}$$

With n the number of pairs of measured ($x_{measured}$) and simulated ($x_{simulated}$) variables.

3.2.4.2 Manual Calibration and Model Evaluation

It is important to calibrate physically based distributed watershed models before they are used in the simulation of hydrologic processes in order to reduce the uncertainty associated with the model prediction. Hence, a thorough attempt was made to tune the parameters of the model so that the predicted values were in very close agreement with available measured data before going for the determination of the hydrologic components Manual calibration was carried out in order to achieve the best possible values for the simulated discharge. This involved the adjustment of hydrologically sensitive parameters after the autocalibration and running the simulation repeatedly until the minimum distance between the observed and simulated discharge was achieved. Simulation of the hydrologic balance is the foundation for all SWAT watershed applications and is usually a product of its application regardless of the focus of the analysis. Most SWAT applications also report some type of graphical or statistical hydrologic calibration, especially for stream flow, and many of the studies also report validation results. SWAT hydrologic predictions have been evaluated by means of a wide range of statistics. The regression correlation coefficient (R^2) and the Nash-Sutcliffe model efficiency (NSE)

coefficient (Nash and Sutcliffe, 1970) so far are the most used statistics for this purpose of hydrologic calibration and validation.

The Nash-Sutcliffe efficiency (NSE) is a normalized statistic determining the relative magnitude to the residual variance compared to the measured data variance (Nash and Sutcliffe, 1970). The NSE is used to assess the predictive power of hydrological model and indicates how well the plot of observed versus simulated data fits the 1:1 line. It is defined as:

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n} \left(Y_i^{obs} - Y_i^{sim} \right)^2}{\sum_{i=1}^{n} \left(Y_i^{obs} - Y^{mean} \right)^2} \right] \tag{3-8}$$

Where Y_i^{obs} is the ith observation for the constituent being evaluated, Y_i^{sim} is the ith simulated value for the constituent being evaluated, Y^{mean} is the mean of observed data for the constituent being evaluated and n is the total number of observations.

Nash–Sutcliffe efficiencies can range from ∞ to 1 . An efficien y o f 1 ($E = 1$) corresponds to a perfect match of modeled discharge to the observed data. An efficiency of 0 ($E = 0$) indicates that the model predictions are as accurate as the mean of the observed data, whereas an efficiency less than zero ($E < 0$) occurs when the observed mean is a better predictor than the model.

The closer the model efficiency is to 1, the more accurate the model is. Nash–Sutcliffe efficiencies can also be used to quantitatively describe the accuracy of model outputs other than discharge. This method can be used to describe the predicative accuracy of

other models as long as there is observed data to compare the model results to. The NSE is considered fit for model evaluation because it is found to be the best objective function for reflecting the overall fit of a hydrograph (Moriasi et al., 2007).

These performance ratings shown in Table 3-2 below give an indication of how well or how poorly the model has performed in simulating the discharge of the rivers under study. If the performance is considered satisfactory, the model can be used for further analysis of different scenarios such as land use and climate scenarios and also for forecasting which will make it a useful tool in prediction of discharge and water balance of the watershed under study.

Table 3-2 Reported performance ratings for NSE

Model	Value	Performance Rating	Modeling Phase	Reference
SWAT	>0.65	Very good	Calibration and validation	Saleh et al. (2000)
SWAT	0.54 to 0.65	Adequate	Calibration and validation	Saleh et al. (2000)
SWAT	>0.50	Satisfactory	Calibration and validation	Santhi et al. (2001); adapted by Bracmort et al. (2006)
SWAT and HSPF	0.65	Satisfactory	Calibration and validation	Singh et al. (2004); adapted by Narasimhan et al. (2005)

Source: Moriasi et al., 2007

The R^2 value measures how well the simulated versus observed regression line approaches an ideal match and ranges from 0 to 1, with a value of 0 indicating no correlation and a value of 1 representing that the predicted dispersion equals the measured dispersion. The regression slope and intercept also equal 1 and 0, respectively, for a perfect fit; the slope and intercept are often not reported. The NSE ranges from $-\infty$ to 1 and measures how well the simulated versus observed data match the 1:1 line (regression line with slope equal to 1). An NSE value of 1 again reflects a perfect fit between the simulated and measured data. A value of 0 or less than 0 indicates that the mean of the observed data is a better predictor than the model output and these statistics provide a good indicator on the hydrologic performance in a wide range of conditions (Gassman et al, 2007).

3.2.5 Scenario Analysis

3.2.5.1 Land Use Scenarios

To explore the sensitivity of SWAT outputs to land use and the effect of land use/land cover changes on the discharge of the Amala and Nyangores Rivers, different land use/land cover scenarios were explored. The percent coverages and details of the conversions are presented in the tables at the end of this section. The land use scenarios included;

I. Partial Deforestation

This scenario involved manipulation of the forest cover reducing it partially by converting the deciduous forest type to small scale or close grown agricultural land

to test the effect it would have on the Nyangores and Amala discharge. This is a realistic scenario following the ongoing trend of deforestation and conversion of the forest land to agricultural land.
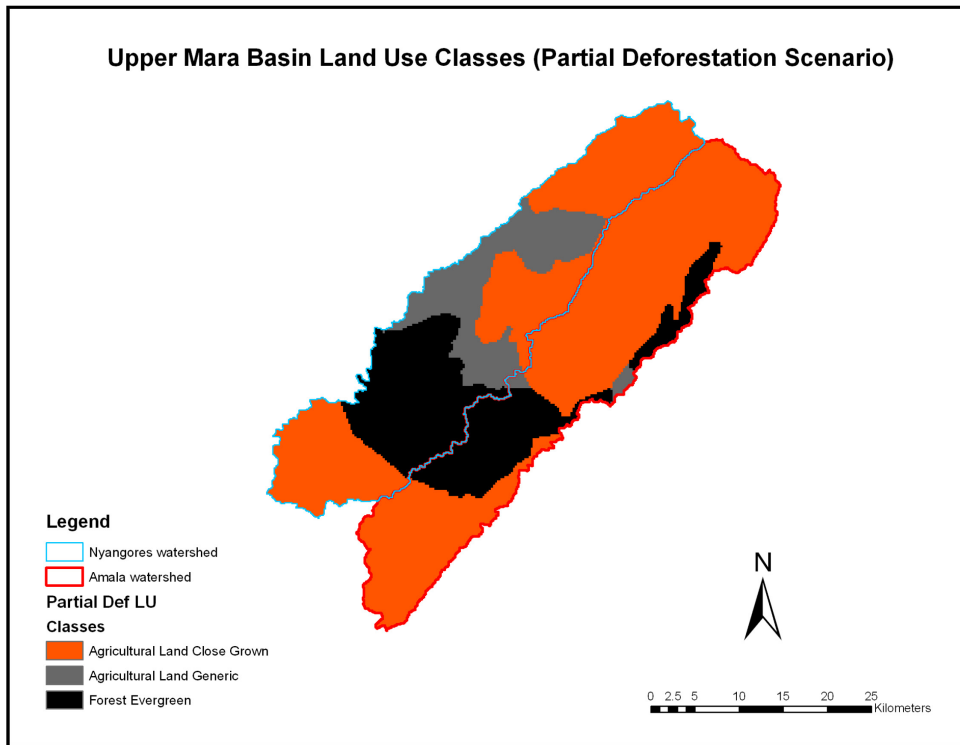
Figure 3-17. Upper Mara Basin Partial deforestation Scenario

II. Complete Deforestation

This scenario involved manipulation of the land cover to examine the effect the absence of forest cover by replacing all the existing forest cover with grassland would have on the discharge of the Nyangores and Amala Rivers.
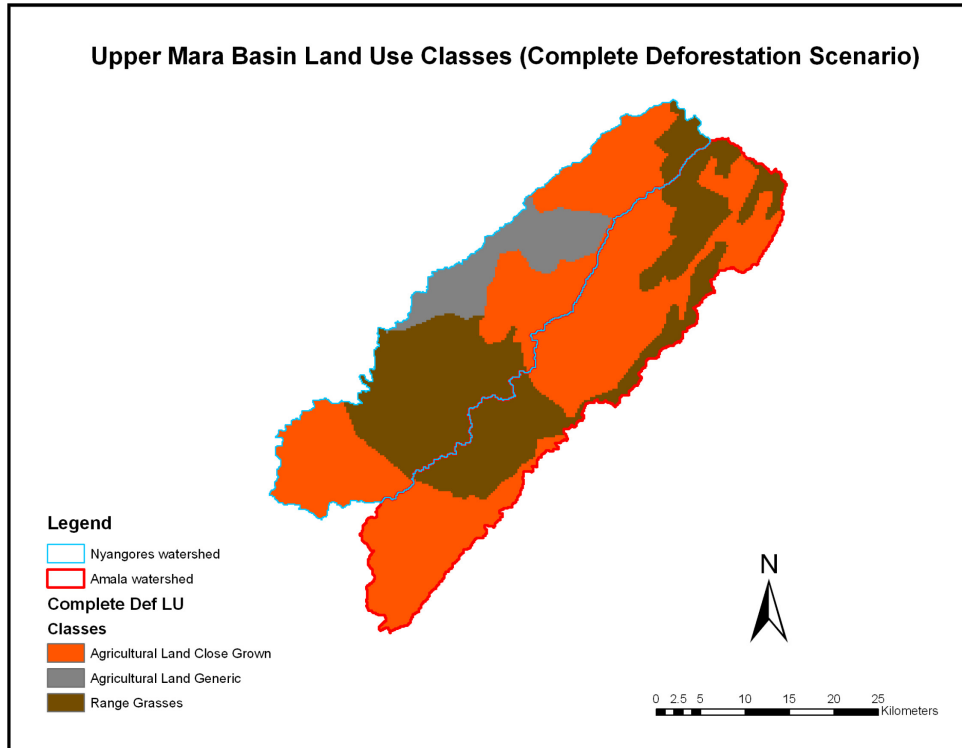
Figure 3-18. Upper Mara Basin Complete Deforestation Scenario

III. Conversion of Forest to Agriculture

Replacement of forest land by agriculture is a common trend within the study area and is seen to be one of the major causes of extreme high and low river flows and increased sediment load in the Nyangores and Amala Rivers. This scenario was carried out by replacing all forest cover with agriculture particularly small scale agriculture.

86
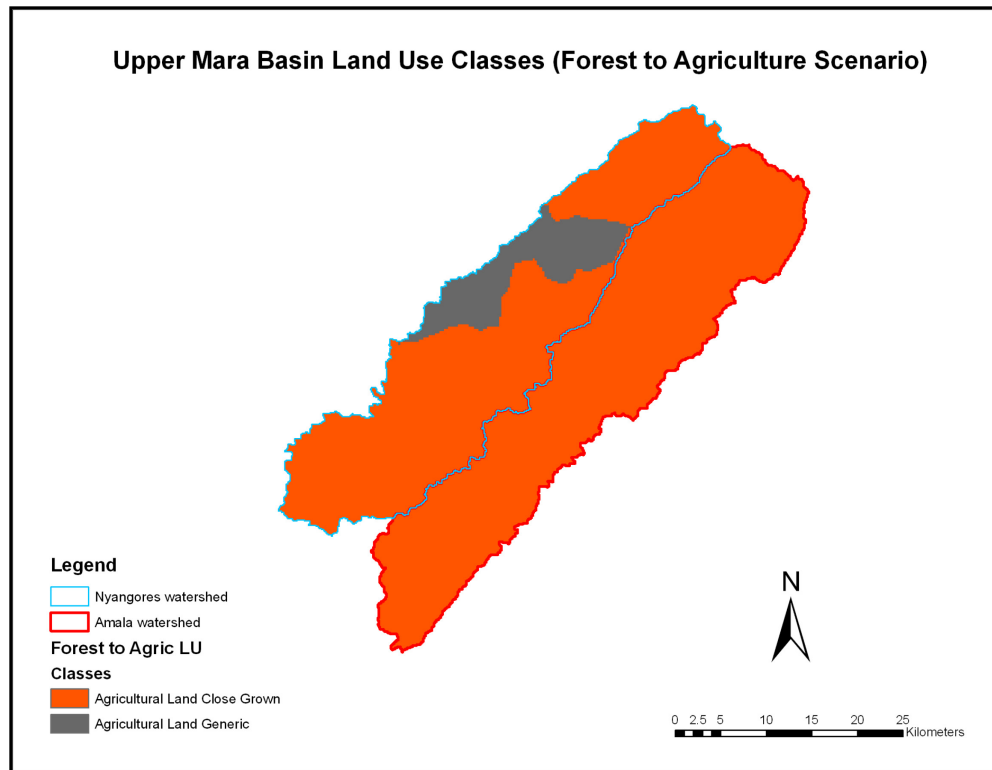
Figure 3-19. Upper Mara Basin Forest to Agriculture Scenario

Table 3-3 Areal Coverage of Land use/ Land Cover

Land Use Scenario / Basin	NY LU 08	AM LU08	NY PD	AM PD	NY CD	AM CD	NY FA	AM FA	Upper Mara
Forest Evergreen	182.4	147.2	182.4	147.2	0	0	0	0	330.26
Forest Deciduous	25.95	94.26	0	0	0	0	0	0	120.61
Forest Mixed	40.11	9.2	0	0	0	0	0	0	49.39
Agricultural Land Generic	121	0	161.09	9.2	121	0	121	0	121.54
Agricultural Land Close Grown	323.03	444.32	349	538.58	323.03	444.32	571.49	694.98	709.45
Range Grasses	0	0	0	0	248.46	250.66	0	0	0
TOTAL (Sq. Km)	692.49	694.98	692.49	694.98	692.49	694.98	692.49	694.98	1331.25

Table 3-3 shows the areal coverage of different land use/ land cover types in the entire study area (Upper Mara Basin), and in the Amala and Nyangores watersheds for the different land use scenarios. Where AM is Amala Basin, NY is the Nyangores Basin, PD is Partial Deforestation, CD is Complete Deforestation, FA is conversion of Forest to Agriculture and LU_08 is the land use in the year 2008.

Table 3-4 Percent Increase/Decrease in Land use/Land cover across scenarios

LAND USE/LAND COVER	NY PD	AM PD	NY CD	AM CD	NY FA	AM FA
Forest Evergreen (%)	-26.59	-41.28	-100	-100	-100	-100
Forest Deciduous (%)	-100	-100	-100	-100	-100	-100
Forest Mixed (%)	-100	-100	-100	-100	-100	-100
Agricultural Land Generic (%)	33.13	0	0	0	0	0
Agricultural Land Close Grown (%)	8.043	21.21	0	0	76.96	56.41

Table 3-4 above shows the percent increase and decrease of the different land use/ land cover types in the various land use scenarios and accounts for those that were reduced, increased, replaced or eliminated in the different scenarios.

Table 3-5 Percent areal coverage of Land Use/Land Cover Type

LAND USE LAND COVER TYPE	NY_LU08	AM LU08	NY PD	AM PD	NY CD	AM CD	NY FA	AM FA	Upper Mara
Forest Evergreen	26.34	21.18	26.34	21.18	0.00	0.00	0.00	0.00	24.81
Forest Deciduous	3.75	13.56	0.00	0.00	0.00	0.00	0.00	0.00	9.06
Forest Mixed	5.79	1.32	0.00	0.00	0.00	0.00	0.00	0.00	3.71
Agricultural Land Generic	17.47	0.00	23.26	1.32	17.47	0.00	17.47	0.00	9.13
Agricultural Land Close Grown	46.65	63.93	50.40	77.50	46.65	63.93	82.53	100.00	53.29
Range Grasses	0	0	0	0	35.88	36.07	0	0	0.00
TOTAL (%)	100	100	100	100	100	100	100	100	100

The table above shows the percent areal coverage of land use/ land cover type in the Upper Mara Basin and also in the Amala and Nyangores Basins.

3.2.5.2 Climate Change Scenarios

According to the Intergovernmental Panel on Climate Change (2007), climate change can be defined as an identifiable change in the state of the climate by change in the mean and/or variability of its properties and that persists for an extended period, typically decades or longer. Trends from 1900 to 2005 have been observed in precipitation and have seen a decrease in precipitation in the Sahel region which has been accounted for in the precipitation reduction scenarios carried out in this study. In the case of projection of future changes of changes in climate for the 21st century and beyond, consideration was given to the scenarios described in the IPCC Special Report on Emission Scenarios and the scenario chosen for the region was A1Fl (rapid economic growth and fossil intensive). This was based on the fact that Kenya is a developing country with a rapid

population growth, is embracing new technology and is dependent on fossil fuel as its major source of energy. The A1Fl scenario projects a temperature increase with a best estimate of 4.0 degrees centigrade with a likely range of between 2.4-6.4 degrees centigrade. Precipitation in the 21st century and beyond is projected to increase by between 5 and 20% for the region. This prompted the precipitation and temperature scenarios below purposely set to capture the effect of both increase and decrease of precipitation and increase in surface temperature.

The different climate scenarios explored included;

I.   A 20% uniform reduction in the RFE Precipitation

II.   A 20% uniform increase in the RFE Precipitation

III.   A 10% uniform reduction in the RFE Precipitation

IV.   A 10% uniform increase in the RFE Precipitation

V.   A 5% uniform increase in the measured air temperature

VI.   Combinations of precipitation decrease and temperature increase

VII.   Combinations of precipitation increase and temperature increase

VIII. A combination of land cover change and precipitation reduction to assess the combined effects of the two

These were carried out by replacing the precipitation and temperature files in the model and running the simulations with the best parameters acquired from the calibration process.

4.0 RESULTS AND DISCUSSIONS

The results of this study are divided into two categories: land cover classification and hydrological modeling. The land cover mapping provides data on the type of land use/ land cover types present within the Amala and Nyangores watersheds. The hydrological modeling makes use of the land cover map and provides data on the water balance of the Amala and Nyangores watersheds and the influence of land use and climate change.

4.1 LAND COVER MAPPING

To investigate the impact of land use/land cover change on the water flux of the Mara River, it was necessary to generate a current land use/land cover map of the study area in order to obtain an accurate representation of the current land use/land cover, run different scenarios and consequently, assess the impact of land use/land cover change on the discharge of the Amala and Nyangores tributaries of the Mara River.

4.1.1 Land Cover Classification by Expert Classifier

The expert classifier was built and this involved the generation of a decision tree by linear partitioning and the process resulted in the production of decision trees based on the training data that was specified for input into the statistical package R. This resulted in production rules used in the expert classifier to classify the image. The resultant decision trees together with the production rules are shown below in Figures 4-1 and 4-2.
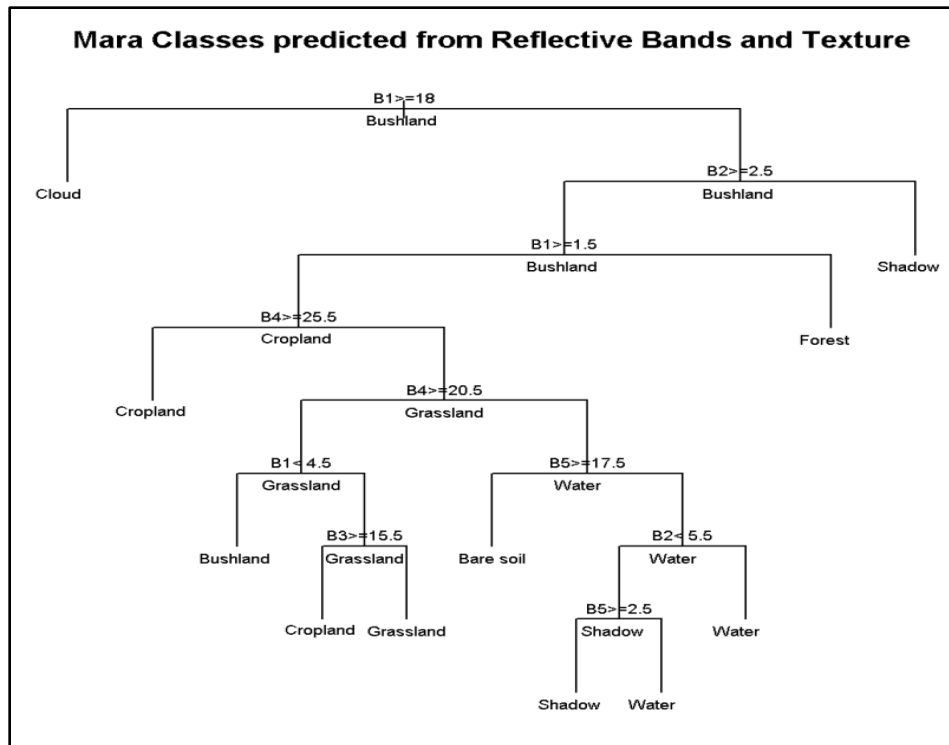
Figure 4-1. Decision tree generated from the reflective and texture bands
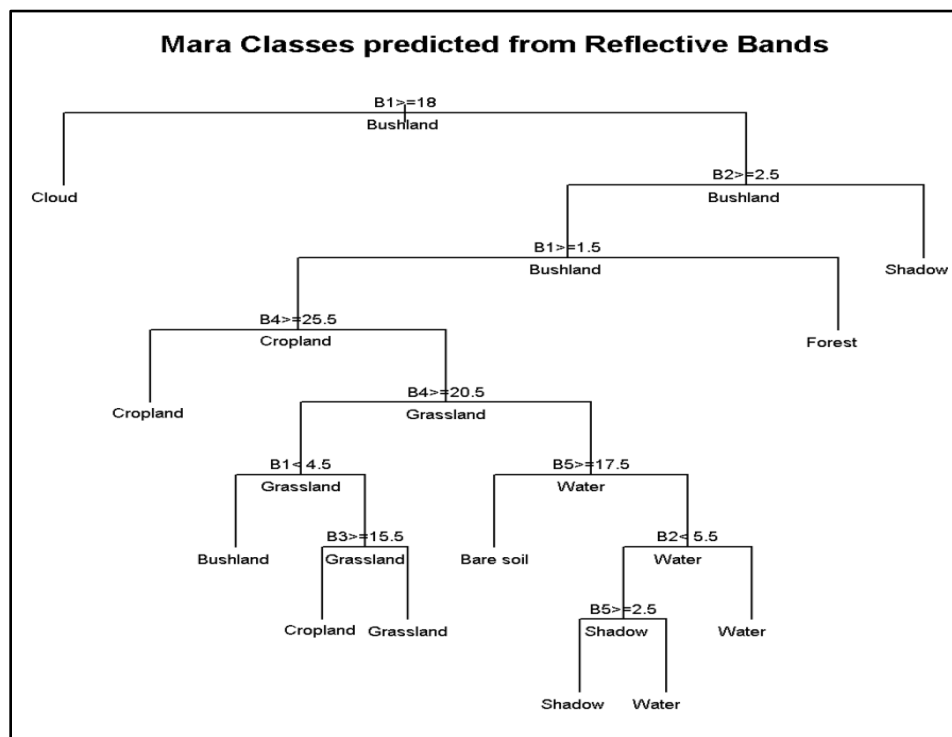
Figure 4-2. Decision tree generated from reflective bands

The application of these production rules in an expert classifier resulted in the classified image below.
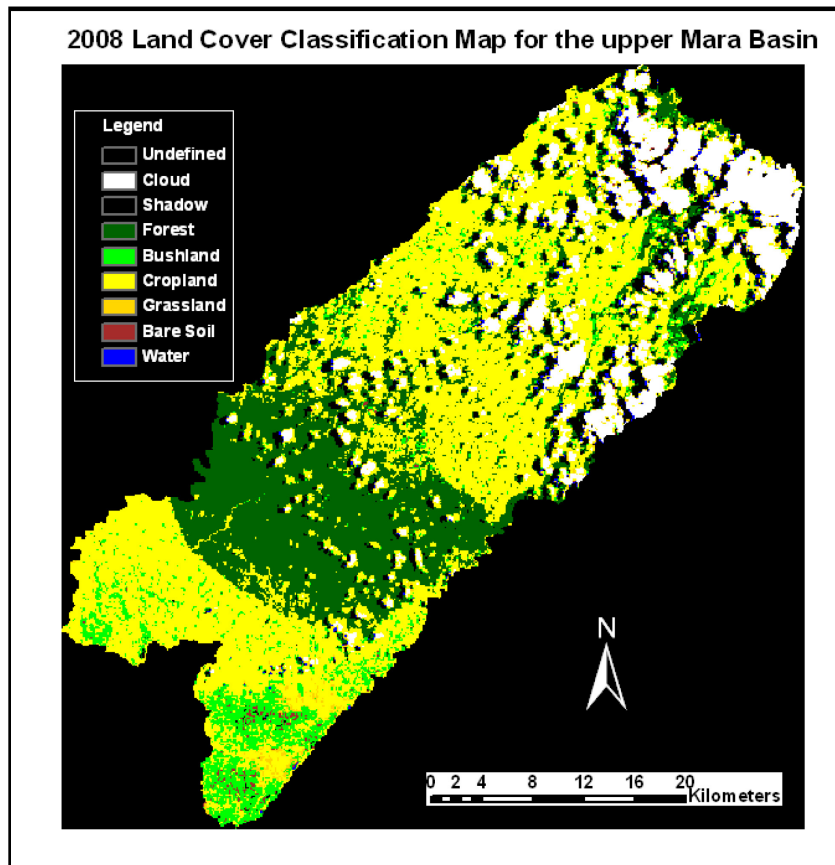
Figure 4-3. 2008 Land Cover Classification Map for the upper Mara Basin

From a visual assessment of the resultant land cover classification map above, cloud cover was a problem in the upper parts of the basin and was a hindrance to the classification in certain areas. The classification was successful in distinguishing different land cover classes in the image but the accuracy of the classification was better determined by the use of an error or confusion matrix

4.1.2 Accuracy Assessment

Table 4-1 Error matrix of the classification map derived from Landsat Thematic Mapper data of the upper Mara Basin, Kenya

	Bare soil	Bushland	Cloud	Cropland	Forest	Grassland	Shadow	Water	Row total
Bare soil	56	2	0	2	0	7	0	9	76
Bushland	3	62	0	1	0	3	0	7	76
Cloud	0	0	92	0	0	0	0	0	92
Cropland	0	5	0	83	4	10	0	5	107
Forest	0	14	0	0	70	0	2	5	91
Grassland	1	12	0	5	0	70	0	0	88
Shadow	2	5	0	0	0	0	86	0	93
Water	0	0	0	0	0	0	0	58	58
Column total	62	100	92	91	74	90	88	84	681

Table 4-2 Producer and User Accuracy

Class	Producer's Accuracy (%)	Omission Error (%)	User's Accuracy (%)	Comission error (%)
Bare soil	56/62 = 90%	10	56/76 = 74%	26
Bushland	62/100 = 62%	38	62/76 = 82%	18
Cloud	92/92 = 100%	0	92/92 = 100%	0
Cropland	83/91 = 91%	9	83/107 = 78%	22
Forest	70/74 = 95%	5	70/91 = 77%	23
Grassland	70/90 = 77%	23	70/88 = 80%	20
Shadow	86/88 = 98%	2	86/93 = 92%	8
Water	58/84 = 69%	31	58/58 = 100%	0

The producer's accuracy shown above shows the probability of a reference pixel being correctly classified and is a measure of omission error while the user's accuracy (or reliability) is the probability that the pixel classified on the map represents the same category on the ground (Jensen, 2005).

The resultant error matrix gave a $\hat{K}$ statistic value of 0.825358 or 82.53% while the overall classification accuracy for the classification was 0.847283 or 84.73%. The reason that the $\hat{K}$ statistic and the overall accuracy differ is because they incorporate different information, the overall accuracy incorporates the major diagonal only and excludes the omission and commission errors. The $\hat{K}$ statistic incorporates the off-diagonal elements as a product of the row and column marginals. Thus these two measures may not agree depending on the amount of error included in the matrix.

The resultant $\hat{K}$ statistic and overall accuracy values of 82.53% and 84.73% can be taken as a fairly good accuracy considering the heterogeneity of the study area that may pose significant difficulties using different classification methods such as a maximum likelihood classifier that assume the data distribution for each class is normally distributed. The error matrix indicates that there was substantial confusion between bushland, forest and grassland which was attributed to the selection of training data and also the fact that use of spectral and texture data alone were not capable of accurately distinguishing these three classes.

The problem with this method of accuracy assessment is that the error evaluation is based on the training pixels, which is based on human or visual interpretation which was used to train algorithm. Therefore, because of this bias, it results in a high accuracy because of the prior knowledge of the analyst, this is usually higher than an accuracy assessment based on ground reference information and is misleading (Muchoney and Strahler, 2002). In this classification however, this was not the case since ground reference data was available and was used as the training data thus eliminating the bias.

4.2 HYDROLOGICAL MODELING

The hydrological modeling exercises were carried out and resulted in discharge simulation values for the Amala and Nyangores watersheds for different rainfall inputs; Rain gauge measurements and radar rainfall estimates (RFE). The hydrographs for the discharge were plotted for daily discharge data for both the uncalibrated and calibrated models.
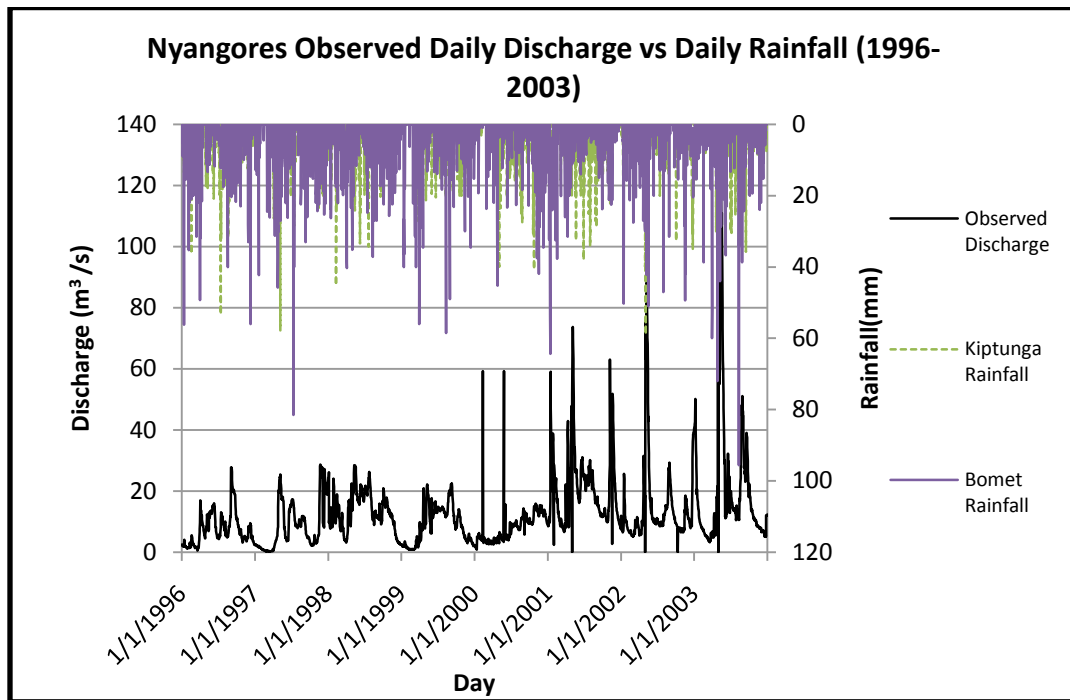
Figure 4-4. Nyangores Observed Daily Discharge vs Daily Rainfall (1996-2003)

From the chart plotted above, a close resemblance in trend can be observed between the discharge and the rainfall where the peaks in rainfall correspond to the peaks in discharge. This can be attributed to the amount of rainfall converted to surface runoff and also direct rainfall on the river channel. The occurrence and amount of rainfall is therefore a significant determinant of the discharge and the discharge hydrograph.
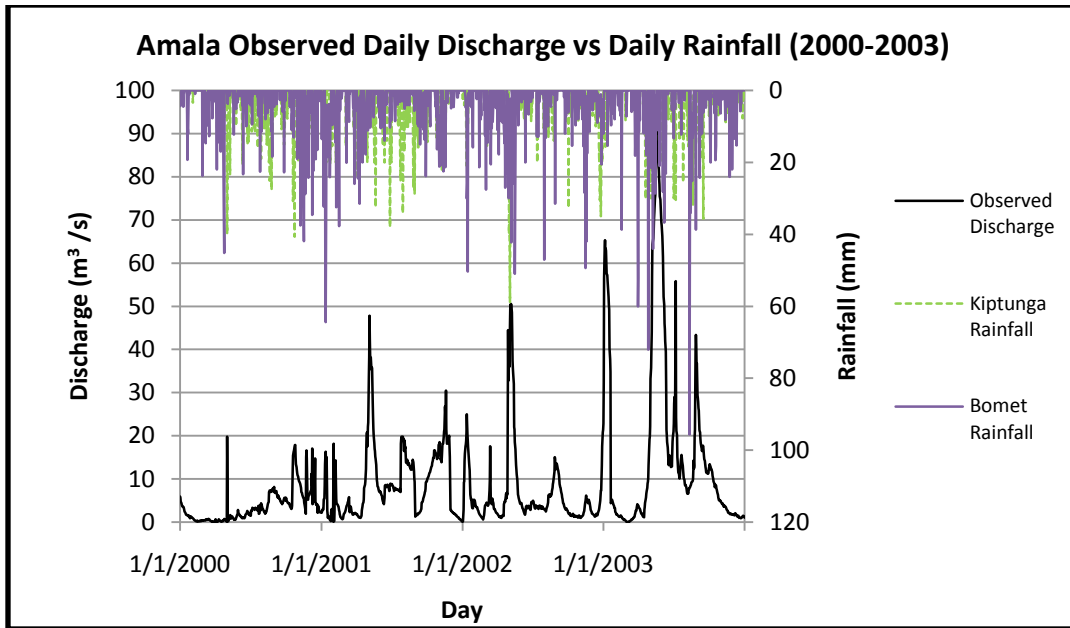
Figure 4-5. Amala Observed Daily Discharge vs Daily Rainfall (2000-2003)
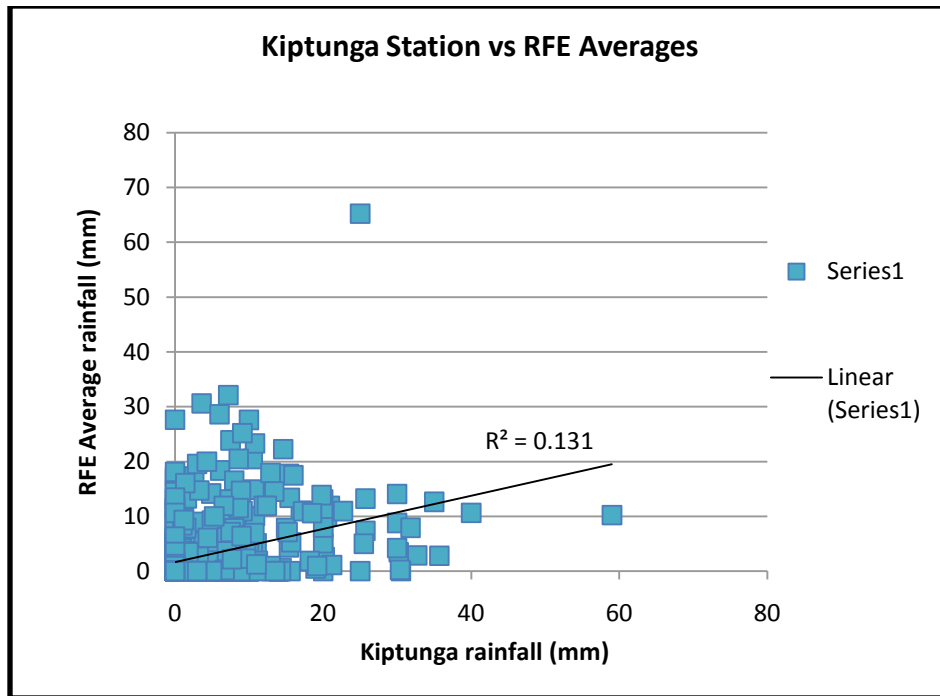
Figure 4-6 Kiptunga Station vs RFE average rainfall
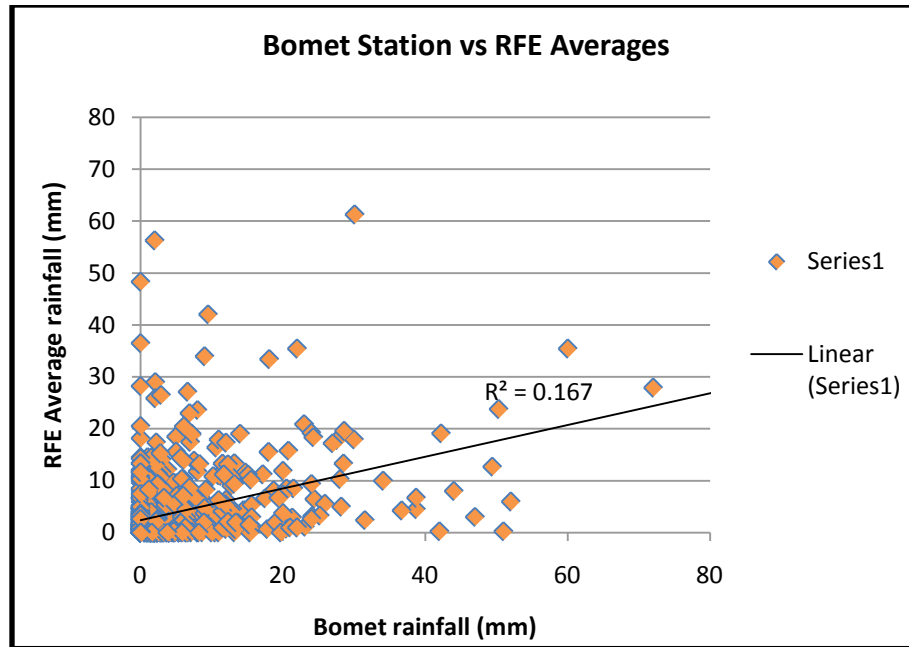
Figure 4-7 Bomet Station vs RFE average rainfall

In the above scatter plots, the rainfall stations are plotted against the average value of the RFE stations that are close to them and may have similar amounts of rainfall for the same time period to test for similarity between the two sources. The R^2 values obtained show a low similarity between the two different sources. This may be due to the fact that the RFE stations although close to the rain gauge stations in question, receive different amounts of rainfall owing to the spatial variability of rainfall. A rainfall cell's mean diameter has been estimated to be between 1-5 and 15 km in diameter and move significantly during rainfall events (Michaud and Sorooshian, 1994) and therefore, the two ground stations in question are therefore not adequate to effectively capture the spatial distribution and heterogeneity of the rainfall occurring within the area. Also, missing values in the ground station data though filled statistically are reliable estimates and cannot replace actual measurements in term of accuracy.

In the same argument of the diameter of rainfall cells it can be argued that the RFE station data is inaccurate since it is based on average values in the form of raster grids but based on the results of the model simulations, the RFE produces a simulation that is of a higher similarity than that of the rainfall stations. Until a dense ground network of rainfall stations is established within the study area, the RFE rainfall appears as a more reliable source of rainfall for semi-distributed and distributed hydrological modeling. Based on the above argument, RFE rainfall was used as the source for the modeling and scenario analyses.

4.2.1 Default Simulations

The hydrographs for the default or uncalibrated simulations are useful for comparison with the observed discharge data and may be used as an indication that the model is capable of simulating the discharge and to what degree the simulation corresponds to or differs from the observed data. From the resulting hydrographs for the Rain gauge data, a comparison made between the observed and simulated discharge shows that the simulated discharge is much higher than the observed discharge with discharge peaks up to 150 m^3/s in the Amala River and up to 200 m^3/s in the Nyangores River.
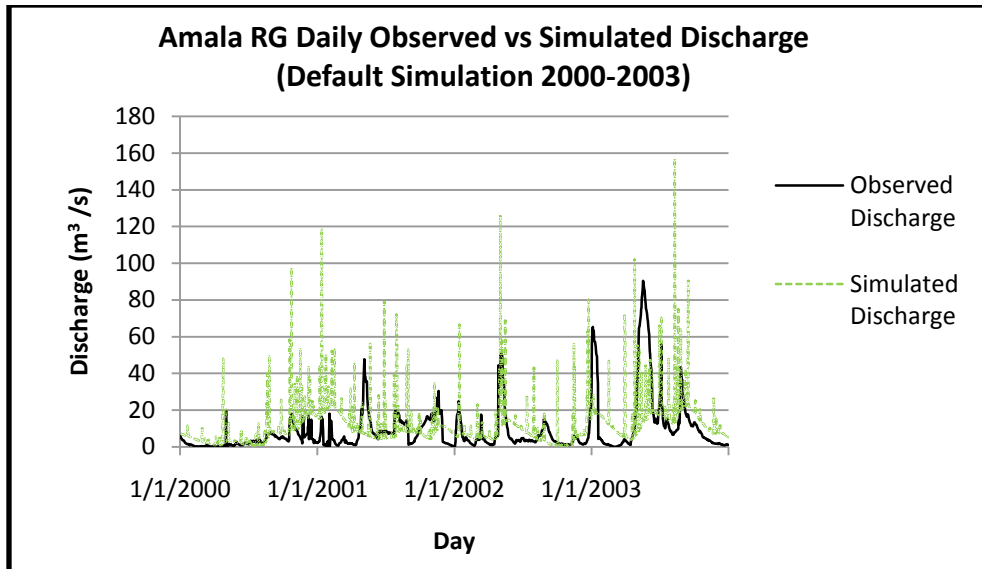
Figure 4-8. Amala Rain gauge Default simulation
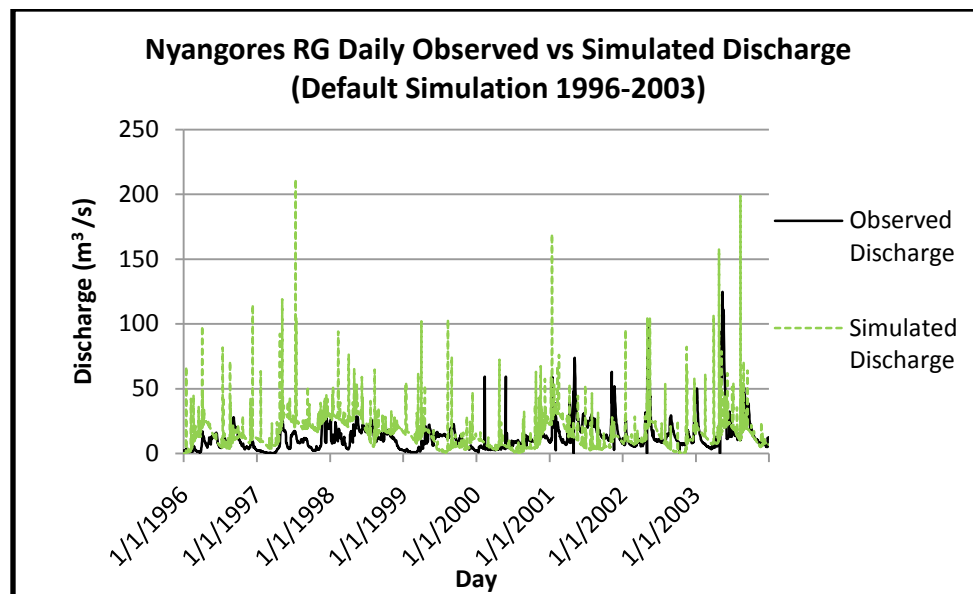
Figure 4-9. Nyangores Rain gauge Default simulation

The difference exhibited between the default simulation and observed discharge hydrographs showed the need for calibration of the model which was carried out using both manual and automatic methods.
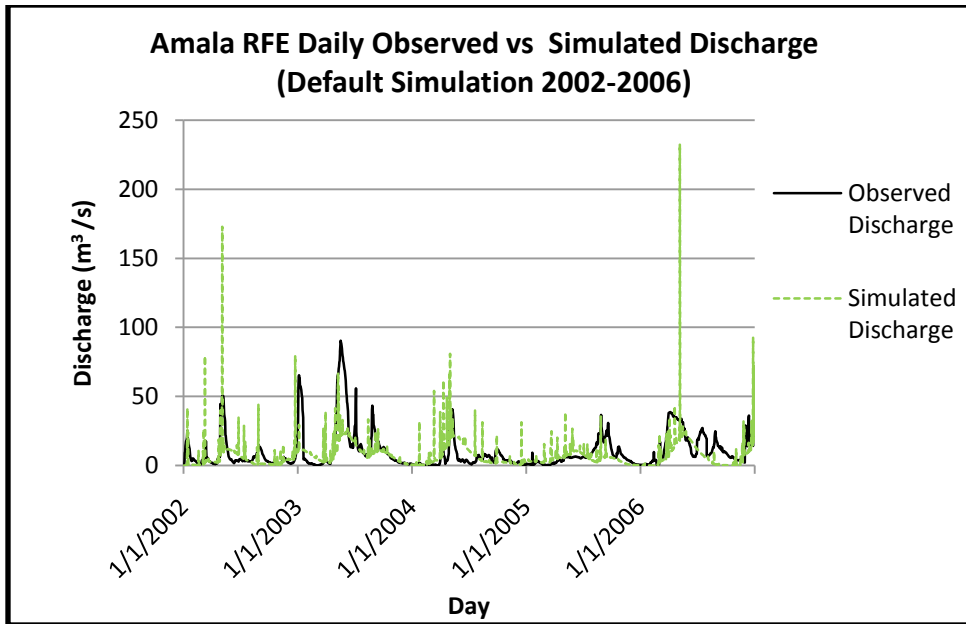
Figure 4-10. Amala RFE Default simulation
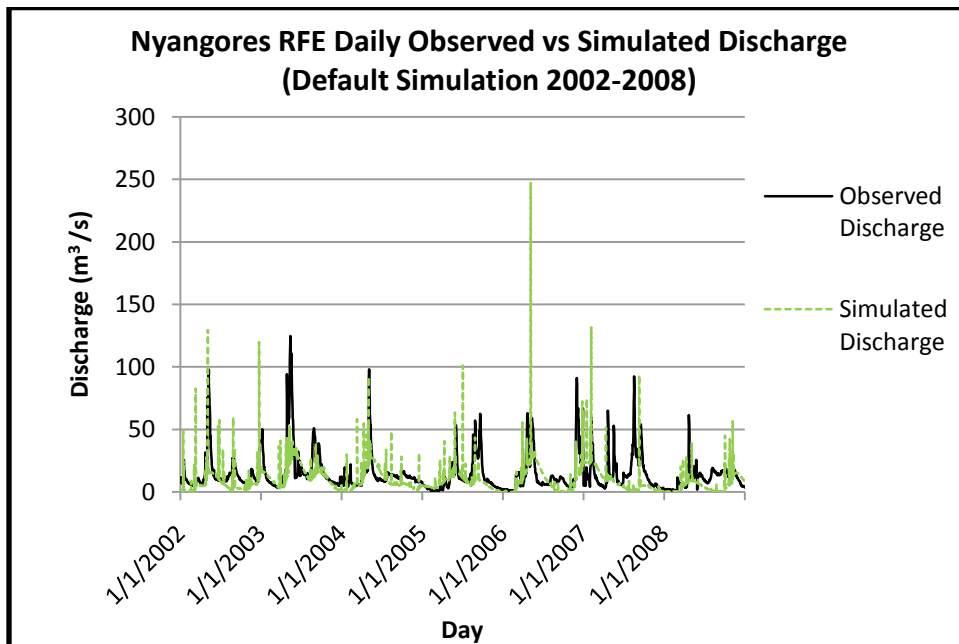
Figure 4-11. Nyangores RFE Default simulation

For the RFE data there was a consistent over-estimation of the discharge with a few extreme peaks, up to 230 m^3/s in the Amala River and up to 240 m^3/s in the Nyangores River. By means of visual comparison, the RFE data simulation showed more similarity to the observed discharge than the Rain gauge data and was expected to result in better calibrated models for both rivers.

4.2.2 Sensitivity Analysis

A sensitivity analysis was carried out and the 10 most sensitive parameters were chosen for calibration of the model. The remaining parameters were found not to have an effect on the simulation of streamflow and were not used for the calibration of the model. The resultant sensitive parameters are shown in Table 4-3 below.

Table 4-3. Sensitivity ranking of parameters towards water flow

SENSITIVITY RANK	AMALA RAIN GAUGE	NYANGORES RAIN GAUGE	AMALA RFE	NYANGORES RFE
1	ESCO	ESCO	CN2	ESCO
2	CN2	CN2	GWQMN	GWQMN
3	GWQMN	ALPHA_BF	ESCO	CN2
4	SOL_Z	GWQMN	SOL_Z	SOL_Z
5	ALPHA_BF	SOL_Z	ALPHA_BF	ALPHA_BF
6	REVAPMN	REVAPMN	SOL_AWC	SOL_AWC
7	SOL_AWC	SOL_AWC	REVAPMN	REVAPMN
8	CANMX	CH_K2	CANMX	CANMX
9	BLAI	BLAI	GW_REVAP	GW_REVAP
10	GW_REVAP	CANMX	SOL_K	BLAI

4.2.3 Calibration and Validation

4.2.3.1 Calibration Statistics

The evaluation of the model and model simulations is best done by statistical comparisons. Statistics such as the Nash-Sutcliffe Efficiency (NSE), Pearson's correlation coefficient (r) and the Coefficient of Correlation (R^2) are able to describe and compare the different datasets (observed and simulated). Moriasi et al., (2007), defines calibration as the process of estimating model parameters by comparison of model predictions or output for a given set of assumed conditions with observed or measured data for the same conditions. In calibration of the model, parameter adjustment was carried out in conjunction with the statistical evaluation until an acceptable correlation or resemblance between the two datasets was achieved. Comparison was carried out for the datasets obtained and the resulting statistics are shown in Table 4-4 and 4-5 below.

Table 4-4 Model evaluation statistics for Daily Discharge

STATISTIC	Rivers							
	Amala				Nyangores			
	RFE		Rain gauge		RFE		Rain gauge	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val
NSE	0.527	0.192	0.004	0.327	0.485	0.0807	-0.445	0.0178
R^2	0.548	0.333	0.206	0.329	0.530	0.233	0.072	0.257
R	0.741	0.57	0.454	0.573	0.728	0.483	0.269	0.507

Table 4-5 Model evaluation statistics for Monthly Discharge

STATISTIC	Rivers							
	Amala				Nyangores			
	RFE		Rain gauge		RFE		Rain gauge	
	Cal	Val	Cal	Val	Cal	Val	Cal	Val
NSE	0.622	0.389	0.076	0.407	0.586	0.094	-0.533	-0.057
R^2	0.654	0.459	0.303	0.413	0.645	0.325	0.085	0.321
R	0.809	0.678	0.550	0.643	0.803	0.57	0.291	0.566

Validation was also carried out for the model simulations and is defined as the process of demonstrating that a given site-specific model is capable of making sufficiently accurate simulations (Refsgaard, 1997).

In the case of the Nyangores and Amala models that made use of the Rain gauge data, there is a clear underperformance of the models in the case of discharge simulation as shown by the different model evaluation statistics. With reference to Table 3-2 in the previous section, these models performed less than satisfactorily in the case of the Rain gauge data with NSE values of less than 0.5. This implies that the rainfall values from the Rain gauge data are not well representative of the actual rainfall that was received in the basins under study. This low NSE value may be the result of a lack of a dense Rain gauge station network within the study area that was unable to capture the different rainfall amounts and account for the spatial variability of the rainfall received.

The monthly values are slightly better for all statistics and this is expected because in this case the discharge values for both the observed and simulated discharge are averaged

over a monthly period and help mute the differences and result in smoothing of the data. This holds true for all the models except for the Nyangores Rain gauge which resulted in negative NSE values. This occurs when the data is moving in opposite directions indicating the mean observed value is a better predictor than the simulated value which is unacceptable (Moriasi et al., 2007). This is an anomalous occurrence but this can be attributed to the increased sample size (n) for the daily data of 8 years as opposed to monthly data for the same period of time.

Ideally, a dataset should be divided into two parts for the calibration and validation processes. The calibration period normally requires a longer period of data in order to capture the variability of the data especially for hydrological data where there may be periods of drought and floods which may have 10 or maybe 20 year return periods. During the calibration period these different events are captured and the input parameters estimated should enable the model to accurately simulate or predict future events. In the above table, the extremely low NSE values for the validation show that the parameters estimated were not able to make accurate predictions and therefore are not suitable for any forecasting of events. This may have been caused by the short calibration periods due to the shortage of measured or observed data causing a failure to capture long term variability of the hydrological processes and also by the uncertainty in the observed data due to estimated missing data values.

Rainfall is the main driving force of the hydrological cycle and when the rainfall for large watersheds such as the Amala and Nyangores watersheds cannot be accurately accounted for this presents a problem in the simulation process and when calibrating the model because this necessitates the rigorous adjustment of parameters which is not only a

time consuming process but also may result in parameter values that may give a good simulation result but are hydrologically unrealistic for the watershed. The hydrographs for the calibration and validation periods follow.

4.2.3.2 Discharge Hydrographs
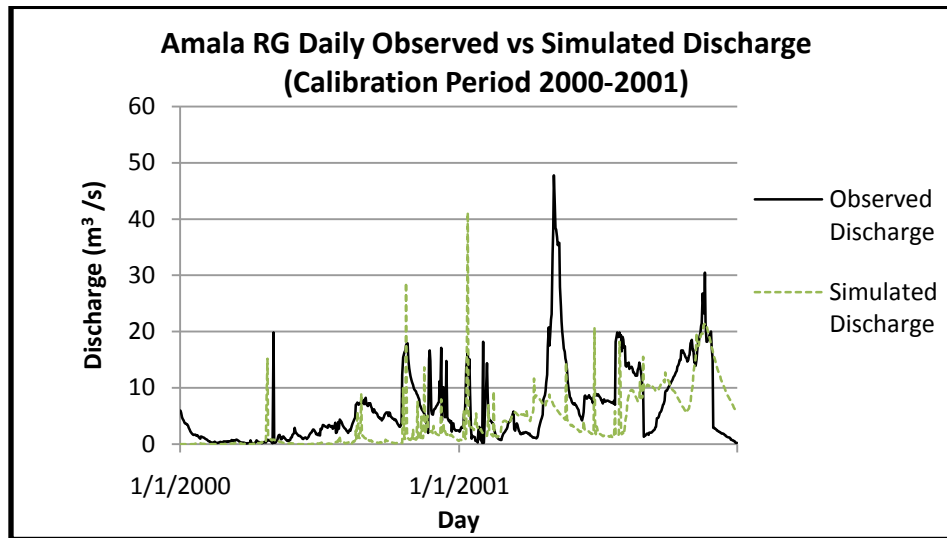
Figure 4-12. Amala Discharge for Rain gauge data
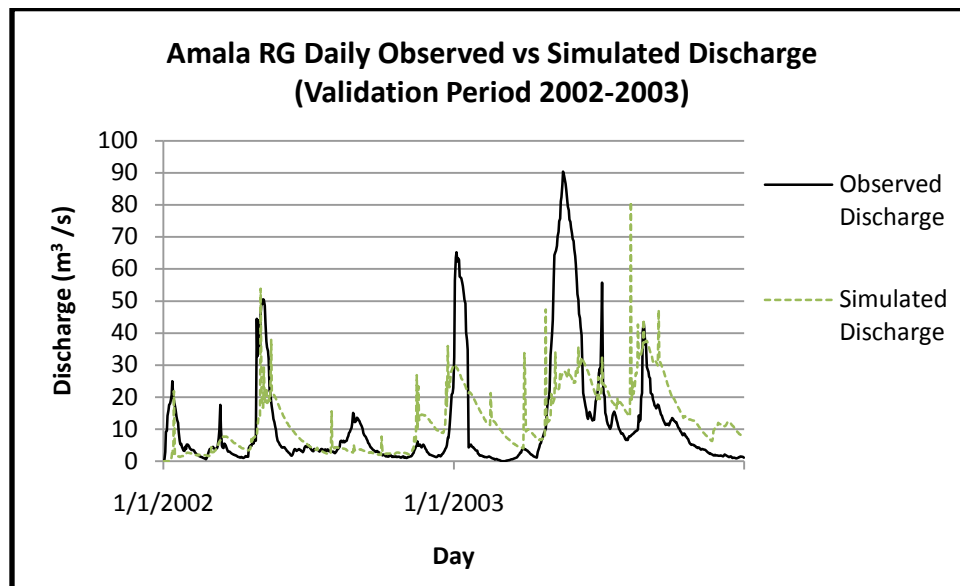
Figure 4-13. Amala Discharge for Rain gauge data for the Validation period (2002-2003)
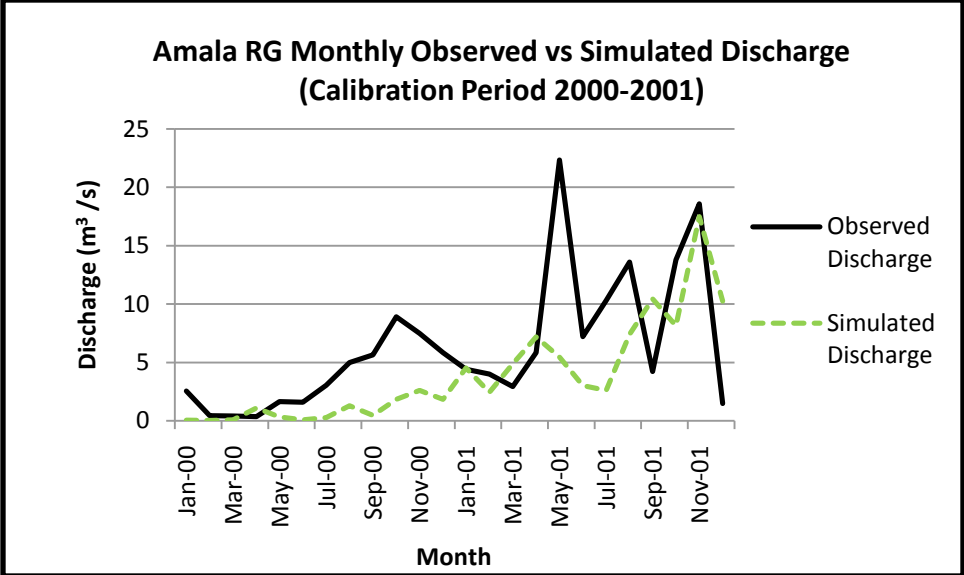
Figure 4-14 Amala Rain gauge Monthly Observed vs Simulated Discharge for the
Calibration period
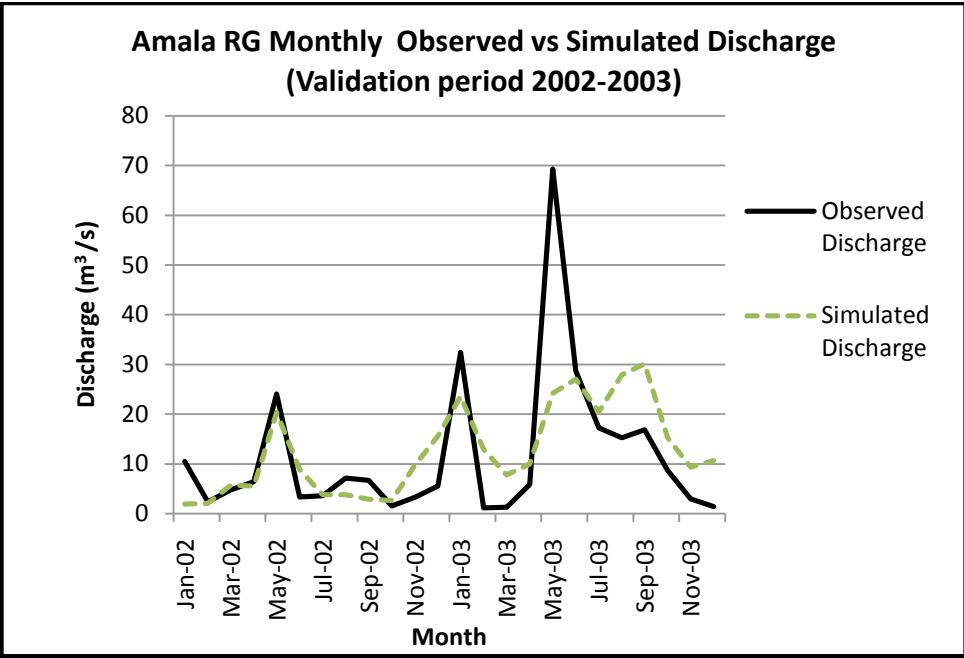
Figure 4-15. Amala Rain gauge Monthly Observed vs Simulated Discharge for
the Validation period
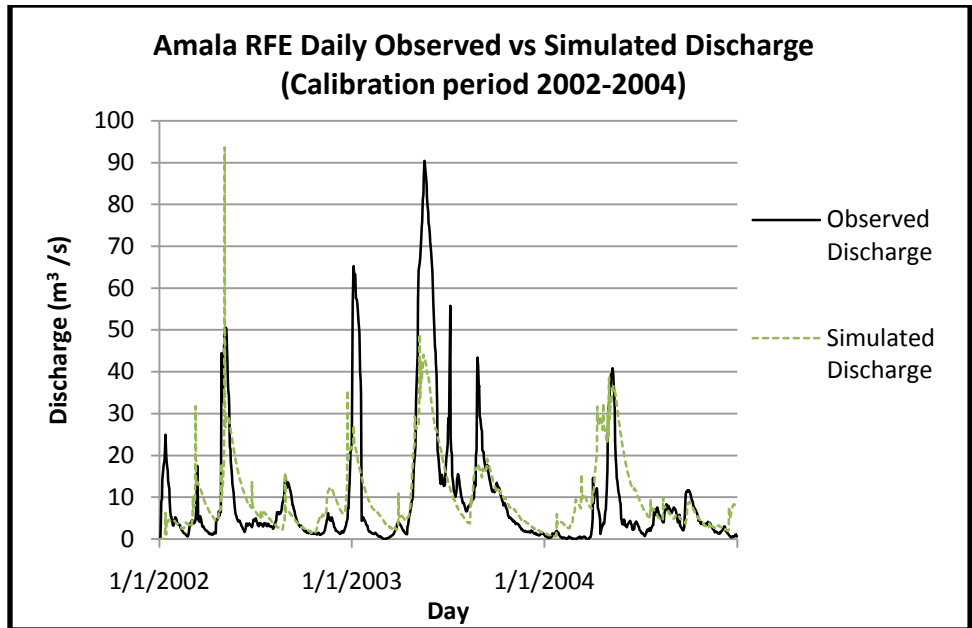
Figure 4-16. Amala Daily Discharge for RFE data (Calibration period)
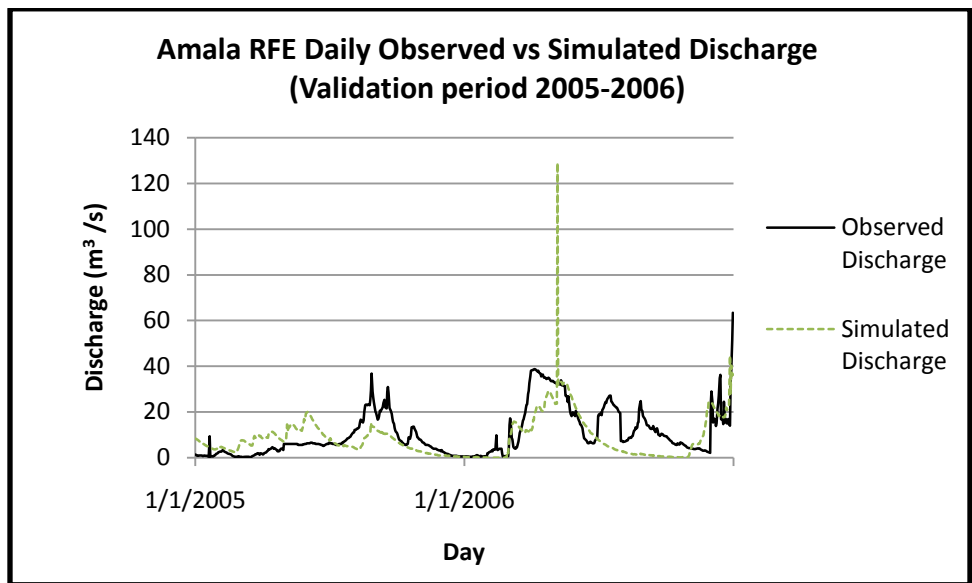
Figure 4-17. Amala RFE Daily Observed vs Simulated Discharge for the Validation period
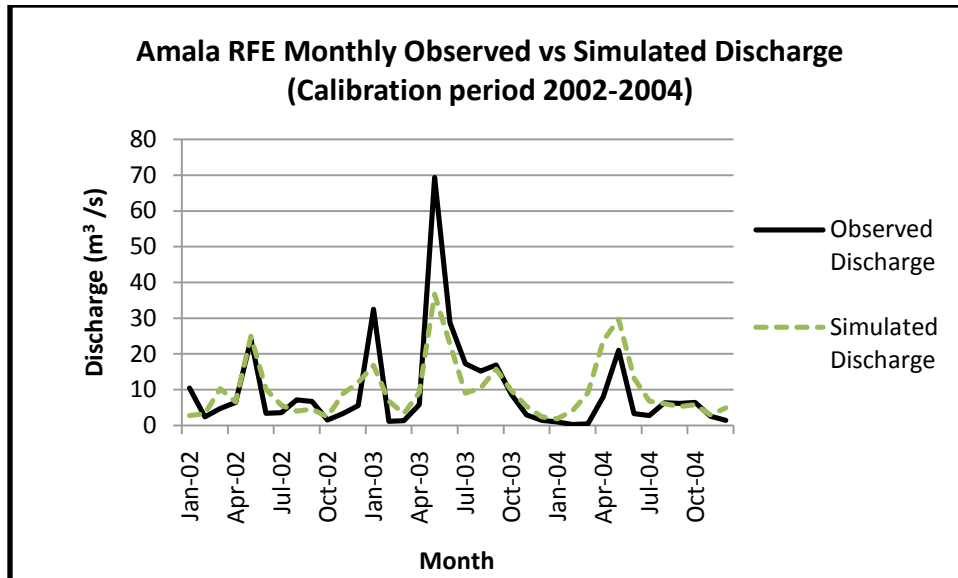
Figure 4-18. Amala RFE Monthly Observed vs Simulated Discharge for the Calibration period
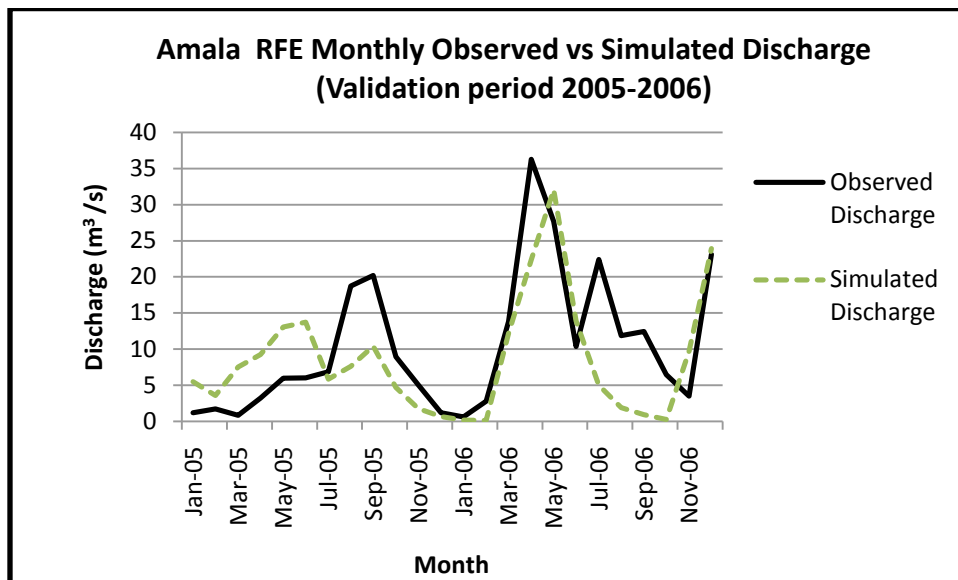
Figure 4-19. Amala RFE Monthly Observed vs Simulated Discharge for the Validation period
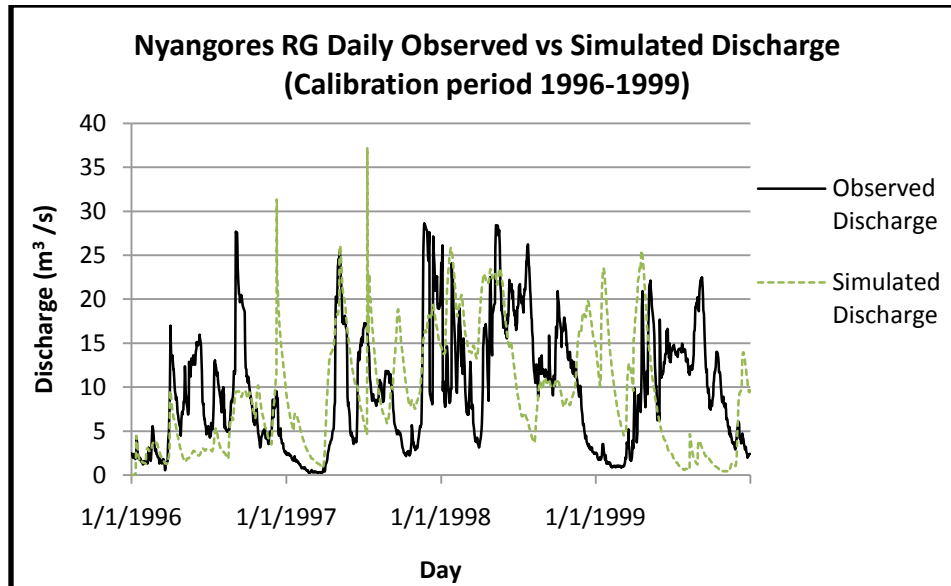
Figure 4-20. Nyangores Rain gauge Daily Observed vs Simulated Discharge for the Calibration period
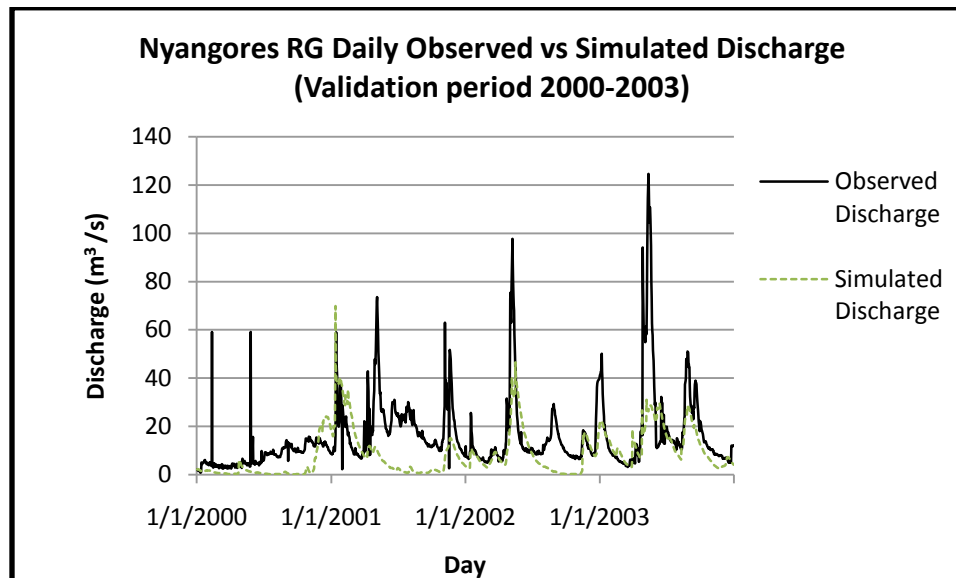
Figure 4-21. Nyangores Rain gauge Daily Observed vs Simulated Discharge for the Validation period
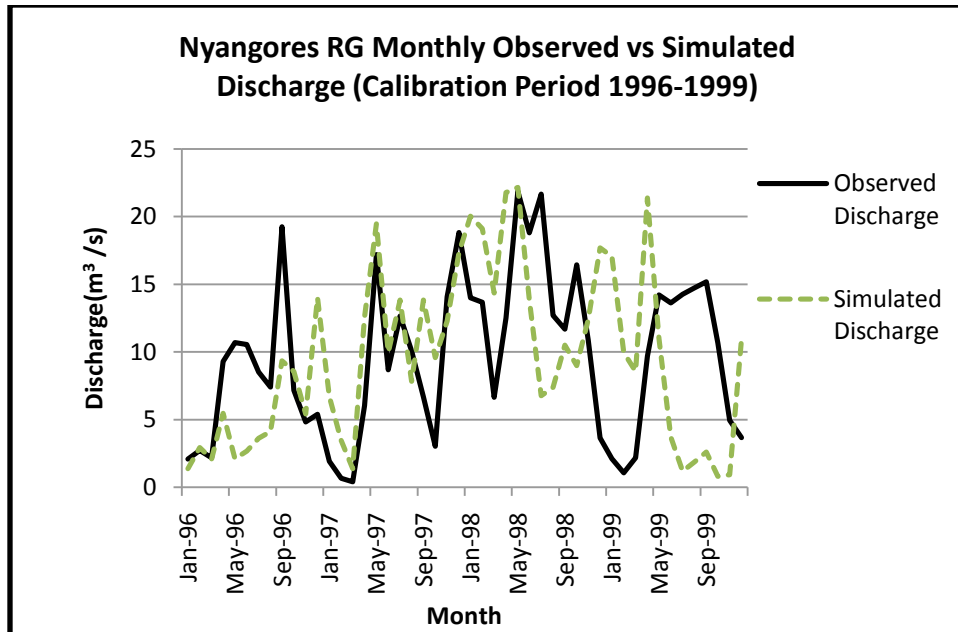
Figure 4-22. Nyangores Rain gauge Monthly Observed vs Simulated Discharge
for the Calibration period
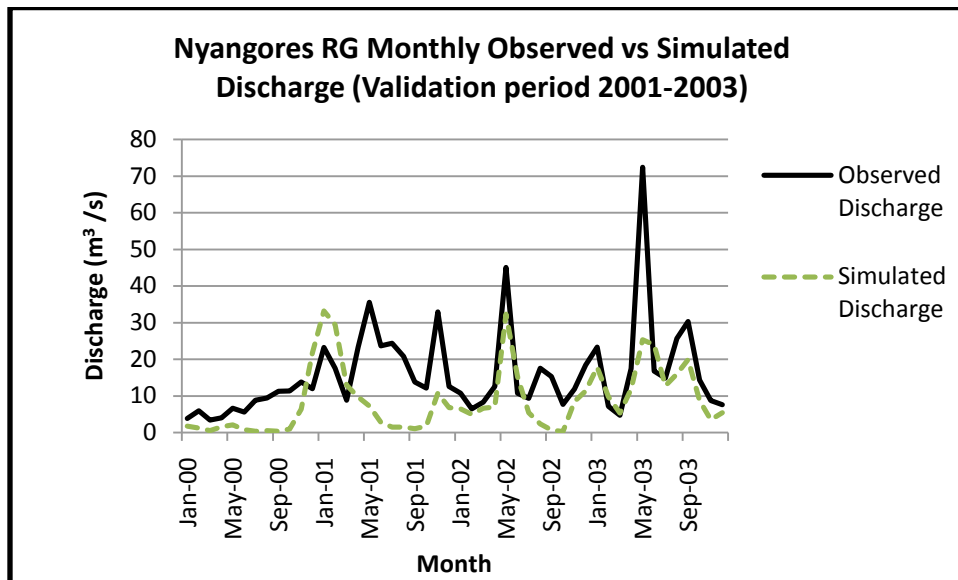
Figure 4-23. Nyangores Rain gauge Monthly Observed vs Simulated Discharge
for the Validation period

111
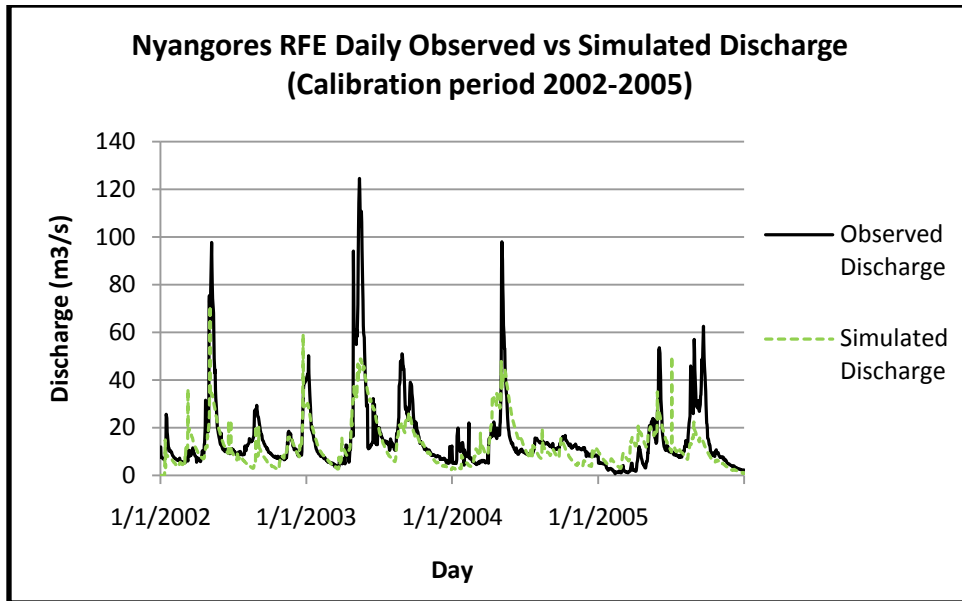
Figure 4-24. Nyangores RFE Daily Observed vs Simulated Discharge for the Calibration period
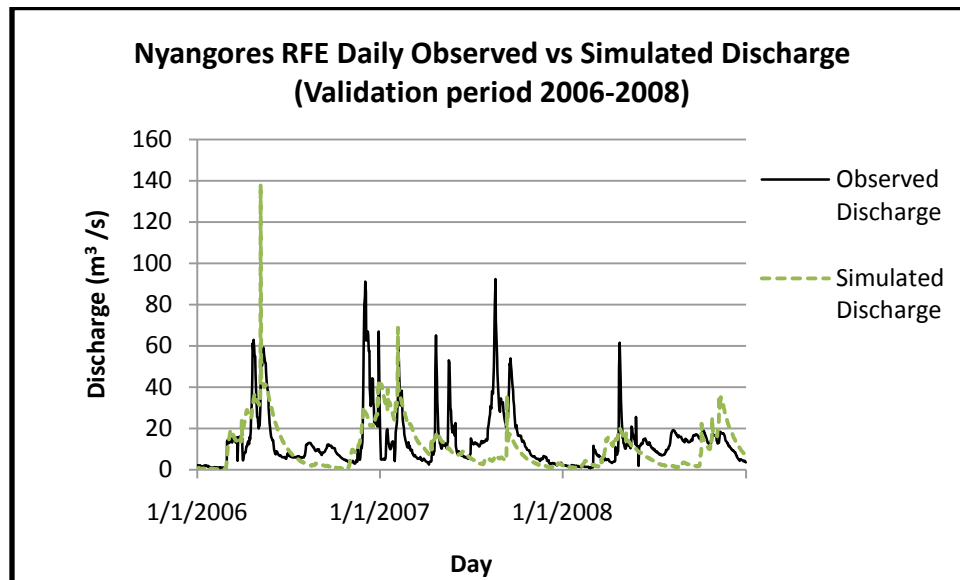
Figure 4-25. Nyangores RFE Daily Observed vs Simulated Discharge for the Validation period
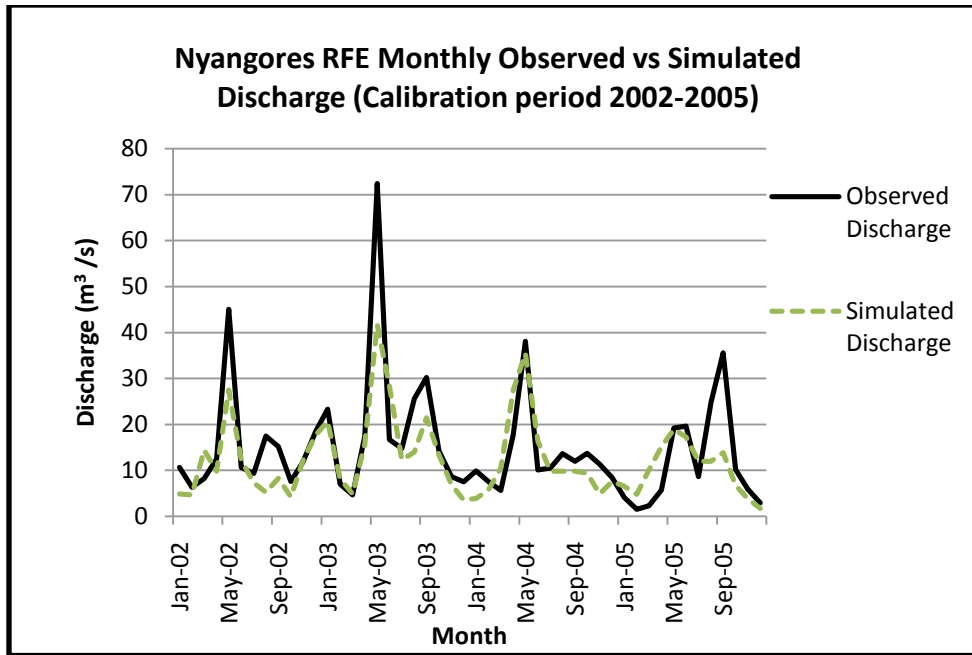
Figure 4-26. Nyangores RFE Monthly Observed vs Simulated Discharge for the Calibration period
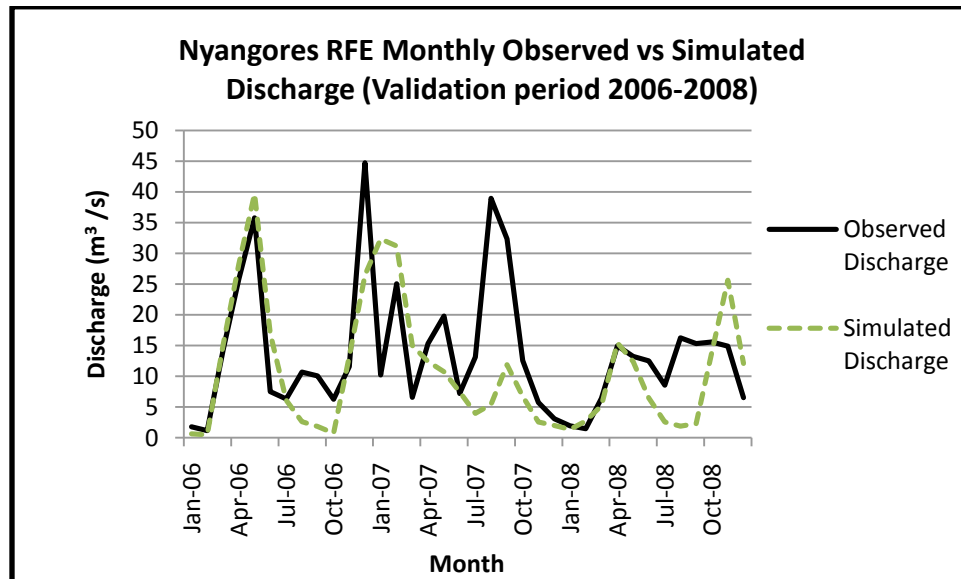
Figure 4-27. Nyangores RFE Monthly Observed vs Simulated Discharge for the Validation period

4.2.3.3 RFE vs Rain gauge Comparison

From the comparison below, the model simulations show a constant underestimation of discharge for the months of September and October which are the dry months of the year in the study area. This is contrary to the observed discharge that is much higher for the mentioned months, this may be attributed to ongoing surface-groundwater interactions or processes that are not captured by the model simulations and may be beyond the scope of this study.
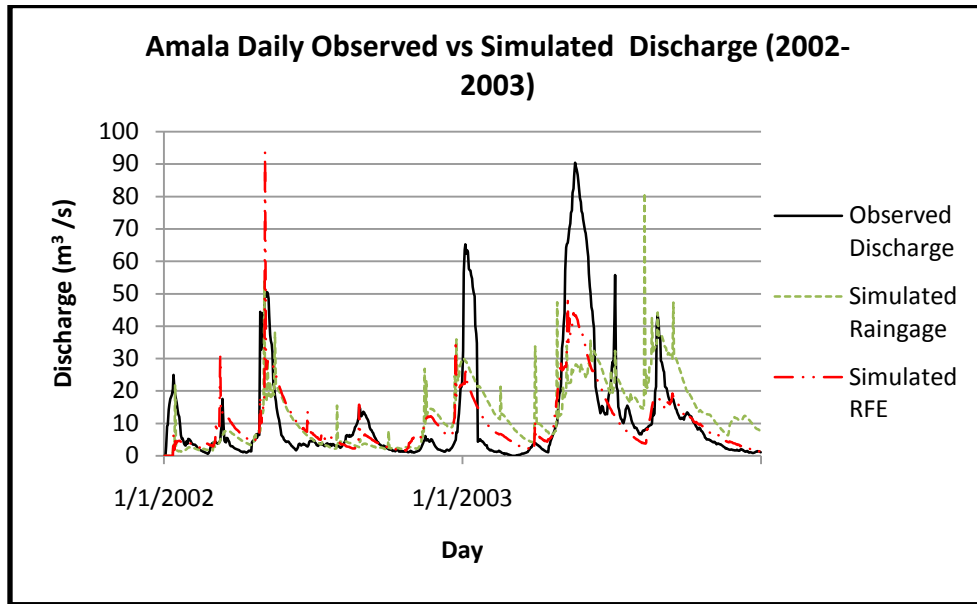
Figure 4-28. Amala Daily observed vs simulated Discharge
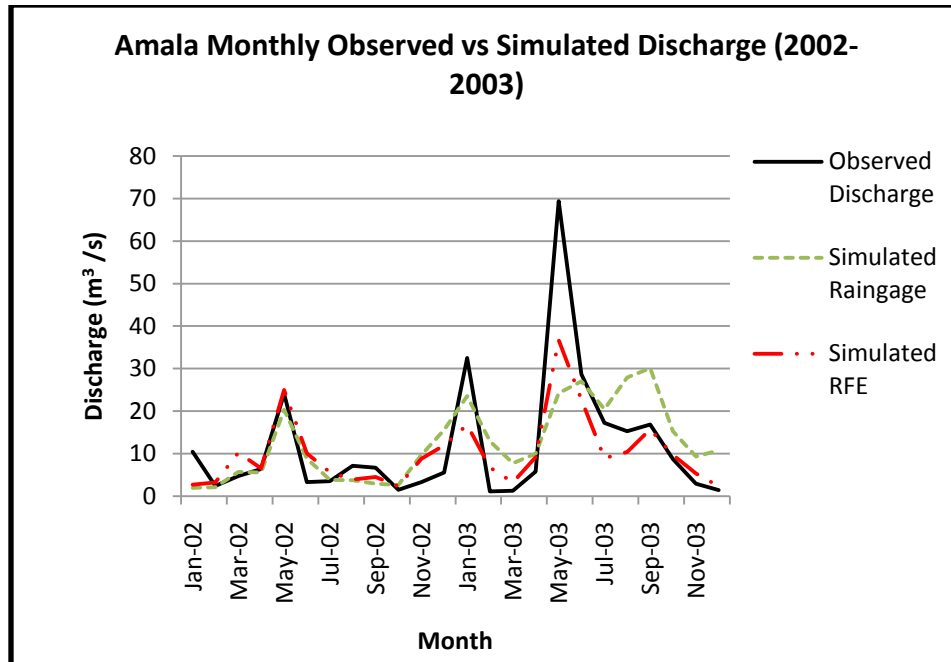
Figure 4-29. Amala Monthly observed vs simulated Discharge
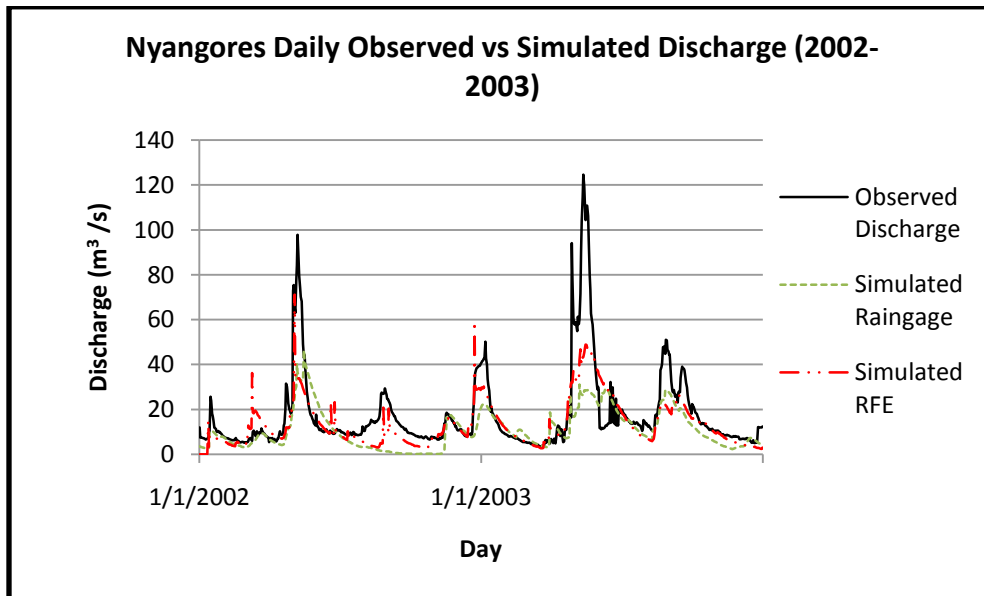
Figure 4-30. Nyangores Daily observed vs simulated Discharge
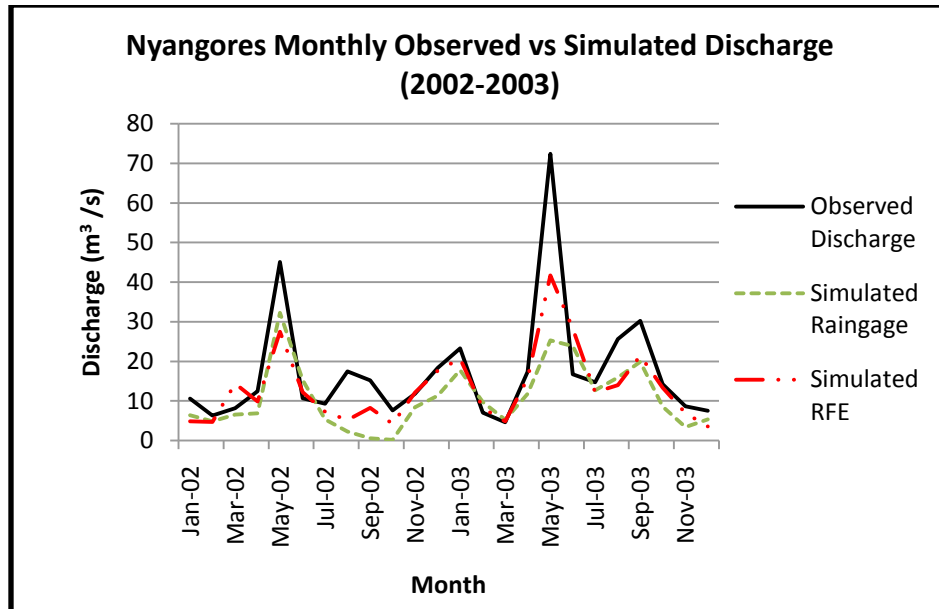
Figure 4-31. Nyangores Monthly observed vs simulated Discharge
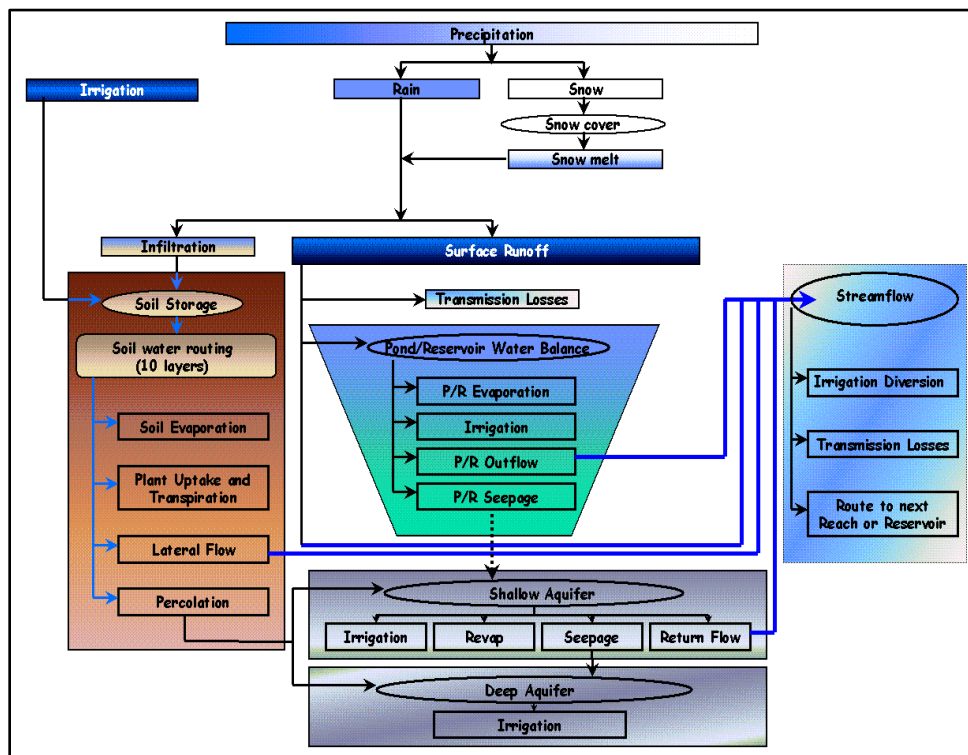
Source: SWAT2005 Manual

Figure 4-32. Schematic of pathways available for water movement in SWAT

The previous diagram shows the pathways that water moves through in the SWAT model and is helpful in interpretation of the resulting average annual water balance components for the basin under study.

Table 4-6 Annual Average Water Balance Components for the calibrated Amala watershed models

COMPONENTS	AMALA RG 2000-2003	AMALA RFE 2002-2006
PRECIP (mm)	1235.3	1009.8
SURQ (mm)	25.64	12.22
LATQ (mm)	37.68	30.99
GW_Q (mm)	367.85	394.26
REVAP (mm)	0	46.86
DA_RCHG (mm)	28.63	20.82
GW_ RCHG (mm)	572.62	416.41
WYLD (mm)	429.58	436.52
PERC (mm)	577.94	444.37
ET (mm)	581.1	508.8
PET (mm)	1258	1171.4
TLOSS (mm)	1.59	0.95
SEDYLD (T/HA)	0.641	0.307

PRECIP=Average total precipitation on sub basin (mmH$_2$0), PET=Potential evapotranspiration (mmH$_2$0), ET=Actual evapotranspiration (mmH$_2$0), PERC=Amount of water percolating out of the root zone (mmH$_2$0), SURQ=Surface runoff (mmH$_2$0), GW_Q=Groundwater discharge into reach or return flow (mmH$_2$0), WYLD=Net water yield to reach (mmH$_2$0), TLOSS=Amount of water removed from tributary channels by transmission (mmH$_2$0), DA RCHG= Amount of water entering deep aquifer from root

zone (mmH$_2$0), REVAP=Water in shallow aquifer returning to root zone (mmH$_2$0), GW_RCHG=amount of water entering both aquifers (mmH$_2$0), SEDYLD=Sediment yield (metric tons/ha), LATQ=Lateral flow contribution to reach (mmH$_2$0)

Table 4-7 Annual Average Water Balance Components for the calibrated Nyangores watershed models

COMPONENTS	NYANGORES RG 1996-2003	NYANGORES RFE 2002-2008
PRECIP (mm)	1329.9	1097.2
SURQ (mm)	15.03	11.51
LATQ (mm)	60.67	43.09
GW_Q (mm)	354.59	481.23
REVAP (mm)	21.89	3.48
DA_RCHG (mm)	22.47	25.33
GW_ RCHG (mm)	449.43	506.63
WYLD (mm)	429.28	535
PERC (mm)	450.02	509.52
ET (mm)	789	530
PET (mm)	1150.3	1179
TLOSS (mm)	1.01	0.82
SEDYLD (T/HA)	0.686	0.704

An examination of the water balance components for both the watersheds, the RFE models in both cases had lower rainfall averages than the Rain gauge models but simulated more net water yield to the reach (WYLD) than the Rain gauge models thus more total discharge. The surface runoff was consistently lower in the RFE models than in the Rain gauge models. This can be attributed to the higher rainfall experienced in the rain gauge models which makes more water available for conversion to surface runoff.

The RFE model on the other hand used 15 simulated stations that were able to capture the spatial variability of the precipitation across the watershed thus reducing the probability of overestimation of surface runoff in high rainfall events that may have been localized in one particular area.

These water balance components account for how the model partitions the available water input from precipitation and by converting them into percent increase and reductions provided an effective way to assess the effect of different climatic inputs and land use on the water balance of the Amala and Nyangores watersheds.

4.2.4 Scenario Analysis

4.2.4.1 Climate Change Scenarios

These scenarios were performed based on the Intergovernmental Panel on Climate Change (IPCC) projections for temperature and precipitation. These scenarios included air temperature increase by 5 % and precipitation increases and reductions by 10% and 20%. All scenarios shown below include a standard temperature increase of 5%.

4.2.4.1.1 Precipitation reduction by 10%

This involved the uniform reduction of the precipitation input by 10 percent to assess the effect the precipitation reduction coupled with a 5% increase in temperature will have on the discharge of the Amala and Nyangores rivers.
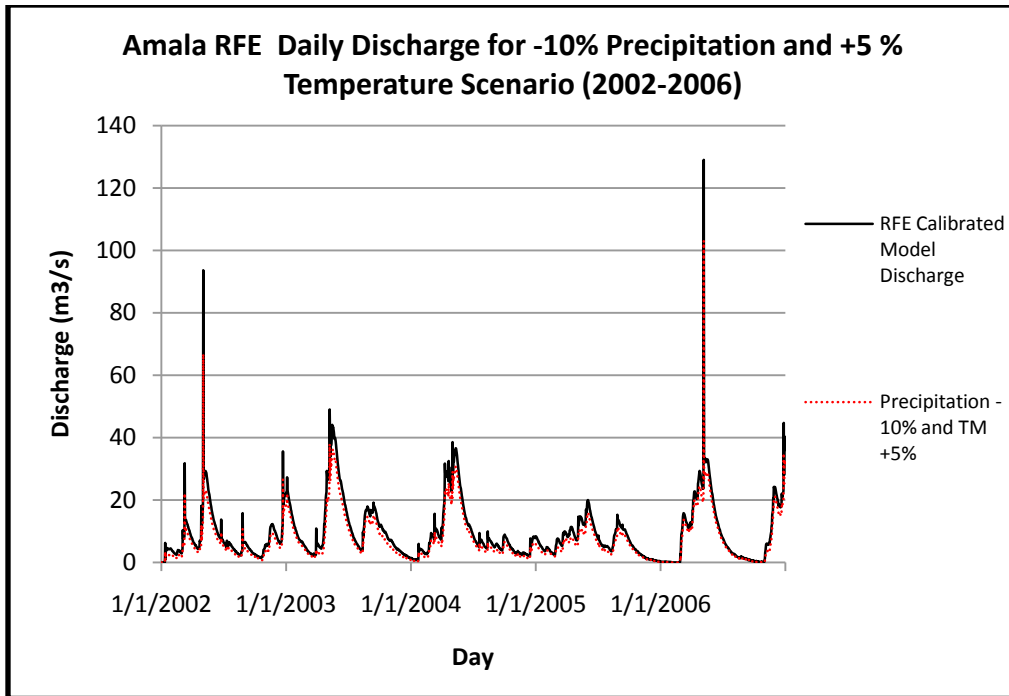
Figure 4-33. Amala RFE Daily Discharge for 10% Precipitation Reduction
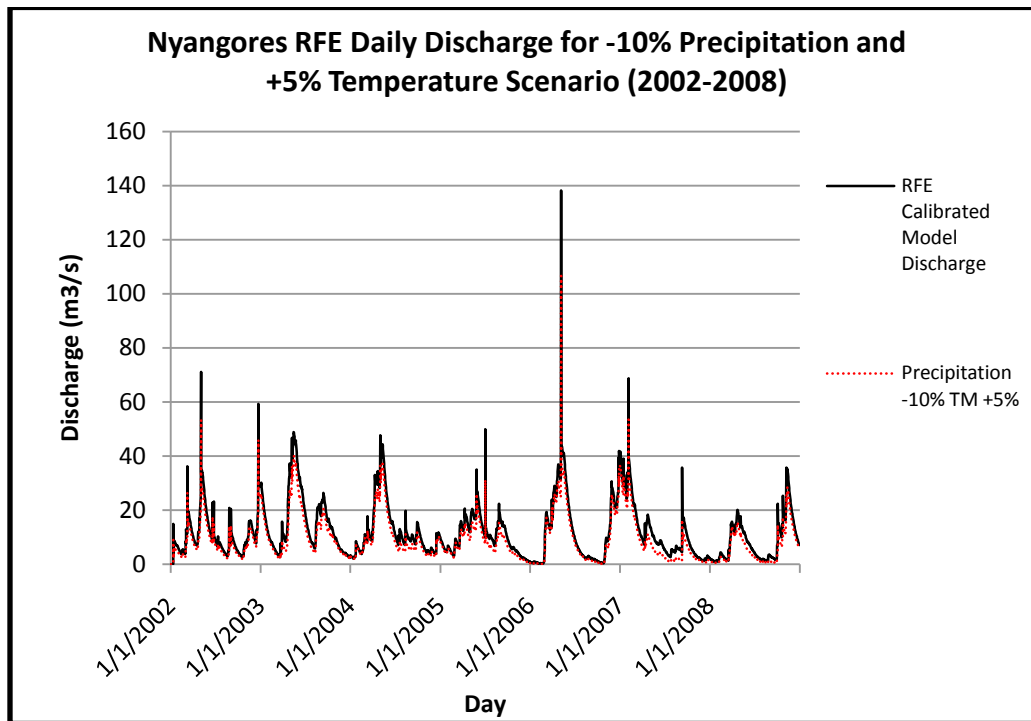
Figure 4-34. Nyangores RFE Daily Discharge for 10% Precipitation Reduction

120

From the hydrographs shown, it is evident that the reduction of precipitation caused an almost uniform decrease in discharge. This holds true for the peak flows and the low flows. The effect is true for both the Amala and Nyangores watersheds. How this reduction affects the water balance components in both watersheds can be seen in Tables 4-8 and 4-9 that follow in this chapter.

4.2.4.1.2 Precipitation reduction by 20%

For this scenario it is evident from the hydrographs that the combination of 20% reduction of precipitation and a 5% temperature increase reduced the discharge in both Amala and Nyangores rivers, the peaks were further reduced than in the previous scenario that simulated a reduction of 10% in precipitation, the baseflow was even further reduced almost to zero.
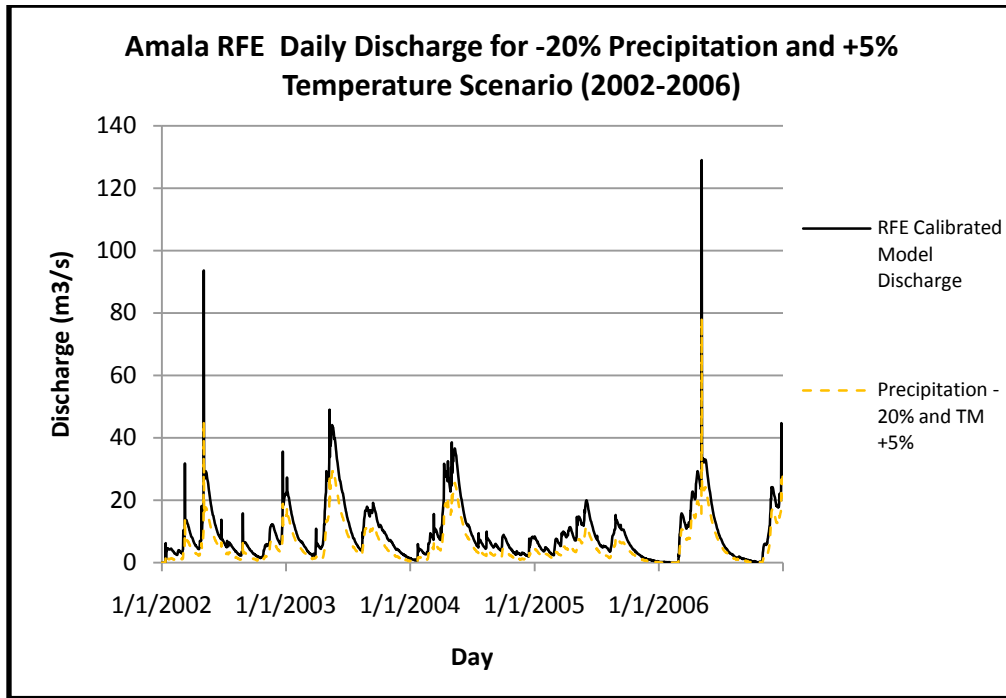
Figure 4-35. Amala RFE Daily Discharge for 20% Precipitation Reduction

121
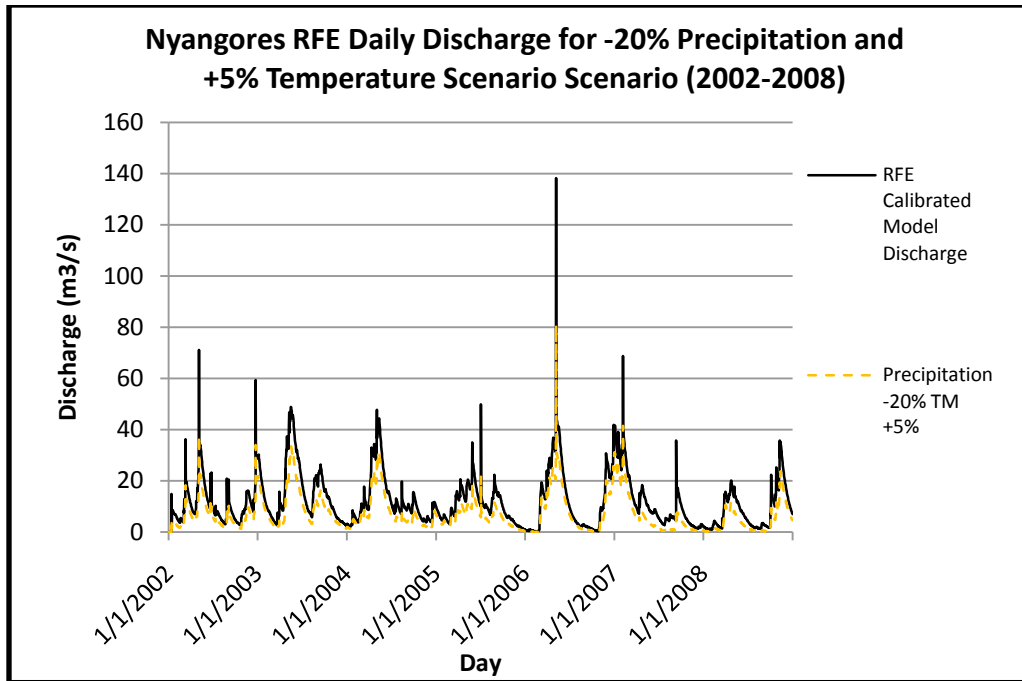
Figure 4-36. Nyangores RFE Daily Discharge for 20% Precipitation Reduction

These simulations underscore the effect of precipitation on the discharge and the direct effect reduction in the total amount received would have on the flow of the Amala and Nyangores rivers. A reduction of this magnitude in precipitation would reduce the baseflow to very low levels and depending on the water demand at that particular time of year, would be a significant threat to the maintenance of critical flows.

4.2.4.1.3 Precipitation Increase by 10 %

An increase in the precipitation by 10% would have a significant impact in both waterheds and this is seen directly in the discharge hydrographs where there are visible increases in the peaks and baseflows of the Amala and Nyangores rivers.
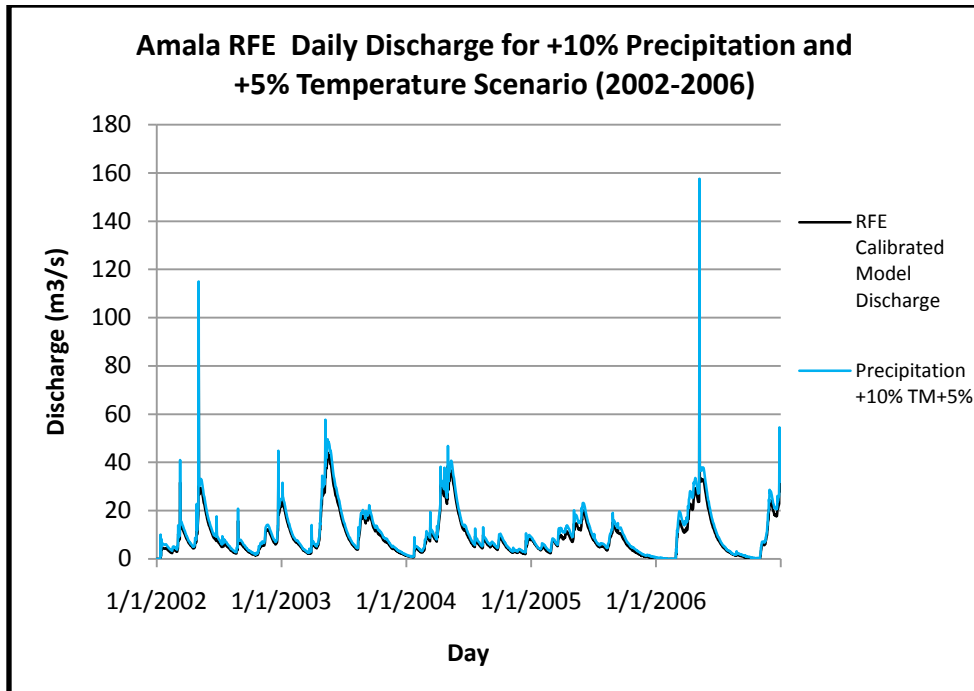
Figure 4-37. Amala RFE Daily Discharge for 10% increase in Precipitation
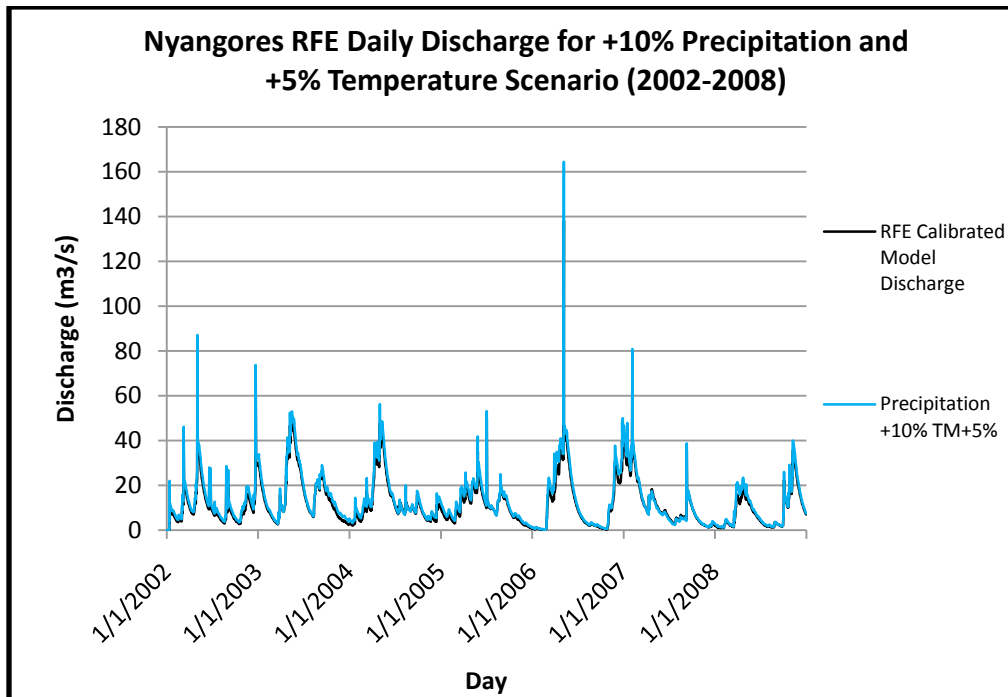
Figure 4-38. Nyangores RFE Daily Discharge for 10% increase in Precipitation

This would have a direct impact on the amount of water available within the two watersheds by increasing it and there would be increased floods and sediment yield from erosion as is observed in the Tables 4-8 and 4-9 that follow later in this chapter.

4.2.4.1.4 Precipitation Increase by 20%

From the hydrographs above, an increase in precipitation by 20% and temperature by 5% would increase the discharge in the Amala and Nyangores rivers with higher peaks during rainfall events and a generally higher baseflow in both rivers. This scenario of climate change would increase the amount of water available in the two watersheds for the water users in the Amala and Nyangores watersheds.
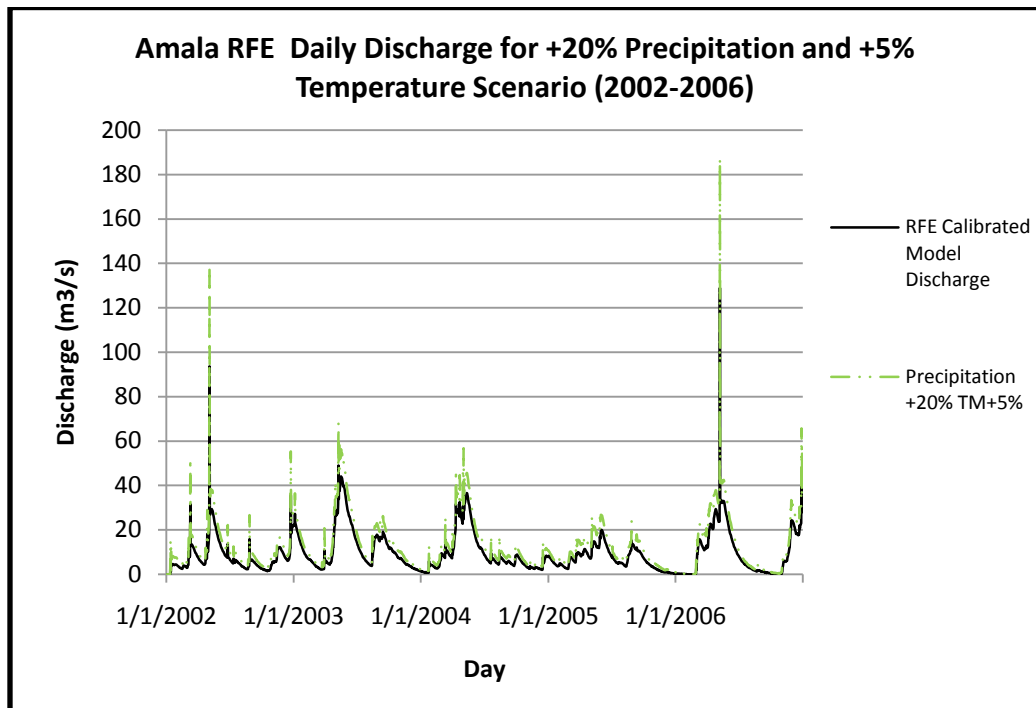
Figure 4-39. Amala RFE Daily Discharge for 20% increase in Precipitation
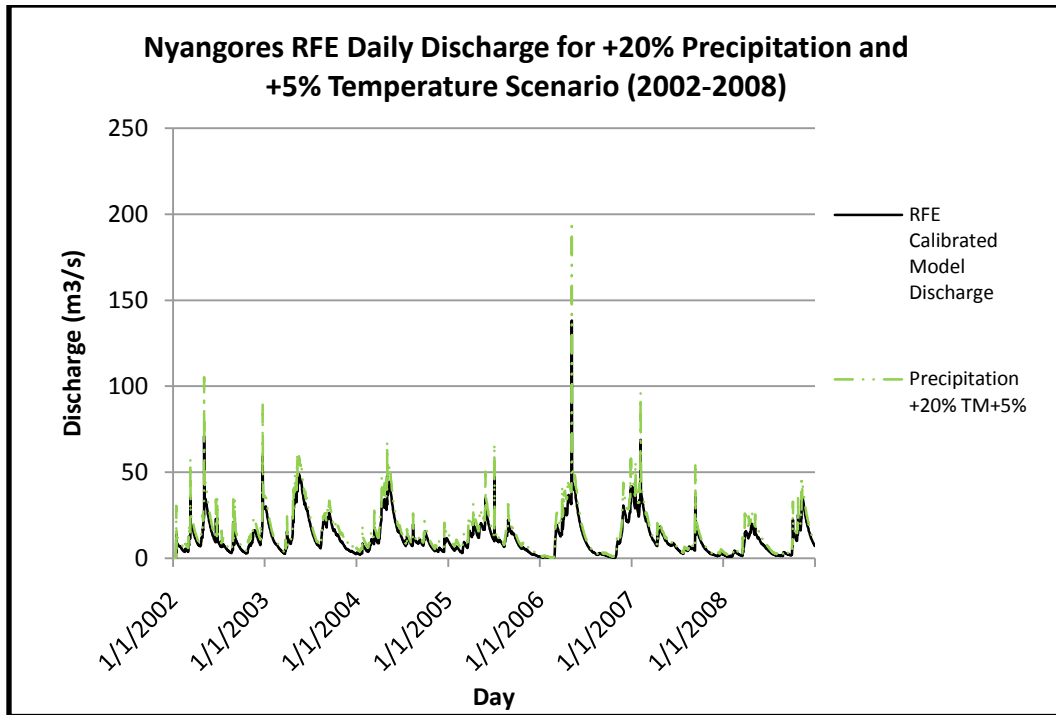
Figure 4-40. Nyangores RFE Daily Discharge for 20% increase in Precipitation

4.2.4.1.5 Complete Daily and Monthly Scenarios

The combined discharge hydrographs for all the climate change scenarios above help single out the impact a single climate change event however unlikely, would have on the discharge of the Amala and Nyangores rivers. The reduction of precipitation would bring about a reduction in available water in the watersheds reducing baseflows to very low levels that may be critical to the users of the water provided by these two rivers. This combined with the temperature increase would affect water availability also across land by increasing evapotranspiration that will ultimately affect the hydrology of the watershed. This can be seen in the tables that follow.
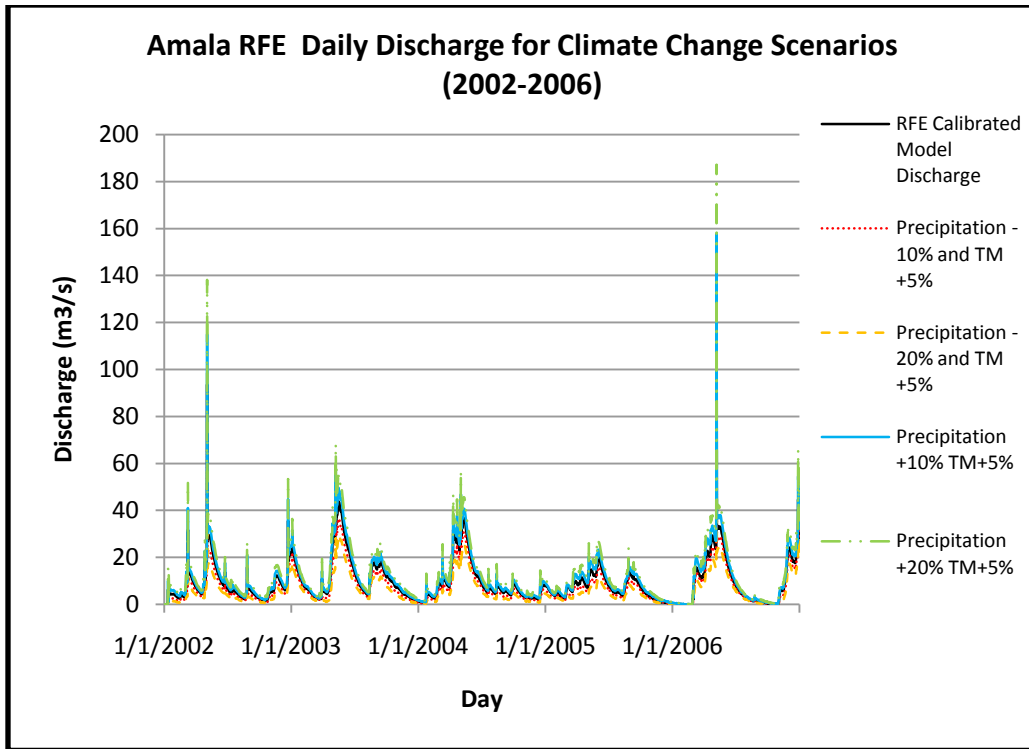
Figure 4-41. Amala Daily Discharge for Climate Change Scenarios
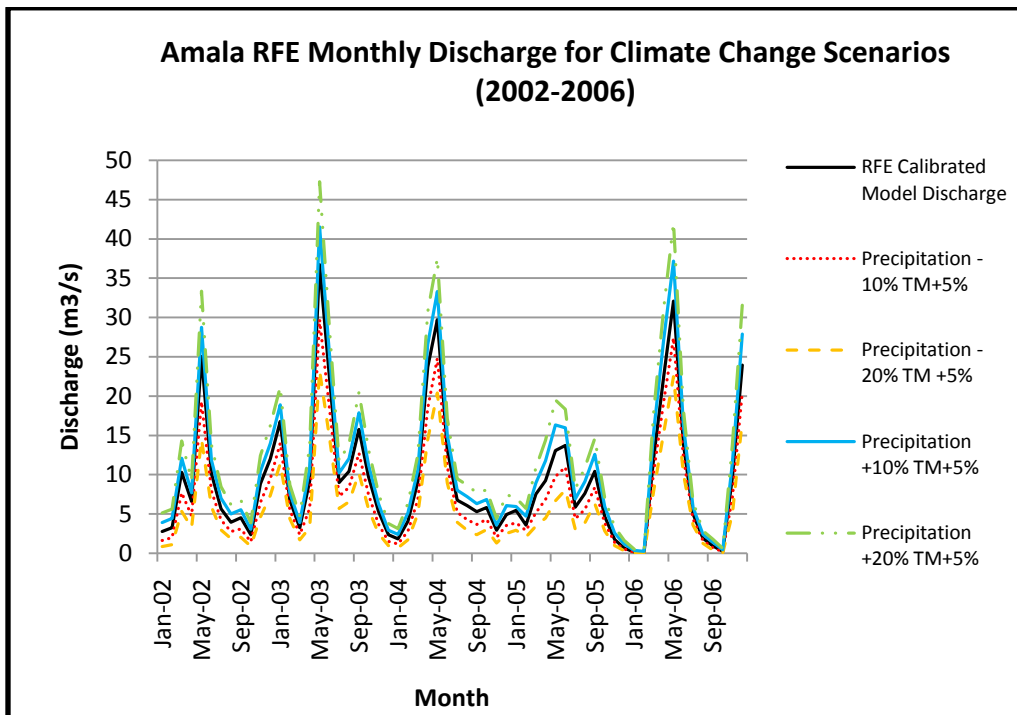
Figure 4-42. Amala Monthly Discharge for Climate Change Scenarios

126

Increases in precipitation by 10 percent and 20 percent would increase the discharge and baseflows in the rivers but on the other hand may have negative effects across land such as erosion and in the reach such as increased sediment load and flooding.
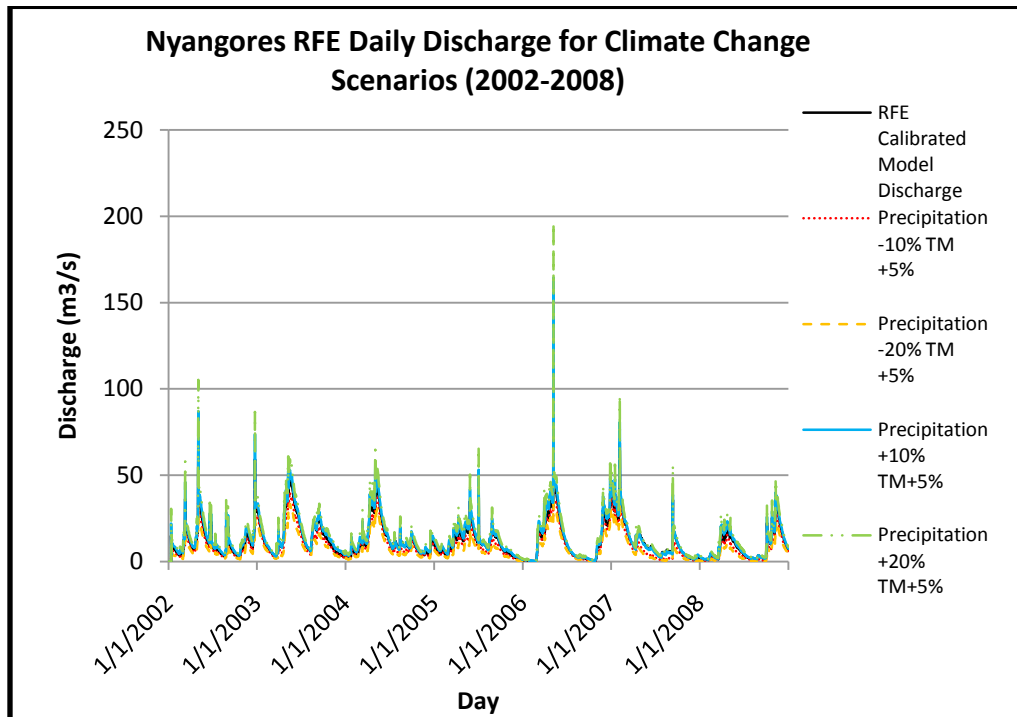
Figure 4-43. Nyangores Daily Discharge for Climate Change Scenarios
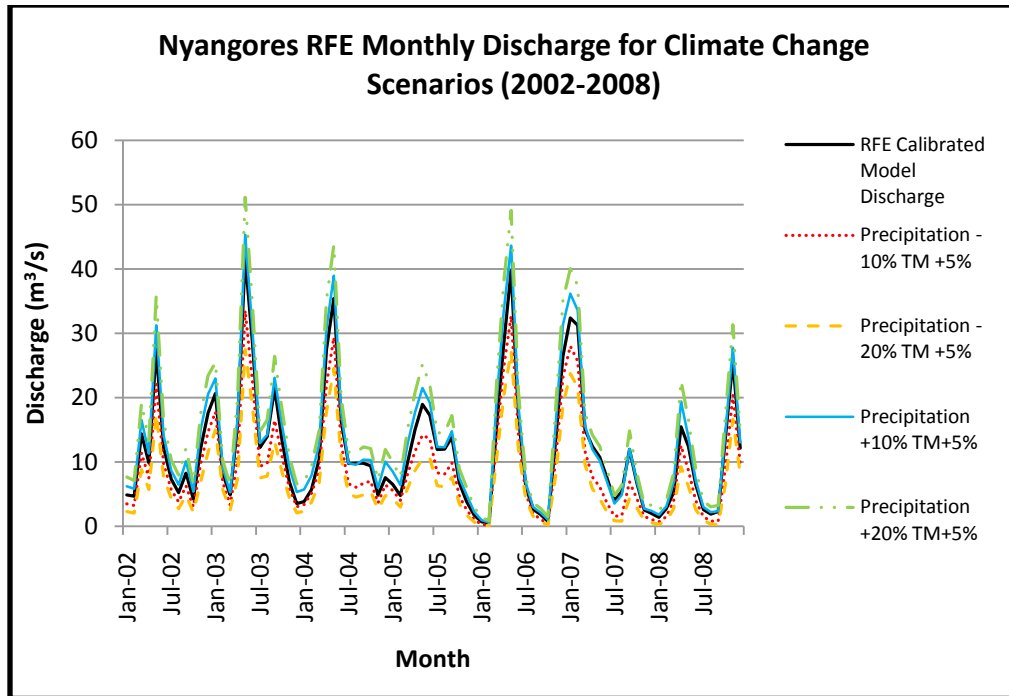
Figure 4-44. Nyangores Monthly Discharge for Climate Change Scenarios

From the monthly hydrographs, the monthly discharge values are constantly higher for the increased precipitation scenarios and are constantly lower for the reduced precipitation scenarios which are an indication that the discharge responds to the variations in precipitation directly and predictably meaning it is highly likely that a reduction or increase in annual precipitation will result in a reduction or increase in discharge of the rivers.

4.2.4.1.6 Annual Average Percent Changes in Water Balance Components

The figures 4-45 and 4-46 show the percent changes in the annual average water balance components for the climate change scenarios and their response in terms of amount. From the plots above it is evident that sediment yield is the most responsive followed by revap, surface runoff and transmission losses.

Figure 4-45. Annual Average Percent Changes for Amala Water Balance Components for Climate Change Scenarios

Figure 4-46. Annual Average Percent Changes for Nyangores Water Balance
Components for Climate Change Scenarios

All the water balance components vary in the a similar manner and from preliminary tests were found to vary linearly which can also be observed from the plots that are almost a mirror image of one another especially those that are of corresponding reduction/increase in precipitation.

Table 4-8 Percent Changes in Annual Averages of Amala Basin Water Balance Components for Climate Change Scenarios

COMPONENTS	AM PR-10	AM PR-20	AM TM+5	AM PR-10 TM+5	AM PR-20 TM+5	AM PR+10 TM5	AM PR+10 TM+5
PRECIP (mm)	-9.91	-20.01	0.00	-9.90	-20.00	10.06	20.02
SURQ (mm)	-34.62	-59.98	-2.95	-35.76	-60.80	40.67	92.72
LATQ (mm)	-15.46	-31.17	-2.23	-17.91	-33.43	13.39	29.07
GW_Q (mm)	-18.61	-37.17	-3.50	-21.80	-40.19	15.30	33.87
REVAP (mm)	-0.02	-0.02	3.12	3.14	3.14	3.14	3.14
DA_RCHG (mm)	-18.59	-37.13	-3.51	-21.76	-40.15	15.27	33.81
GW_ RCHG (mm)	-18.59	-37.13	-3.49	-21.77	-40.14	15.29	33.83
WYLD (mm)	-18.81	-37.35	-3.40	-21.90	-40.25	15.85	35.10
PERC (mm)	-18.13	-36.25	-3.32	-21.15	-39.10	14.97	33.00
ET (mm)	-2.06	-4.68	3.09	0.79	-1.99	5.07	6.84
PET (mm)	-0.01	-0.01	3.13	3.14	3.14	3.14	3.14
TLOSS (mm)	-29.47	-52.63	-2.11	-30.53	-53.68	29.47	65.26
SEDYLD (T/HA)	-44.63	-71.01	-4.89	-45.93	-71.66	62.21	148.21

PR10=Precipitation reduced by 10%, PR20= Precipitation reduced by 20%, TM5=Air Temperature increase by 5%

Table 4-9 Percent Changes in the Annual Averages of Nyangores Basin Water Balance Components for Climate Change Scenarios

COMPONENTS	NY PR-10	NY PR-20	NY TM5	NY PR-10TM5	NY PR-20TM5	NY PR+10TM5	NY PR+20TM5
PRECIP (mm)	-9.92	-19.98	0.00	-9.92	-19.98	10.07	19.98
SURQ (mm)	-34.14	-60.38	-5.04	-38.23	-63.51	37.71	90.79
LATQ (mm)	-14.97	-29.89	-2.32	-17.15	-31.84	12.90	27.80
GW_Q (mm)	-17.31	-34.50	-4.09	-21.29	-38.26	13.52	30.76
REVAP (mm)	-35.92	-63.51	-6.61	-41.38	-68.39	34.48	77.30
DA_RCHG (mm)	-17.29	-34.50	-4.07	-21.28	-38.26	13.54	30.75
GW_ RCHG (mm)	-17.31	-34.50	-4.09	-21.29	-38.26	13.52	30.76
WYLD (mm)	-17.47	-34.66	-3.97	-21.31	-38.26	13.97	31.77
PERC (mm)	-17.27	-34.42	-4.08	-21.24	-38.17	13.49	30.69
ET (mm)	-1.94	-4.49	3.91	1.83	-0.98	5.70	7.28
PET (mm)	0.00	0.00	3.15	3.15	3.15	3.15	3.15
TLOSS (mm)	-25.61	-50.00	-1.22	-28.05	-52.44	28.05	62.20
SEDYLD (T/HA)	-39.49	-67.05	-3.69	-43.47	-69.89	50.43	124.15

Table 4-10 Ratio of Water Balance Components to Precipitation for the Amala Basin Climate Change Scenarios

	AM RFE	AM PR_10	AM PR-20	AM TM+5	AM PR-10 TM+5	AM PR-20 TM+5	AM PR+10 TM+5	AM PR+20 TM+5
PREC (mm)	1	1	1	1	1	1	1	1
SURQ (mm)	0.012	0.009	0.006	0.012	0.009	0.006	0.015	0.019
LATQ (mm)	0.031	0.029	0.026	0.030	0.028	0.026	0.032	0.033
GW_Q (mm)	0.390	0.353	0.307	0.377	0.339	0.292	0.409	0.436
REVAP (mm)	0.046	0.052	0.058	0.048	0.053	0.060	0.043	0.040
DP AQ RCHRG (mm)	0.021	0.019	0.016	0.020	0.018	0.015	0.022	0.023
TOTAL AQ RCHRG (mm)	0.412	0.373	0.324	0.398	0.358	0.309	0.432	0.460
WYLD (mm)	0.432	0.390	0.339	0.418	0.375	0.323	0.455	0.487
PERC (mm)	0.440	0.400	0.351	0.425	0.385	0.335	0.460	0.488
ET (mm)	0.504	0.548	0.600	0.519	0.564	0.617	0.481	0.449
PET (mm)	1.160	1.288	1.450	1.196	1.328	1.496	1.087	0.997
TLOSS (mm)	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
SED (T/HA)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001

Table 4-11 Ratio of Water Balance Components to Precipitation for the Nyangores Basin Climate Change Scenarios

	NY RFE COMP	NY PR-10	NY PR-20	NY TM+5	NY PR-10 TM+5	NY PR-20 TM+5	NY PR+10 TM+5	NY PR+20 TM+5
PREC (mm)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SURQ (mm)	0.010	0.008	0.005	0.010	0.007	0.005	0.013	0.017
LATQ (mm)	0.039	0.037	0.034	0.038	0.036	0.033	0.040	0.042
GW_Q (mm)	0.439	0.403	0.359	0.421	0.383	0.338	0.452	0.478
REVAP (mm)	0.003	0.002	0.001	0.003	0.002	0.001	0.004	0.005
DP AQ RCHRG (mm)	0.023	0.021	0.019	0.022	0.020	0.018	0.024	0.025
TOTAL AQ RCHRG (mm)	0.462	0.424	0.378	0.443	0.403	0.356	0.476	0.503
WYLD (mm)	0.488	0.447	0.398	0.468	0.426	0.376	0.505	0.536
PERC (mm)	0.464	0.426	0.381	0.445	0.406	0.359	0.479	0.506
ET (mm)	0.483	0.526	0.577	0.502	0.546	0.598	0.464	0.432
PET (mm)	1.075	1.193	1.343	1.108	1.230	1.385	1.007	0.924
TLOSS (mm)	0.001	0.001	0.000	0.001	0.001	0.000	0.001	0.001
SED (T/HA)	0.001	0.000	0.000	0.001	0.000	0.000	0.001	0.001

From Table 4-8 and 4-9 shown previously, it is evident that any reduction in precipitation resulted in a reduction of SURQ, LATQ, GW_Q, DA_RCHG, WYLD, PERC, TLOSS and SEDYLD. REVAP was more or less constant for all the simulations. Evapotranspiration reduced steadily with a reduction in precipitation but increased in the simulation that involved increasing the air temperature by 5 percent while keeping all the other inputs constant. The PET was affected by the temperature and minimally by the amount of rainfall but mainly by the maximum and minimum temperature which increases the PET with an increase and vice-versa. TLOSS was also affected by the amount of precipitation where it is high in cases of high precipitation and reduced with a reduction in the same. The high precipitation increased the flow in the reach which increased the water available to be lost and vice versa.

Naturally, the water balance is driven by precipitation and this is shown by the number of components that respond to a reduction in precipitation as exhibited in the Tables 4-8 and 4-9 above. This is an indication that climate change, more importantly precipitation, will have a significant impact on the water balance of the Amala and Nyangores watersheds which will be manifested in the discharge hydrographs of the Amala and Nyangores rivers as shown by the model simulations above.

4.2.4.2 Land Use Change Scenarios

4.2.4.2.1 Partial Deforestation
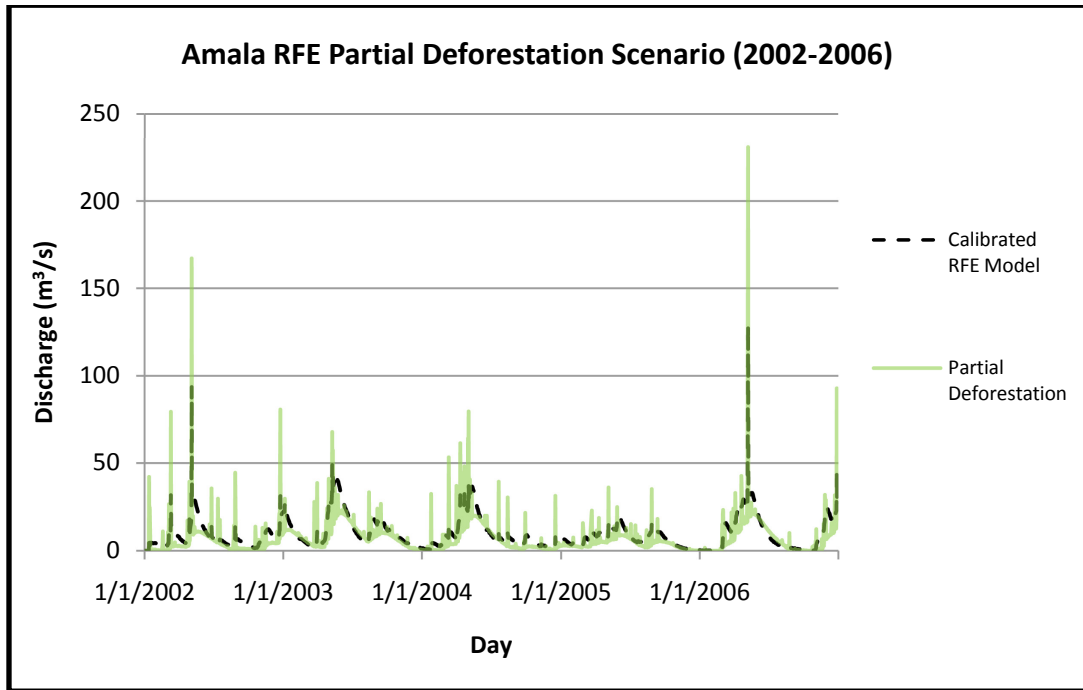
Figure 4-47. Amala RFE Partial Deforestation Scenario

From the resultant hydrographs above, it is seen that the peaks in the partial deforestation hydrograph are more in number and also higher in magnitude when compared to the calibrated model the lower dips in the hydrographs are indicative of lower baseflows in the drier periods. The higher peaks and reduced baseflow result from increased surface runoff due to the removal of forest cover that slows it down and facilitates infiltration into the soil during rainfall events leading to reduced recharge of groundwater that finds its way back to the reach or stream channel as return flow. This is confirmed by the values in table 4-12 and 4-13 that follow later in this chapter.
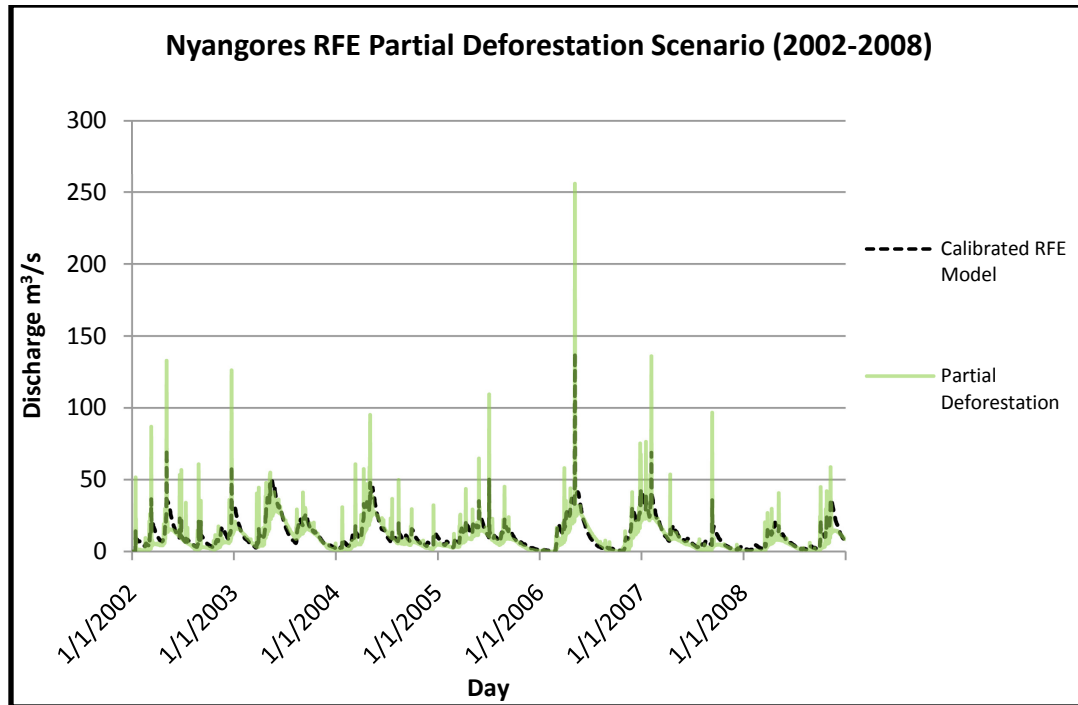
Figure 4-48. Nyangores RFE Partial Deforestation Scenario

From the tables 4-12 and 4-13, both watersheds recorded an increase in surface runoff (3% in the Amala watershed and 6.9% in the Nyangores watershed) which was attributed to the reduction in forest cover that reduces the initial abstractions from the rainfall received by the watershed which increases the amount of rainfall getting to the ground surface to be converted to runoff. There was also an increase in evapotranspiration, transmission losses which increased as a result of the increased flow of the river. Sediment yield also increased as a result of the increased runoff which the model uses in the simulation of erosion and sediment yield.

REVAP is determined by the soil temperature and land cover particularly deep rooted plants that remove water from the shallow aquifer and in this case reduces because of the partial reduction of forest cover in this scenario, The rest of the components (LATQ,

GW_Q, DA RCHG, GW_RCHG, PERC) recorded a decrease in amount as they depend on the amount of water infiltrating and penetrating to the shallow and deep aquifer which was slightly reduced because of the reduction in forest which increased the conversion of precipitation to runoff in both watersheds. This shows that the amount of water in the system more than anything determines the amounts that will be distributed among the different pathways of the water balance components.

4.2.4.2.2 Complete Deforestation
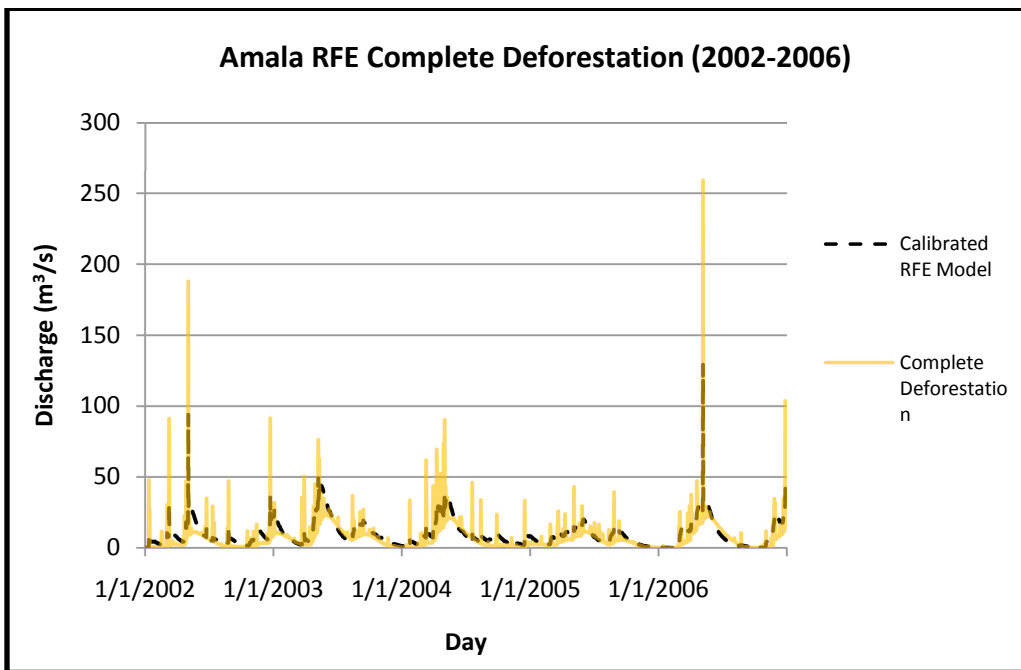
Figure 4-49. Amala RFE Complete Deforestation
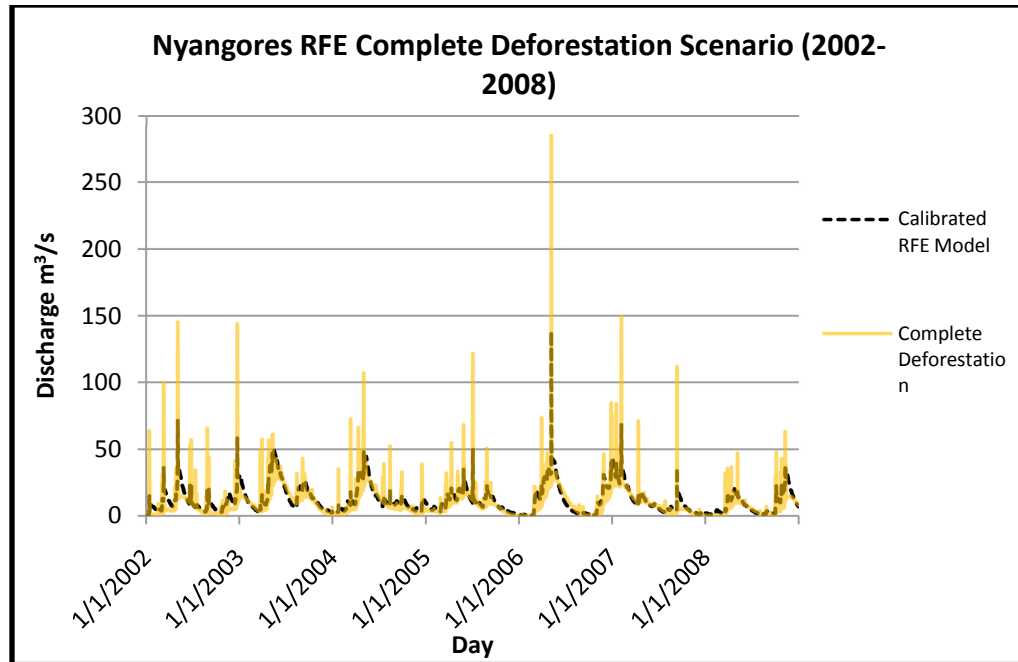
Figure 4-50. Nyangores RFE Complete Deforestation

From the hydrographs shown, it is evident that there are increased peaks both in number and magnitude that are even higher than the partial deforestation scenario. Reduced flows in the dry periods show a decrease in baseflow. This is a result of the reduced infiltration and increased runoff during rainfall events that contribute to the channel flow. The low baseflow is an effect of the reduced groundwater recharge.

Surface runoff (SURQ) increased and was a lot more (Amala 12.4%, Nyangores 20.57%) than in the partial deforestation scenario (Amala 3%, Nyangores 6.9%) this is due to the absence of forest cover which was replaced by grassland. The net water yield to the reach (WYLD) also increased minimally (0.25% in the Amala and 2.93%) in the Nyangores watersheds and this can be attributed to the increased surface runoff that makes its way to the reach. This scenario also had an increase in TLOSS which naturally

139

increased with an increase in water amount in the river channel because there was more water to be lost. SEDYLD also increased because of the loss of forest cover and increased surface runoff. Total water yield to the stream reach also increased as a result of the increased runoff which had percentages of 12% (Amala) and (20%) Nyangores.

Evapotranspiration decreased in the Nyangores watershed and this was seen to have been caused by reduction of forest cover specifically evergreen forest that is a large source of transpiration and evaporation of rainfall intercepted by the canopy.

There was a decrease in the other water balance components. REVAP decreases because of the absence of forest cover which significantly reduced the amount of deep rooted plants in the watershed that could remove water from the shallow aquifer. The LATQ, GW_Q, DA_RCHG, GW_RCHG and PERC all recorded a decrease in amount due to a decrease in the amount of water making its way to the shallow and deep aquifer as groundwater recharge.

4.2.4.2.3 Replacement of Forest by Agriculture

Replacement of forest land by agriculture is a common trend within the study area and is seen to be one of the major causes of extreme high and low river flows and increased sediment load in the Nyangores and Amala Rivers.
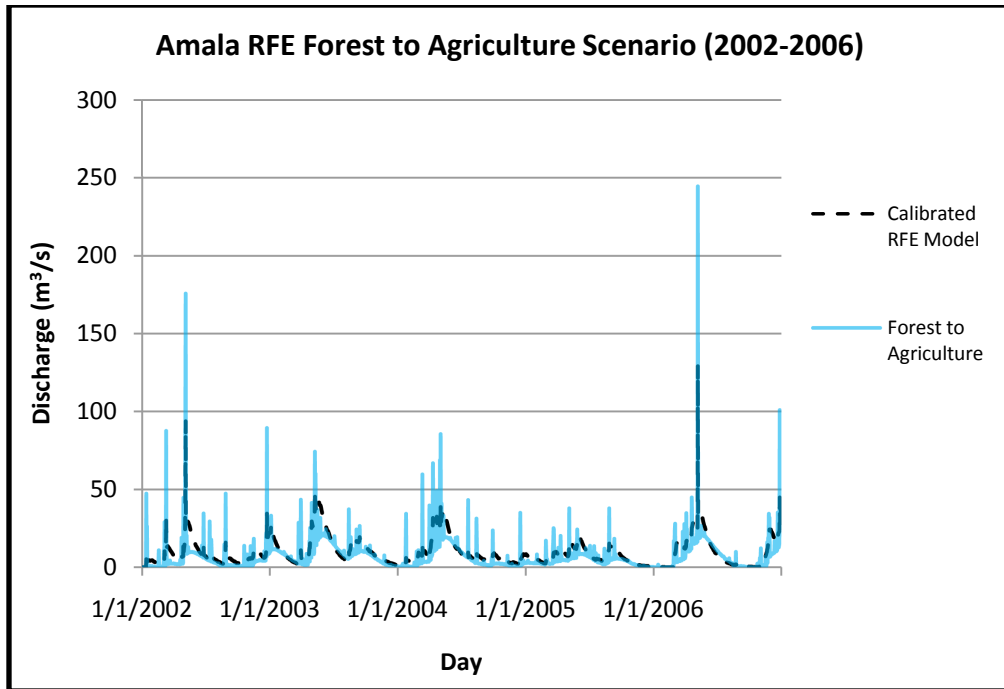
Figure 4-51. Amala RFE Forest replaced by Agriculture
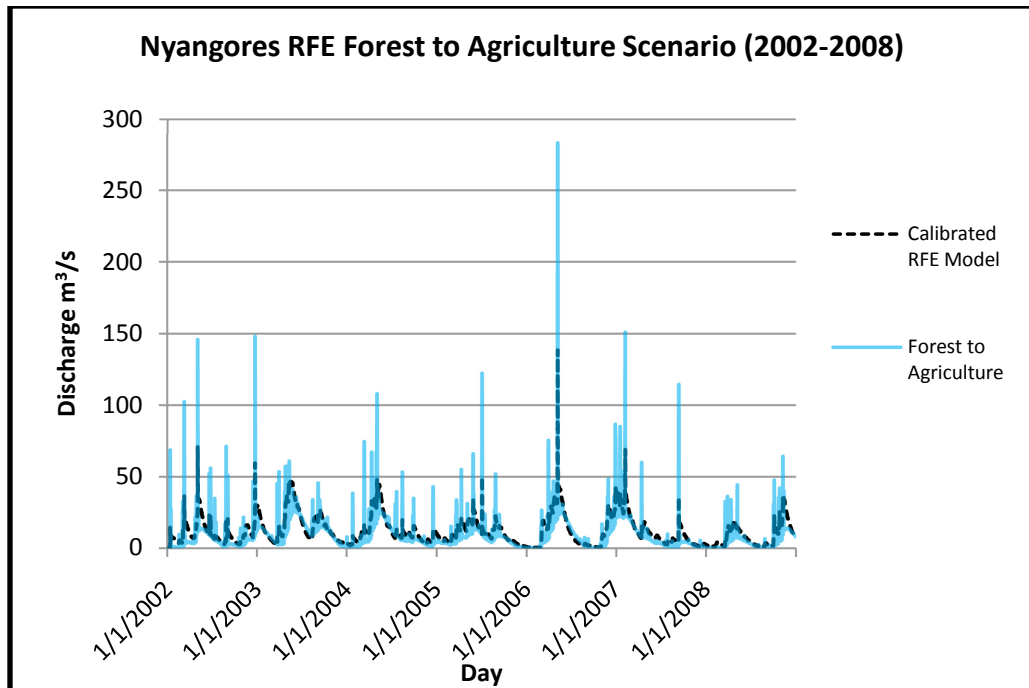
Figure 4-52. Nyangores RFE Forest replaced by Agriculture

The hydrographs reveal some of the effects that this change would have on the watersheds. Increased peaks of up to 10 m^3 in some cases are an indication that there will be increased runoff during rainfall events and the low flows also indicate that there will be a significant reduction in the baseflow of the Amala and Nyangores rivers during the dry periods causing a reduction in water yield of both rivers.

This scenario showed the greatest increase in the amount of surface runoff (Amala 13.7% and Nyangores 31.49%). There was also an increase in evapotranspiration for both watersheds (3.05% Amala and 2.46% Nyangores) which resulted from the change of land cover to agriculture. Transmission losses also increased in both watersheds as a result of the increased runoff and water in the reach. Sediment yield increased significantly in both watersheds due to the land cover change that increased the surface runoff hence erosion of the soil. A reduction in LATQ, GW_Q, REVAP, DA RCHG, GW_RCHG, WYLD and PERC indicated a reduction in water penetration to the shallow and deep aquifers of the Amala and Nyangores watersheds.
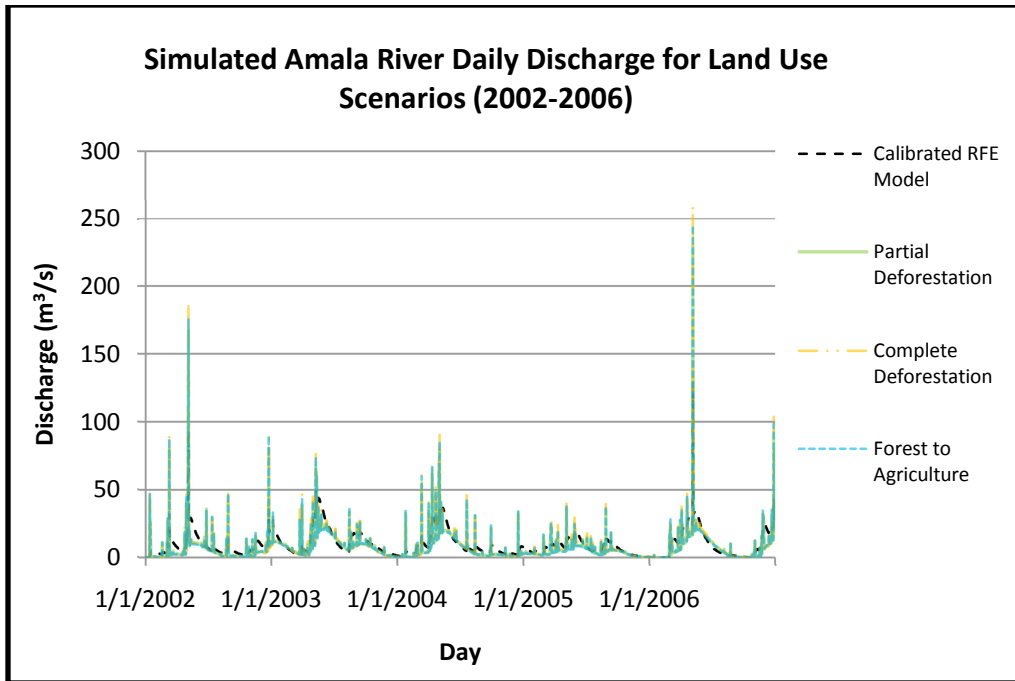
Figure 4-53. Simulated Amala River daily discharge for different land use scenarios
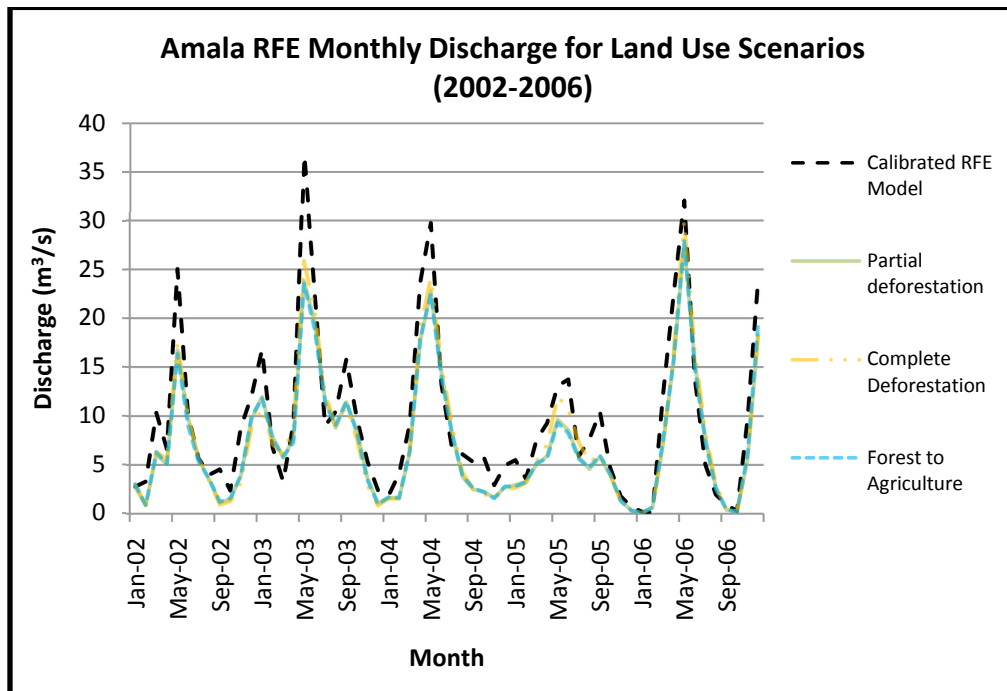
Figure 4-54. Simulated Amala River monthly discharge for different land use scenarios

143
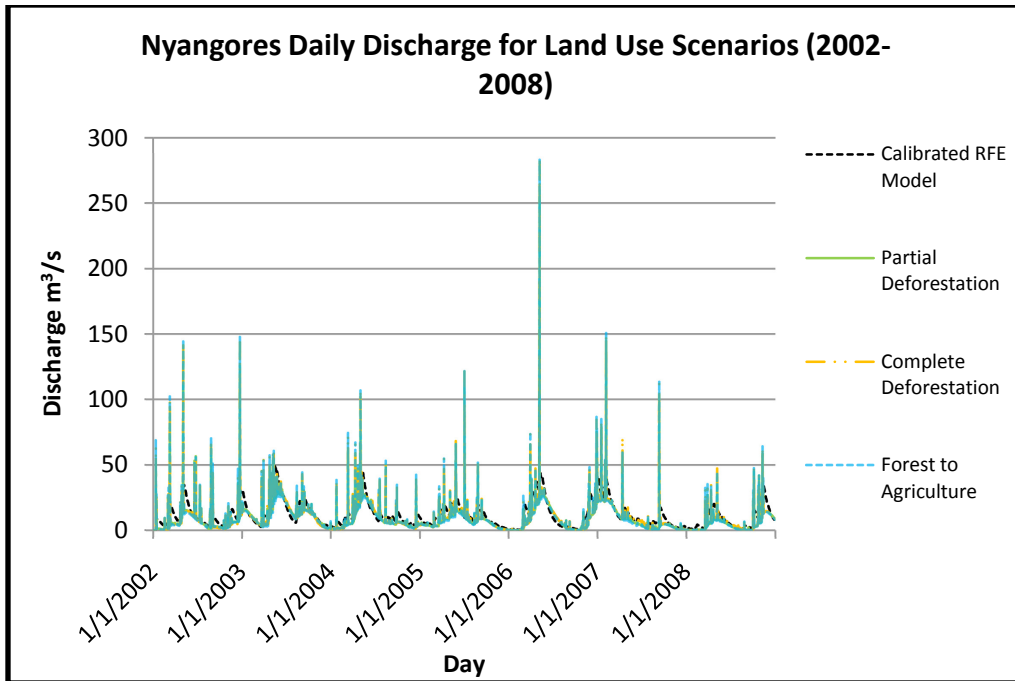
Figure 4-55. Simulated Nyangores River daily discharge for different land use scenarios
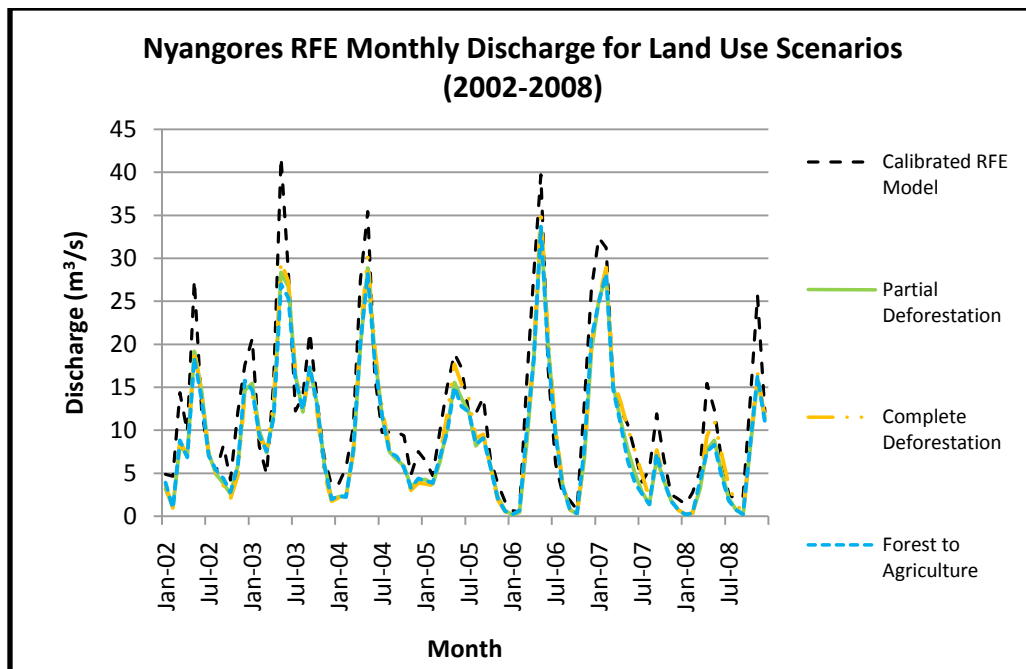
Figure 4-56. Simulated Nyangores River monthly discharge for different land use

scenarios

From the monthly hydrographs for the land use change scenarios above, it is clear that all the scenarios result in reduced monthly discharge in both the Amala and Nyangores rivers. In terms of the type of land use change and its effect, the monthly averages are consistent with the daily hydrographs where the conversion of forest land to agricultural land would have the most impact on the watersheds by reducing the output of both Amala and Nyangores rivers the most.

Table 4-12 Percent Changes in the Annual Averages of Amala Basin Water Balance Components for Land Use-Climate Change Scenarios

	AM PD	AM CD	AM FA	AM PD PR-10 TM+5	AM PD PR-20 TM+5	AM CD PR-10 TM+5	AM CD PR-20 TM+5	AM FA PR-10 TM+5	AM FA PR-20 TM+5
PRECIP (mm)	0.00	-9.91	-20.00	0.00	-9.91	-20.00	0.00	-9.91	-20.00
SURQ (mm)	3.15	-27.00	-51.39	12.40	-19.36	-45.38	13.70	-20.00	-47.15
LATQ (mm)	-2.55	-21.98	-37.71	-2.70	-19.50	-34.74	0.80	-17.90	-34.13
GW_Q (mm)	-6.28	-33.22	-53.73	-3.54	-27.94	-48.25	-10.51	-35.82	-55.77
REVAP (mm)	-2.51	-16.93	-31.74	-2.73	-15.03	-29.31	-4.02	-17.08	-32.27
DA_RCHG (mm)	-6.03	-32.13	-52.28	-3.60	-27.35	-47.28	-10.07	-34.56	-54.19
GW_ RCHG (mm)	-5.99	-32.12	-52.29	-3.59	-27.32	-47.28	-10.02	-34.56	-54.21
WYLD (mm)	-3.88	-31.14	-52.27	0.25	-25.47	-46.82	-4.28	-31.18	-52.56
PERC (mm)	-5.95	-31.19	-50.79	-3.93	-26.94	-46.39	-10.41	-34.00	-52.96
ET (mm)	2.57	3.30	-0.05	0.24	-0.02	-3.23	3.05	3.45	0.20
PET (mm)	0.00	3.14	3.14	0.00	3.14	3.14	0.00	3.14	3.14
TLOSS (mm)	10.83	-9.13	-29.51	14.86	-5.10	-25.48	38.00	12.95	-13.16
SEDYLD (T/HA)	15.72	-19.79	-49.43	14.38	-24.31	-52.79	55.02	5.51	-34.59

PD=Partial Deforestation, CD=Complete Deforestation, FA=Forest replaced by Agriculture, AM=Amala Watershed

Table 4-13 Percent Changes in the Annual Averages of Nyangores Basin Water Balance Components for Land Use-Climate Change Scenarios

	NY PD	NY CD	NY FA	NY PD PR-10 TM+5	NY PD PR-20 TM+5	NY CD PR-10 TM+5	NY CD PR-20 TM+5	NY FA PR-10TM+5	NY FA PR-20 TM+5
PRECIP (mm)	0.00	-9.92	-19.98	0.00	-9.92	-19.98	0.00	-9.92	-19.98
SURQ (mm)	6.94	-24.08	-48.85	20.58	-13.61	-41.47	31.49	-5.99	-36.31
LATQ (mm)	-1.56	-20.65	-35.94	-0.66	-17.55	-32.60	-14.91	-30.31	-43.44
GW_Q (mm)	-4.16	-30.49	-49.05	-0.75	-24.38	-43.01	-9.39	-33.28	-51.00
REVAP (mm)	-1.84	-12.76	-26.76	-0.89	-10.03	-23.28	-2.80	-12.76	-27.37
DA_RCHG (mm)	-4.04	-29.68	-48.09	-0.75	-23.79	-42.21	-9.06	-32.33	-50.00
GW_ RCHG (mm)	-4.06	-29.71	-48.12	-0.79	-23.78	-42.21	-9.09	-32.35	-49.99
WYLD (mm)	-2.00	-28.50	-47.85	2.93	-21.96	-41.88	-3.00	-28.49	-47.98
PERC (mm)	-4.22	-29.92	-48.36	-1.10	-24.07	-42.49	-9.71	-32.91	-50.50
ET (mm)	1.64	3.46	0.23	-1.93	-1.33	-4.20	2.46	3.50	0.35
PET (mm)	0.00	3.15	3.15	0.00	3.15	3.15	0.00	3.15	3.15
TLOSS (mm)	9.90	-10.15	-29.21	23.76	1.49	-19.80	45.30	20.30	-4.46
SEDYLD (T/HA)	22.39	-12.37	-42.73	11.90	-22.78	-50.15	41.48	1.19	-33.21

NY=Nyangores Watershed

Table 4-14 Ratio of Water Balance Components to Precipitation for the Amala Basin Land Use-Climate Change Scenarios

	AM RFE	AM PD	AM CD	AM FA	AM PD PR-10 TM+5	AM CD PR-10 TM+5	AM FA PR-10 TM+5	AM PD PR-20 TM+5	AM CD PR-20 TM+5	AM FA PR-20 TM+5
PREC (mm)	1	1	1	1	1	1	1	1	1	1
SURQ (mm)	0.085	0.088	0.096	0.097	0.069	0.076	0.076	0.052	0.058	0.056
LATQ (mm)	0.026	0.025	0.025	0.026	0.023	0.023	0.024	0.020	0.021	0.021
GW_Q (mm)	0.236	0.221	0.228	0.211	0.175	0.189	0.168	0.137	0.153	0.131
REVAP (mm)	0.013	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.012	0.011
DP AQ RCHRG (mm)	0.013	0.013	0.013	0.012	0.010	0.011	0.010	0.008	0.009	0.008
TOTAL AQ RCHRG (mm)	0.269	0.253	0.260	0.242	0.203	0.217	0.196	0.161	0.177	0.154
WYLD (mm)	0.343	0.330	0.344	0.328	0.262	0.284	0.262	0.205	0.228	0.203
PERC (mm)	0.287	0.270	0.276	0.257	0.219	0.233	0.210	0.177	0.192	0.169
ET (mm)	0.582	0.597	0.583	0.600	0.667	0.646	0.668	0.727	0.704	0.729
PET (mm)	1.160	1.160	1.160	1.160	1.328	1.328	1.328	1.496	1.496	1.496
TLOSS (mm)	0.005	0.005	0.005	0.006	0.005	0.005	0.006	0.004	0.004	0.005
SED (T/HA)	0.003	0.003	0.003	0.004	0.002	0.002	0.003	0.002	0.002	0.002

Table 4-15 Ratio of Water Balance Components to Precipitation for the Nyangores Basin Land Use-Climate Change Scenarios

	NY DEF	NY PD	NY CD	NY FA	NY PD PR-10 TM+5	NY CD PR-10 TM+5	NY FA PR-10 TM+5	NY PD PR-20 TM+5	NY CD PR-20 TM+5	NY FA PR-20 TM+5
PREC (mm)	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
SURQ (mm)	0.072	0.077	0.087	0.094	0.061	0.069	0.075	0.046	0.053	0.057
LATQ (mm)	0.040	0.040	0.040	0.034	0.036	0.037	0.031	0.032	0.034	0.029
GW_Q (mm)	0.285	0.273	0.282	0.258	0.220	0.239	0.211	0.181	0.203	0.174
REVAP (mm)	0.013	0.013	0.013	0.013	0.013	0.013	0.013	0.012	0.013	0.012
DP AQ RCHRG (mm)	0.016	0.015	0.016	0.014	0.012	0.013	0.012	0.010	0.011	0.010
TOTAL AQ RCHRG (mm)	0.316	0.303	0.313	0.287	0.246	0.267	0.237	0.205	0.228	0.197
WYLD (mm)	0.393	0.385	0.405	0.381	0.312	0.341	0.312	0.256	0.286	0.256
PERC (mm)	0.315	0.301	0.311	0.284	0.245	0.265	0.234	0.203	0.226	0.195
ET (mm)	0.567	0.576	0.556	0.581	0.651	0.621	0.651	0.710	0.679	0.711
PET (mm)	1.075	1.075	1.075	1.075	1.230	1.230	1.230	1.385	1.385	1.385
TLOSS (mm)	0.004	0.004	0.005	0.005	0.004	0.004	0.005	0.003	0.004	0.004
SED (T/HA)	0.005	0.006	0.005	0.006	0.004	0.004	0.005	0.003	0.003	0.004

4.2.4.2.4 Combination of Land use and Climate Change Scenarios

The shown hydrographs display the discharge outputs of the land use-climate change scenarios which are actually more realistic scenarios in terms of future projections of land use change and climate change. The most plausible land use scenarios in the case of the Upper Mara basin are the three covered in this study with the complete deforestation being the least likely to happen among the three. Among the climate change the most plausible scenarios are the combinations of temperature increase and precipitation increase as projected by the IPCC though precipitation reduction is a more often occurrence in the study area as of today and therefore were included in the scenarios analysis.
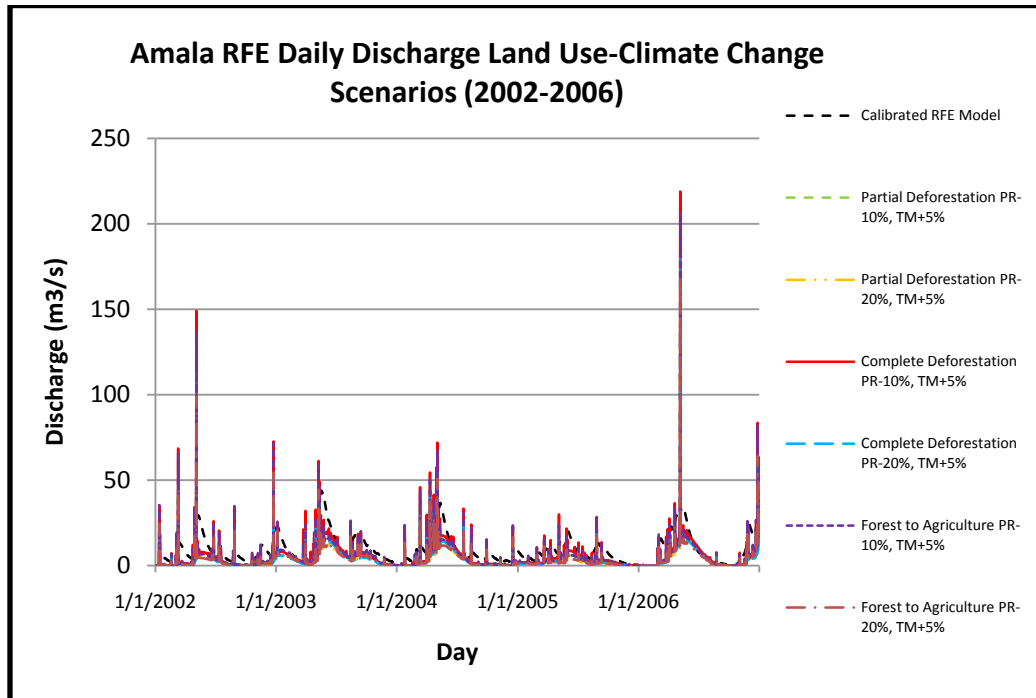
Figure 4-57. Amala Daily Discharge for Land Use-Climate Change Scenarios
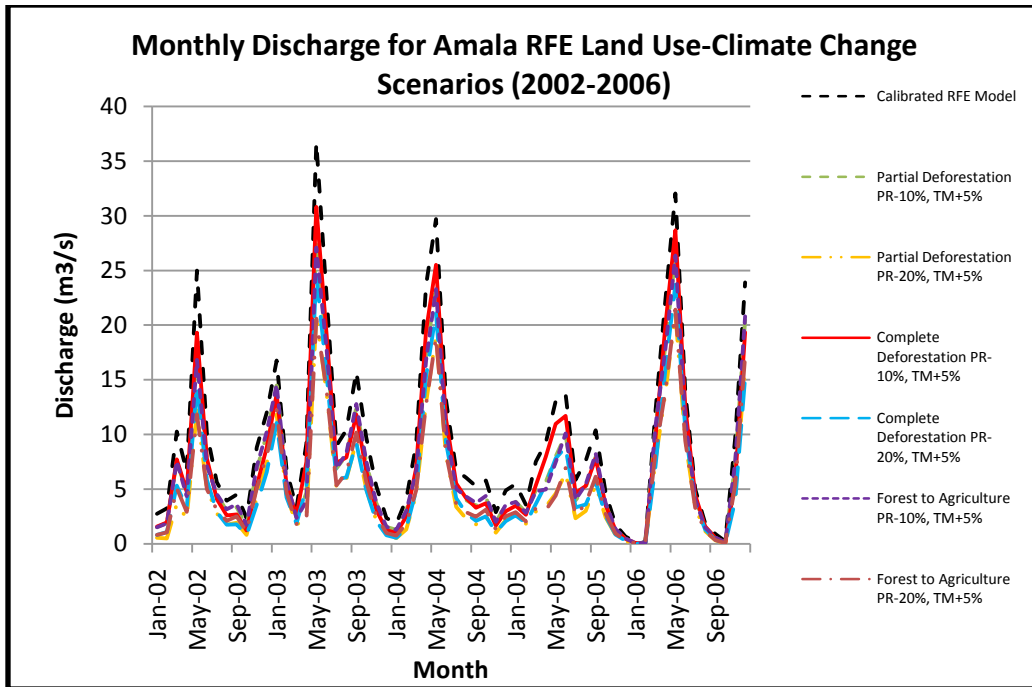
Figure 4-58. Amala Monthly Discharge for Land Use-Climate Change Scenarios
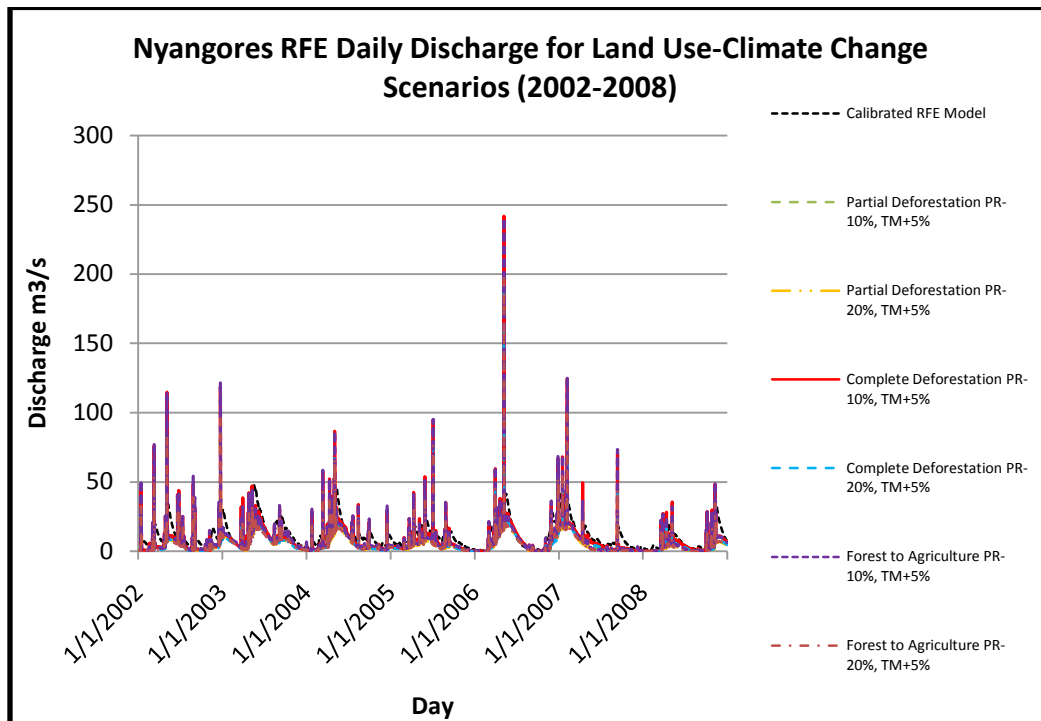
Figure 4-59. Nyangores Daily Discharge for Land Use-Climate Change Scenarios

151
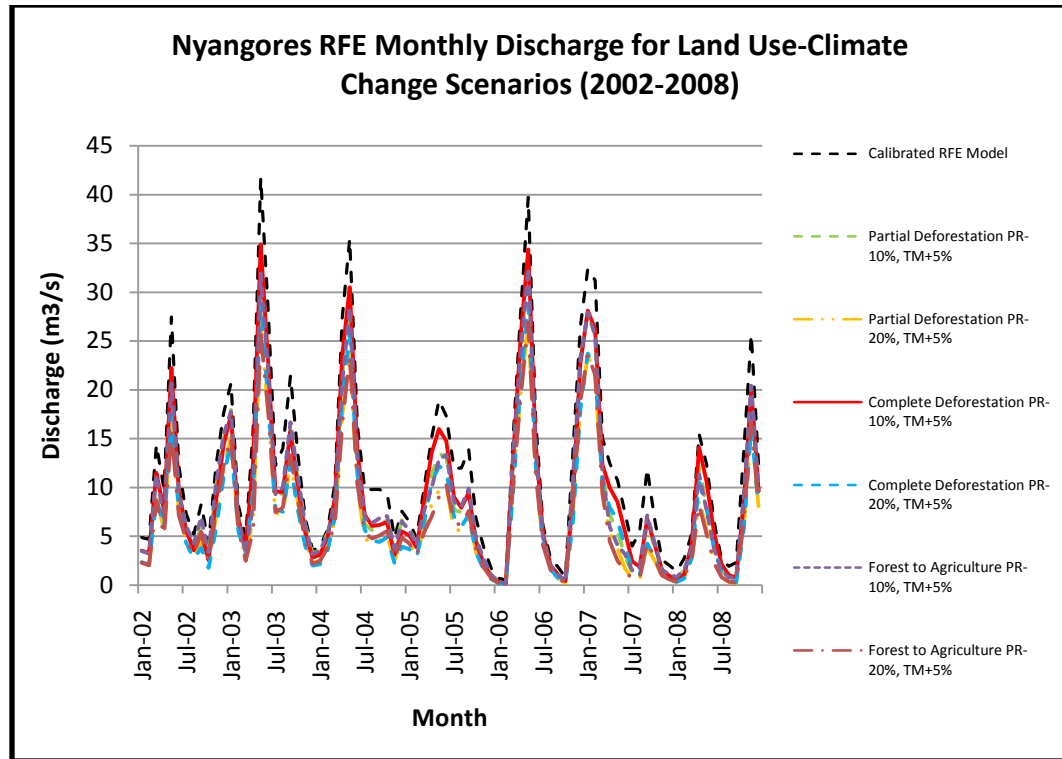
Figure 4-60. Nyangores Monthly Discharge for Land Use-Climate Change Scenarios

The resulting hydrographs and water balance components were able to graphically show the effects of the combined scenarios in terms of stream response. From the hydrographs above and the tables 4-10 and 4-11 above, the effect of land use on the discharge hydrographs and water balance components was evident. The conversion of forest to agriculture scenario had the lowest baseflows and the cause for this is the reduction in ground water recharge which is shown in the table with reductions of up to 49.99% when precipitation is reduced by 20% and a reduction of 32% when precipitation is reduced by 10% and a reduction of 48% at normal precipitation. Complete deforestation on the other hand saw an increase in surface runoff and this occurred in the 20% precipitation reduction in both Amala and Nyangores watersheds.

152
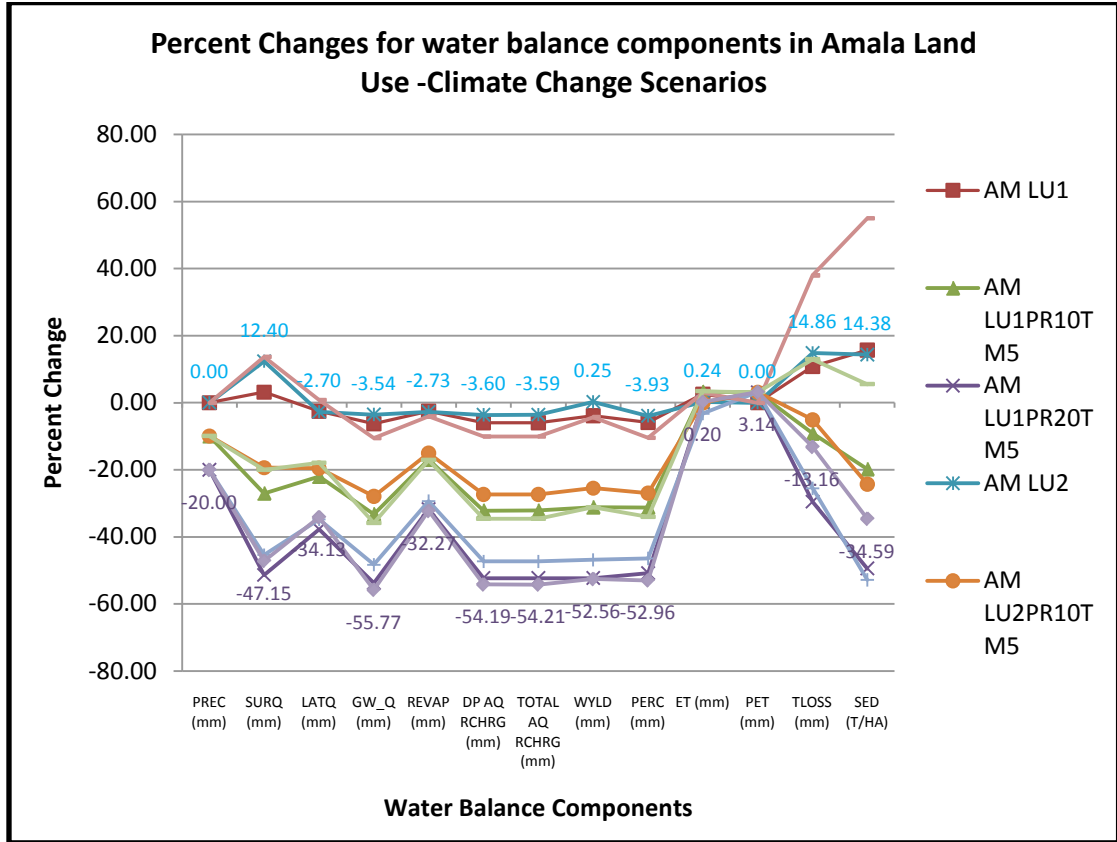
Figure 4-61 Percent Changes for water balance components in Amala Land Use-Climate

Change Scenarios

For the percent changes in water balance components, the figures 4-61 and 4-62 display the variation in the water balance components across different land uses. From the tables 4-12 and 4-13 above, the different land use scenarios affect the water balance components differently and where these differences are most pronounced are in the surface runoff and groundwater recharge as earlier discussed.
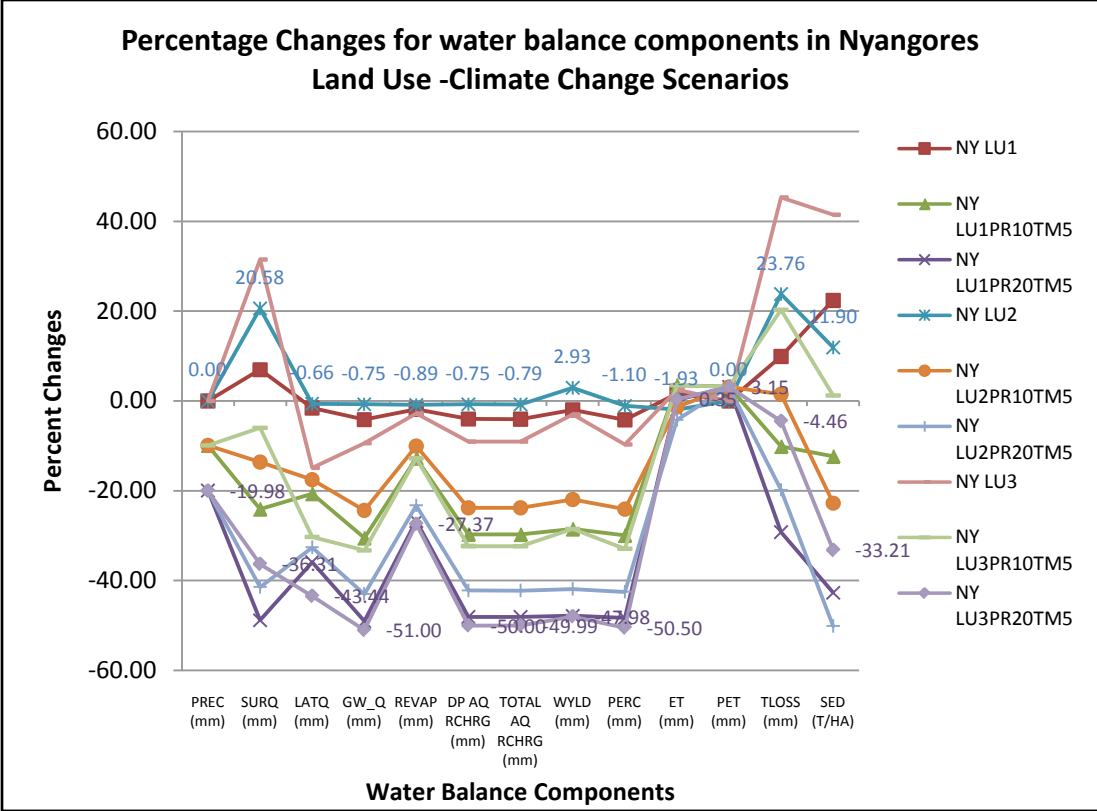
Figure 4-62 Percent Changes for water balance components in Nyangores Land Use-Climate Change Scenarios

5.0 CONCLUSIONS

It is undeniable that land use and climate change have an effect on the hydrology of the upper Mara River basin. The different scenarios have shown the potential effects climate and land use change or transitions will have on the flow of the Mara tributaries. This study however did not take into account the effect of better land and soil management practices on these watersheds will have in the face of these changes. The implementation of better management practices may completely alter the effect these changes may have on the natural system either positively or negatively. The natural system may respond differently in ways we have no knowledge of and therefore it is safe to say that this study examines the potential and not actual impacts of these land use and climate changes on the water flux of these two tributaries.

This study was able to achieve the objectives for which it was designed to a reasonably high degree of success. With regards to mapping land use/ land cover, the expert classifier was able to classify the image at a high accuracy of 84% and this was able to achieve an accurate map of the highly variable study area using far less time and effort than conventional algorithms such as a maximum likelihood classification would have taken. This method is useful in situations where a trip to the field to obtain ground reference data cannot be made as was in the case of this study. The recursive partitioning step and cross validation method is an exceptional way of classifying spectral imagery and this methodology should be further developed to aid in classification of similar areas with highly variable land cover.

From the performance of the models it can inferred that the set-up and calibration of a semi-distributed hydrological model such as SWAT in a large watershed with variable

land cover, soils and topography is a feasible task and will yield satisfactory results given reliable data and proper attention to manual or automatic calibration. The calibration and validation of these models further revealed that the RFE models consistently performed better than the rain gauge models in this study. This is because only two rain gauge stations were used to provide rainfall data for the whole watershed and there lacks a dense ground network of stations that are able to capture the spatial variability or rainfall events such as localized storm cells that may measure between 1 and 15 kilometers in diameter. The RFE data were able to capture the spatial variability of the rainfall in the watersheds resulting in acceptable simulations of the river flow and water balance components. The validation of the models however was not satisfactory and this can be attributed to the parameter values obtained in the calibration process not being good enough for performing further simulations.

This study revealed that land use and climate change scenarios will significantly impact the water flux in the upper Mara River. The climate change scenarios revealed that the variation of precipitation has the greatest impact on the amount of discharge, sediment yield, surface runoff (a reduction of 20% in precipitation will reduce the surface runoff by half its amount in both the Amala and Nyangores watersheds) and generally to the water balance components in the watersheds. As seen in Tables 4-10, 4-11, 4-14 and 4-15, the ratio of the water balance components to precipitation reduces drastically with the reduction of precipitation. This is expected as precipitation is the main driving force of the hydrological cycle and any change in the amount will be directly reflected in the flow of the Mara River. Temperature increase impacts the discharge less directly than decrease in precipitation but nonetheless has an impact by increasing evapotranspiration and plant

production which will ultimately affect land cover in the long-term. More realistic climate change simulations that combined precipitation reduction and temperature increase had the most effect on the discharge and water balance components with a reduction in river discharge in both wet and dry seasons, reduction in total water yield, groundwater discharge, surface runoff and groundwater contribution to the river channel. As previous studies that have found precipitation and slope to be the main factors affecting streamflow in small watersheds, the model results show that Amala and Nyangores are no different and imply that climate change alone will have profound effects on the upper Mara River flow and the human and wildlife inhabitants of the Mara River basin.

The simulation of land use scenarios showed the impact that land use change will have on the discharge of the Amala and Nyangores rivers and water balance components. In these land use-climate change combined scenarios, there is a reduced amount of groundwater recharge, surface runoff, and total water yield to the stream meaning the water balance will be significantly affected by the reduction of precipitation, increased temperature and altered land cover. The long term effects are not well known at this point and a simulation of long term effects may turn out entirely different results.

The model simulations predict that the upper Mara River flow will be significantly affected in the face of the climate and land use change scenarios posing difficulties in adaptation to the altered flow regimes of the Amala and Nyangores rivers. It is therefore prudent to work towards establishing and maintaining adequate minimum flows that would mitigate the effects of reduced baseflows and put in place measures to maintain adequate sustained river flows to the benefit of the stakeholders of the Mara River basin.

6.0 RECOMMENDATIONS

From the encouraging performance of the RFE rainfall data in the simulation of the water balance and discharge of the Amala and Nyangores rivers in this study, additional studies into the feasibility of RFE data should be carried out in order to further validate it and develop methodologies to seamlessly integrate it into hydrological and other resource related studies.

Consistent land use/land cover classification/ mapping coupled with predictive land use transition models should be carried out in order to quantify and characterize true land use/land cover changes. This will help establish trends and enable resource managers to project realistic change scenarios helpful for natural resource management in the Mara River basin. From the study, the only land use change scenarios explored were those that were already perceived to have a negative effect on the water output of the Amala and Nyangores, further work should be done to identify land use changes or land uses with a positive effect on the hydrology of the Mara system which managers and stakeholders can work to establish at a minimum cost to the stakeholder and the environment.

Additional research is needed to improve and extend the findings of this study. An increased spatial and temporal collection of river discharge and sediment data will enable researchers and water resource managers to better understand the characteristics of the Mara River. There is need to 1) quantify these impacts at a river basin scale and 2) consider in more detail the interactions between surface and groundwater in order to properly simulate the water balance of the watershed.

The research questions that emerge from this study are: Is there land use/land cover change occurring in the Mara River basin or is it a combination of plant phenological

cycles and land management land practices? Is land use change by itself a major cause of erratic and reduced flows or is the amount of precipitation the major cause? Is there good land use change and if there is what are these land use types? There are many issues in the Mara River basin that are a constant source of conflict the major two being forest conservation and land use practices. Further research into proper water and land management practices and conservation practices would go a long way in integrated water resource management and also in protecting and ensuring the integrity of the Mara River ecosystem for all the stakeholders who depend on it for their livelihood. Building further on the already existing research work carried out in the Mara Basin will be most beneficial to the residents of the Mara River basin.

An important question however, is to what extent can these changes be quantified and how likely is it that these changes will occur? It is therefore important to have a good sense or picture of the current and potential land use transitions that may occur within the upper Mara River basin and their effects on the water output of the upper Mara tributaries. Results from this study will add to the existing knowledge base of the Mara River basin and will be an important piece in efforts to manage water and other valuable natural resources in similar watersheds worldwide.

REFERENCES

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R., 2006. Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology*. 333, 413– 430

Abbot, M.B. and Refsgaard, J.C. (eds) 1996. Distributed Hydrological Modelling. Dordrecht, Kluwer Academic Publishers.

Adinarayana, J., Flach, J. D., and Collins, W., G. 1994. Mapping Land Use Patterns in a River Catchment using Geographical Information Systems. *Journal of Environmental Management*, 42, 55-61.

Alder-Golden, S.M., Matthew, M. W., Bernstein, L. S., Levine, R. Y., Berk, A., Richtsmeier, S. C., Acharya, P. k., Anderson, G. P., Felde, G., Gardner, J., Hoke, M., Jeong, L. S., Pukall, B., Mello, J., Ratkowski, A. and H. H. Burke, 1999. Atmospheric Correction for Short-wave Spectral Imagery based on MODTRAN4, *SPIE Proceedings on Imaging Spectrometry V*, (3753), 10p.

Anderson, R. J., Hardy, E. E., Roach, J. T., and Witmer, R. E. 1976. A Land Use And Land Cover Classification System For Use With Remote Sensor Data. United States Geological Survey. Washington DC, United States Government Printing Office.

Arnell, N. (1996) Global warming, river flows and water resources. Institute of Hydrology. John Wiley & Sons Ltd: Chichester.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., Williams, J. R., 1998. Large area hydrologic modeling modeling and assessment. Part I: Model development. *Journal of the American Water Resources Association* 34, 73-89.

Beven, K. J. and Kirkby, M. J., 1979, A physically-based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, 24, 1-10.

Beven, K. J. and Moore, I. D., (eds), 1993, *Terrain analysis and distributed modeling in hydrology*. Chichester, John Wiley and Sons.

Bewket, W. and Sterk, G. 2005. Dynamics in land cover and its effect on stream flow in the Chemoga watershed, Blue Nile basin, Ethiopia. *Hydrol. Process.* **19**, 445–458 (2005)

Bratko, I., Kononenko, I., Lavrac, N., Mozetic, I., and Roskar, E., 1989. Automatic synthesis of knowledge: Ljubljana research, *Machine and Human Learning (Y. Kodratoff and A. Hutchinson, editors)*, GP Publishing, Inc., Columbia, Maryland, pp. 25-33.

160

Cihlar, J., Latifovic, R., Beaubien, J., Guindon, B. and M. Palmer. 2003. Thematic mapper (TM) based accuracy assessment of a land cover product for Canada derived from SPOT VEGETATION (VGT) data. *J. Remote Sensing*, Vol. 29, No. 2, pp. 154–170.

Congalton, R. G. and R. A. Mead, 1983, A Quantitative Method to Test for Consistency and Correctness in Photo interpretation, *Photogrammetric Engineering & Remote Sensing*, 49 (1);69-74.

Congalton, R. G., 1991, "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data," *Remote Sensing of Environment*, 37:35-46.

Congalton, R.G., and K. Green, 1999, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Boca Raton, FL: Lewis Publishers, 137p.

Crawley, J. M., 2007. The R Book. John Wiley and Sons Ltd, West Sussex, England.

Crist, E. P. and R. C. Cicone, 1984, Application of the Tasseled Cap Concept to Simulated Thematic Mapper Data, Photogrammetric Engineering & Remote Sensing, 50:343-352

Di Luzio, M., Srinivasan, R., Arnold, J.G. 2002. Integration of Watershed Tools and the SWAT Model into BASINS, *Journal of the American Water Resources Association*, 38(4), 1127-1141.

Dietterich, T.G., and R.S. Michalski, 1983. A comparative review of selected methods for learning from examples, *Machine Learning, Vol. 1* (R.S. Michalski, J.G. Carbonell, and T.M. Mitchell, editors), Morgan Kaufmann, Inc., San Mateo, California.

Duan, Q., S. Sorooshian, H. V. Gupta, A. N. Rousseau, and R. Turcotte. 2003. Advances in Calibration of Watershed Models,AGU, Washington, DC, 2003.Freer, J., K. Beven, and N. E. Peters, Multivariate seasonal period model rejection within the Generalised Likelihood Uncertainty Estimation procedure, in Calibration of Watershed Models, edited by Q. Duan, H. V. Gupta, S. Sorooshian, A. N. Rousseau, and R. Turcotte, pp. 69-87, AGU, Washington, DC, 2003.

Duerr, W. A., Teeguarden, D. E. et al., 1979, Forest resource management: decision-making principles and case. Philadelphia, W. B. Saunders.

ERDAS Incorporated. 2006. ERDAS IMAGINE Tour Guides: ERDAS IMAGINE V9.3. ERDAS Worldwide Headquarters. Atlanta, Georgia. Pp 662.

ESRI. 2006. What is ArcGIS 9.2? ESRI. Redlands, California. Pp119

Food and Agriculture Organization (FAO) of the United Nations. 2005. "Kenya Country Report" In: *Irrigation in Africa in Figures, AQUASTAT Survey 2005*. Rome: FAO

Foody, G. M., 1996, Approaches for the Production and Evaluation of Fuzzy Land Cover Classifications from Remotely Sensed Data, *International Journal of Remote Sensing*, 17 (7): 1317-1340.

Foody, G. M., 2002, Status of Land Cover Classification Accuracy Assessment, *Remote Sensing of Environment*, 80:185-201.

Gann, D. 2003 Land-cover detection and landscape structure analysis in the Pachitea Basin, Peruvian Amazon. Thesis. Florida International University, Miami, FL. USA. pp 189.

Gassman, P. W., Reyes, M. R., Green, C. H., Arnold. J. G., 2007. The Soil And Water Assessment Tool: Historical Development, Applications, and future research directions. American Society of Agricultural and Biological Engineers Vol. 50(4): 1211-1250

Gebert, W.A., Krug, W.R., 1996. Streamflow trends in Wisconsin's Driftless Area. *Water Resources Bulletin* 32 (4), 733–744.

Gereta, E., Wolanski, E., Borner, M. and S. Serneels, 2002, Use of an ecohydrology model to predict the impact on the Serengeti ecosystem of deforestation, irrigation and the proposed Amala Weir water Diversion Project in Kenya. *Ecohydrology and Hydrobiology* Vol. 2 No 1-4, 135-142.

Green WH, Ampt G. A. 1911. Studies on soil physics, 1. The flow of air and water through soils. *Journal of Agricultural Sciences*; 4: 11-24

Grizzetti, B., Bouraoui, F., De Marsily, G., 2005. Modelling nitrogen pressure in river basins: A comparison between a statistical approach and the physically-based SWAT model. *Physics and Chemistry of the Earth*, 30, 508–517

Haddeland. I, T. Skaugen, and D. P. Lettenmaier. 2006. Hydrologic effects of land and water management in North America and Asia:1700–1992. *Hydrol. Earth Syst. Sci. Discuss.*, 3, 2899–2922.

Hargreaves, G. H. and Z. A. Samani. 1982. Estimating potential evapotranspiration. Tech. Note, *J. Irrig. And Drain. Engr*. 111 (2): 113-124.

Hoffman, C. M. 2007. Geospatial mapping and analysis of water availability-demand-use within the Mara River Basin. Thesis. Florida International University, Miami, FL, USA. pp 114.

Holvoet, K. et al. 2007. Modifications to the SWAT code for modelling direct pesticide losses, *Environ. Model. Softw*, doi:10.1016/j.envsoft.2007.05.002

Huang, X. and J. R. Jensen, 1997, "A Machine-Learning Approach to Automated Knowledge-base Building for Remote Sensing Image Analysis with GIS Data," *Photogrammetric Engineering & Remote Sensing*, 63 (10): 1185-1194.

Intergovernmental Panel on Climate Change. 2007. Climate Change 2007: Synthesis Report. Cambridge Press. Cambridge.

Intergovernmental Panel on Climate Change. 2001. Climate Change 2001: Synthesis Report. Cambridge Press. Cambridge.

International Water Management Institute 2007. *Water for Food, Water for Life: A Comprehensive Assessment of Water Management in Agriculture.* London: Earthscan, and Colombo:

Jackson, P., 1990. *Introduction to Expert Systems, Second Edition,* Addison-Wesley Publishing Company, Wokingham, England.

Jacobs, J.H., Angere, J., Vitale, J., Srinivasan, R., and R. Kaitho, 2007**,** Mitigating Economic Damage in Kenya's Upper Tana River Basin: An Application of Arc-View SWAT., *Journal of Spatial Hydrology* Vol.7, No.1.

Jayakrishnan, R., Srinivasan, R., Santhi, C.and Arnold, J., 2005. Advances in the application of the SWAT model for water resources management. *Hydrol. Process.* 19, 749–762 (2005)

Jensen, J.R., 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective, Third Edition,* Prentice-Hall, Englewood Cliffs, New Jersey.

Juckem, P. F., Hunt, R. J., Anderson, M. P., Robertson, D. M. 2008. Effects of climate and land management change on streamflow in the driftless area of Wisconsin. *Journal of Hydrology* 355, 123– 130

Kauth R. J and G.S Thomas, 1976, The Tasseled Cap-A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat, *Proceedings, Symposium on Machine Processing of Remotely Sensed Data,* West Lafayette, IN: LARS, 41-51

Kenya. Central Bureau of Statistics (CBS). 2006 Ministry of Planning and National Development. *Kenya Facts and Figures, 2006 Edition*. Nairobi: CBS

Kirby, M.J., Naden, P. S., Burt, T. P. and Butcher, D. P., 1993, Computer *Simulation in Physical Geography*. Chichester, John Wiley and Sons.

Krug, W.R., 1996. Simulation of temporal changes in rainfall runoff characteristics, Coon Creek Basin, Wisconsin. *Water Resources Bulletin* 32 (4), 745–752.

Landis, J. and G. Koch, 1977, "The Measurement of Observer Agreement for Categorical Data", *Biometrics*, 33:159-174

Liebowitz, J., 1988, Introduction to Expert Systems, Santa Cruz, CA: Mitchell Publishing, p. 3

Liu X. H., & Liu, Y., 2008, *The accuracy assessment in areal interpolation: An empirical investigation*, Science in China Series E: Technological Sciences

Mati B.M., Mutie S., Home P. Mtalo F and Gadain H. 2005. Land Use Changes in the Transboundary Mara Basin: A Threat to Pristine Wildlife Sanctuaries in East Africa. A Paper presentation at the: 8th International River Symposium, Brisbane, Australia, September 6-9, 2005

McLachlan G. J., Do, K.A., and C. Ambroise. 2004. *Analyzing microarray gene expression data*. John Wiley & Sons Ltd.

McKay M.D., Beckman, R.J. and Conover, W.J., 1979. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2), 239-245

McKay M.D., 1988. Sensitivity and Uncertainty Analysis Using a Statistical Sample of Input Values. Uncertainty Analysis, Y. Ronen, ed. CRC press, Inc. Boca Raton, Florida, 145-186.

Melesse, A.M., Hydrological Modeling to Guide Water management in the Mara River Basin, 2006. Global Water for Sustainability Program, Florida International University. 3pp

Michaud, J. D. and Sorooshian, 1994. S. Effects of rainfall sampling errors on simulations of desert flash floods. *Wat. Resour. Res.*, 30, 10, 2765-2775.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. and T. L. Veith. 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, 50:(3), 885–900.

Monteith, J. L. 1965. Evaporation and the environment. P. 205-234. *In* The state and movement of water in living organisms, XIXth Symposium. Soc. For Exp. Biol., Swansea, Cambridge University Press.

Mutie, S., Mati, B., Gadain, H., and Gathenya, J., 2006. Evaluating land use change effects on river flow using USGS geospatial stream flow model in Mara River basin, Kenya. A Paper presentation at the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover. Bonn, 28-30 September 2006.

Mutie S. M., Mati B., Gadain H., and Home, P. 2005 Land-cover Change Effects on Flow regime of Mara River. A paper presented and published in the workshop proceedings of Information System for Crisis Management (ISCRAM), Belgium. April 2005.

Nash, J. E. and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I-A discussion of principles , *Journal of Hydrology*, 10 (3), 282–290.

Neitsch S. L, Arnold JG, Kiniry JR, Williams JR. Soil and Water Assessment Tool, Theoretical Documentation: Version 2005. Temple, TX. USDA Agricultural Research Service and Texas A&M Blackland Research Center, 2005.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R. 2001.Soil and Water Assessment Tool Theoretical Documentation, Blackland Research Center, Texas Agricultural Experiment Station, Temple, TX.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R. 2001. Soil and Water Assessment Tool User's Manual. Blackland Research Center, Texas Agricultural Experiment Station. Temple, TX.

Nile Basin Initiative (NBI) Shared Vision Program. 2001. Nile River Basin: Transboundary Environmental Analysis. United Nations Development Programme, World Bank, Global Environment Facility

Ogutu, O.J., Piepho, H. P., Dublin, H. T., Bhola, N. and R. S. Reid. 2007. El Nino-Southern Oscillation, rainfall, temperature and Normalized Difference Vegetation Index Fluctuations in the Mara-Serengeti ecosystem. *Afr. J. Ecol*., 46, 132-143

Olsson, L. and P. Pilesjo. 2002. Approaches to spatially distributed hydrological modelling in a GIS environment: In Skidmore, A., 2002, (editors). *Environmental modeling with GIS and Remote Sensing*. London. Taylor and Francis.

Paine, D. P. and J. D. Kiser, 2003, "Chapter 23: Mapping Accuracy Assessment", Aerial Photography and Image Interpretation, 2nd ed., New York: John Wiley & Sons, 465-480.

Pal, M. and Mather, P., M. 2001. Decision Tree based Classification of Remotely Sensed Data. Paper presented at the 22nd Asian conference on remote Sensing, 5-9 November 2001, Singapore.

Phoon, S. Y., Shamseldin, A. Y. and Vairavamoorthy, K. 2004. Assessing impacts of climate change on Lake Victoria Basin, Africa. People-Centred Approaches To Water And Environmental Sanitation: 30th WEDC International Conference, Vientiane, Lao PDR, 2004

Potter, K.W., 1991. Hydrological impacts of changing land management practices in a moderate-sized agricultural catchment. *Water Resources Research* 27 (5), 845–855.

Priestley, C. H. B. and Taylor R. J. On the assessment of surface heat flux and evaporation using large-scale parameters, 1972. Mon Weather. Rev. 100: 81-92.

R Development Core Team, 2009. An Introduction to R: Notes on R: A Programming Environment for Data Analysis and Graphics Version 2.10.1 (2009-12-14) Available online at URL http://www.r-project.org/

Refsgaard, J. C. 1997. Parameterisation, calibration, and validation of distributed hydrological models. *J. Hydrology* 198(1): 69-97.

Rosenfield, G. H. and K. Fitzpatrick-Lins, 1986, "A Coefficient of Agreement asa a Measure of Thematic Classification Accuracy", *Photogrammetric Engineering & Remote Sensing*, 52 (2):223-227.

Robock, A., Turco, R.P., Harwell, M.A., Ackerman, T.P., Anderssen, R., Chang, H.-S. and Sivakumar, M.V.K. 1993. Use of general circulation model output in the creation of climate change scenarios for impact analysis. *Climatic Change*. Vol. 23, pp. 293-335

Russo, J.M & Zack, J.W. (1997). Downscaling GCM output with a mesoscale model. *Journal of Environmental Management.* Vol. 49, pp. 19-29

Santhi, C., Arnold, J.G., Williams, J.R., Srinivasan, R., 2006. A modeling approach to evaluate the impacts of water quality management plans implemented in a watershed in Texas. *Environmental Modelling & Software* 21, 1141e1157

Serneels, S., Said, M. Y., and Lambin, E. F. 2001. Land cover changes around a major east African wildlife reserve: the Mara Ecosystem (Kenya). *Int. J. Remote Sensing,* 2001, vol. 22, no. 17, 3397–3420

Setegn, S. G., Srinivasan, R., Dargahi, B., 2008. Hydrological Modelling in the Lake Tana Basin, Ethiopia using SWAT model. *The Open Hydrology Journal* Vol. 2(2008): pp 25-38

Shirmohammadi, A., T. W. Chu, H. Montas, and T. Sohrabi. 2006. SWAT model and its applicability to watershed nonpoint-source pollution assessment. ASAE Paper No. 012005. St. Joseph, Mich.: ASAE.

Singler, H., and McClain, M., 2006. A Water Quality Baseline Assessment of the Mara River Basin, Kenya-Tanzania. Global Water for Sustainability Program, Florida International University. 66pp

Skidmore, A., 2002, (editors). Environmental modeling with GIS and Remote Sensing. London. Taylor and Francis.

Sorooshian, S. and Gupta, V. K. 1983. Evaluation of Maximum Likelihood Parameter Estimation Techniques for Conceptual Rainfall-Runoff Models: Influence of Calibration Data Variability and Length on Model Credibility. Water Resources Research, 19(1), 251-259.

Stehman, S.V. and R. L. Czaplewski, 1998, Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles, *Remote Sensing of Environment*, 64:331-334

Story, M. and R. Congalton, 1986. Accuracy Assessment: A User's Perspective, *Photogrammetric Engineering & Remote Sensing*, 52 (3): 397-399

Survey of Kenya (SoK). 2003. *National Atlas of Kenya, Fifth Edition*. Nairobi: SoK.

Thornthwaite, C.W. and Mather, J.R., 1955, The water balance. Centerton, NJ, Laboratory for climatology publications in climatology, 8, 1-86

USDA Soil Conservation Service (SCS). National Engineering Handbook Section 4 Hydrology, Chapters 4-10, 1972

United Republic of Tanzania, 2007. Realising Water's Potential to Support Growth in Tanzania: Brief 4; Poverty and Human Development Report (PHDR). Dodoma. The Poverty Eradication Division of the Ministry of Planning and Economic Affairs.

Van der Meer, F., 2000, Imaging spectrometry for geological applications. In: G Meyers, R. A. (Ed.), Encyclopaedia of Analytical Chemistry: Applications, Theory and Instrumentation, (New York: Wiley), Vol. A2310, 31pp

Van Griensven A, Meixner T, Grunwald S, Bishop T, Diluzio M, Srinivasan R. 2006. A global sensitivity analysis tool for the parameters of multi-variable catchment models. *J Hydrol* ; 324(1-4): 10-23

Van Loon, A., R. Lasage, H. Mathijssen, P. Droogers 2007, Water Management Support Methodologies: State of the Art, WatManSup Research Report No 1, General Foulkesweg 28, 6703 BS Wageningen, The Netherlands.

Vieux, B. E., 2001. Distributed Hydrologic modeling Using GIS. Kluwer Academic Publishers. Dordrecht, The Netherlands

Ward, A. D. and Trimble, W. S. 2004. Environmental Hydrology. 2nd ed. Lewis Publishers. Boca Raton.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics*, *1*, 80-83.

Williams, J.R. 1969. Flood routing with variable travel time or variable storage coefficients. *Transactions of the ASAE* 12 (1): 100-103.

Williams, J.R. 1995. Chapter 25: The EPIC model. P. 909-1000. *In* V.P. Singh (ed). Computer models of watershed hydrology. Water Resources Publications, Highlands Ranch, CO.

Woodcock, E.W., Macomber, S. A., and L. Kumar. Vegetation mapping and monitoring: In Skidmore, A., 2002, (editors). *Environmental modeling with GIS and Remote Sensing*. London. Taylor and Francis.

World Resources Institute, Department of Resource Surveys and remote Sensing, Ministry of Environment and Natural resources, Kenya; Central Bureau of Statistics, Ministry of Planning and Development, Kenya; and International Livestock Research Institute. 2007. *Nature's Benefits in Kenya: An Atlas of Ecosystems and Human Well-Being*. Washington DC and Nairobi: World Resources Institute.

World Resources Institute, Department of Resource Surveys and Food and Agriculture Organization (FAO) of the United Nations. 2003. Resources and Freshwater Ecosystems-Kenya: Earth Trends Country Profiles: Accessed at URL http://earthtrends.wri.org

Wu, W. and G. Shao, 2002, "Optimal Combinations of Data, Classifiers, and Sampling Methods for Accurate Characterizations of Deforestation", *Canadian Journal of Remote Sensing*, 28(4):601-609.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-Year monthly analysis based on gauge observations, Satellite Estimates, and numerical model outputs. *Bulletin of the American Meteorological Society,* 78, 11 2539-11 2558

APPENDIX

APPENDIX 1: Kenya's Internal Water Resources

Internal Renewable Water Resources (IRWR),1977-2001, in cubic km	Kenya	Sub-Saharan Africa
Surface water produced internally	17	3,812
Groundwater recharge	3	1,549
Overlap (shared by groundwater and surface water)	0	1,468
Total internal renewable water resources (surface water + groundwater - overlap)	20	3,901
Per capita IRWR, 2001 (cubic meters)	633	5,705
Natural Renewable Water Resources(includes flows from other countries)		
Total, 1977-2001 (cubic km)	30	X
Per capita, 2002 (cubic meters per person)	947	X
Annual river flows:		
From other countries (cubic km)	10	X
To other countries (cubic km)	X	X
Water Withdrawals		
Year of withdrawal data	1990	
Total withdrawals (cubic km)	2.0	X
Withdrawals per capita (cubic m)	87	X
Withdrawals as a percentage of actual renewable water resources	9.2%	X
Withdrawals by sector (as a percent of total) {a}		
Agriculture	76%	X
Industry	4%	X
Domestic	20%	X

Source: WRI, 2003

Appendix 2: Tanzania's Internal Water resources

Internal Renewable Water Resources (IRWR),1977-2001, in cubic km	Tanzania	Sub-Saharan Africa
Surface water produced internally	80	3,812
Groundwater recharge	30	1,549
Overlap (shared by groundwater and surface water)	28	1,468
Total internal renewable water resources (surface water + groundwater - overlap)	82	3,901
Per capita IRWR, 2001 (cubic meters)	2,227	5,705
Natural Renewable Water Resources (includes flows from other countries)		
Total, 1977-2001 (cubic km)	91	X
Per capita, 2002 (cubic meters per person)	2,472	X
Annual river flows:		
From other countries (cubic km)	X	X
To other countries (cubic km)	X	X
Water Withdrawals		
Year of withdrawal data	1994	
Total withdrawals (cubic km)	1.2	X
Withdrawals per capita (cubic m)	39	X
Withdrawals as a percentage of actual renewable water resources	1.6	X
Withdrawals by sector (as a percent of total) {a}		
Agriculture	89%	X
Industry	2%	X
Domestic	9%	X

Source: WRI, 2003