UNESCO-IHE
INSTITUTE FOR WATER EDUCATION

MODELING THE IMPACT OF LAND USE CHANGE ON RIVER HYDROLOGY IN DATA SCARCE ENVIRONMENTS

THE CASE OF THE MARA RIVER BASIN KENYA

Allan Caleb ABWOGA

MSc Research Thesis (WM.12.10)
April 2012

UNESCO-IHE
Institute for Water Education

MODELING THE IMPACT OF LAND USE CHANGE ON RIVER
HYDROLOGY IN DATA SCARCE ENVIRONMENTS

THE CASE OF THE MARA RIVER BASIN KENYA

Master of Science Thesis
by
Allan Caleb Abwoga

Supervisors
Prof. Michael McClain (UNESCO-IHE)
Dr. Marloes Mul (UNESCO-IHE)

Examination committee
Prof. Michael McClain (UNESCO-IHE)
Dr. Markus Hrachowitz (TU Delft)
Dr. Marloes Mul (UNESCO-IHE)

This research is done for the partial fulfilment of requirements for the Master of Science degree at the
UNESCO-IHE Institute for Water Education, Delft, the Netherlands

Delft
March 2012

This MSc work is dedicated to my late parents, Jared and Beldina Abwoga

Abstract

The Mara basin is a trans-boundary basin cutting across Kenya and Tanzania. It traverses the internationally acclaimed Masaai Mara and Serengeti game reserves. It covers an area of 13,750 km^2 and the Mara river runs for 395 km from the Napuiyapui swamp on the Mau escarpment in Kenya to Musoma bay in Tanzania. The basin supports a human population of approximately 800,000 inhabitants. Seventy five (75%) percent of the Mara river flow emanates from the upper Mara in Kenya. The Nyangores and Amala sub basins form the upper part of the catchment. Many livelihoods and ecosystems depend on the Mara river and yet it faces numerous interactions that require effective management to ensure sustainability of the water resource.

Over the past 40 years with population growth, there has been significant change in land use in the basin. Forest cover has been lost (32%) and land under agriculture more than tripled in size. This rapid land use change is believed to have impacted on the hydrology of the Mara River and there is need to mitigate the hydrological impacts of changing land use to ensure sustainability of the Mara ecosystem.

The objectives of this research were to develop a hydrological model for the upper Mara which would subsequently be used to simulate the impacts of land use change on the Mara hydrograph, and in addition compare the modelling efficiency of this research to other previous modelling efforts in order to generate information that would be used for sustainable management of the Mara river basin.

A model was developed based on the STREAM model (Spatial Tools for River Basins and Environment and Analysis of Management Options) that was successfully used to model the Zambezi river basin. The model was coded using the environmental modelling language of PCRaster. Input data included hydromet data from the Kenya metrological department and discharge data from the Water Resources Management Authority. Soils, land use and location maps were processed in ARCview (GIS). The data was processed, by filling the missing data gaps through multiple linear regressions and converting it to a form compatible with PCRaster language. The chosen modelling period was January 1999 to December 2003. The modelling was carried at daily time steps and the results aggregated to monthly level.

The modelling efficiency results were an NSE of 0.59 and 0.56 for Nyangores and Amala respectively for calibration and 0.52 and 0.35 for Nyangores and Amala respectively for validation. Previous modelling efforts in the Mara by Mango in 2011 registered efficiencies of -0.53 and 0.08 for calibration for Nyangores and Amala respectively and - 0.06 and 0.41 for Nyangores and Amala respectively for validation. This was done using similar data sets and time periods. STREAM performs significantly better in modelling Nyangores but registers similar achievement for Amala catchment.

The model was subsequently applied to three (3) land use change scenarios. A partial deforestation scenario where up to 25 % of forest cover is lost, a re-afforestation scenario, where forest cover is increased by up to 26 % and a land use change scenario of converting forests to Agricultural use. Results of land use change scenarios indicate that under reforestation there will be a reduction in river flows, whereas for deforestation there is increase in peak flows but a reduction in base flows. The conversion of all forest land to agricultural land has the highest impact on Mara flows with an increase in stream flows between 8% and 10% but with reduced contribution of attendant precipitation to base flow, meaning reduced flows during the dry periods.

Comparison of STREAM modelling efficiency to SWAT and GeoSFM shows that it performs fairly well, out performing both models in Nyangores and achieving comparable results for the Amala river. Of importance is that both SWAT and STREAM indicate the biggest impact on the Mara hydrograph for the scenario with forest land converted to agricultural use. This is confirmed by the GeoSFM model which was used to carry out a historical analysis of flow change due to changes in land use over a 30 year period.

Keywords: Mara river basin, Hydrological processes, Mara hydrograph, Modelling efficiency Land use change.

Acknowledgements

It is an honour and great pleasure to have the opportunity to appreciate and thank all those who have enabled me to successfully undertake and complete my MSc. studies.

I wish to start by giving thanks to almighty GOD because it is by His will that all things are possible.

I extend my appreciation and sincere thanks to my supervisor Prof. Michael McClain for the support, advice and encouragement throughout the period of this MSc research; in the same breath, I express my utmost appreciation to Prof. Stefan Uhlenbrook for valuable advice and guidance.

My mentor Dr. Marloes Mul has been an invaluable resource throughout this research. I wish to thank her most sincerely for her advice and guidance throughout this research from inception to completion and for sharing with me her vast experience in hydrological modelling.

Members of staff in Integrated Water Systems and Governance department have been a valuable resource providing backstopping services whenever called upon. In this respect I wish to thank Prof. Pieter van der Zaag for his invaluable advice and critic during the MSc research seminars, Dr. Masih Ilyas for advice and guidance and Ms. Susan Grass for advice guidance and encouragement during proposal preparation and throughout the research period.

I extend my appreciation to Patrick Khisa and Fidelis Kilonzo for sharing with me valuable data.

I wish to thank all my colleagues at UNESCO - IHE for creating an enabling environment that enabled me to undertake this research. I specifically extend special mention to my colleagues and friends from Integrated Water Systems and Governance department Ayu Wijayanti, Eric Muala, Edwin Afosa, Fadzai Mukororira, Fortune Nyatsanza, Le Tuan, Tennakoon Sobana, Lilian Masilago and David Bosuben for encouragement and sharing of experiences.

I wish to extend my appreciation to the Kenyan government and the Netherlands government through Nuffic for providing a scholarship to undertake these studies in Netherlands.

Last but by all means not least, i wish to extend my heartfelt appreciation to my dear family, Judy and Anthony, for their love, prayers, support and encouragement as I undertook my studies away from home.

To all those that I have not mentioned by name, you are equally as important. Thank you for the support.

God bless you all.

Table of Contents

List of Figures

List of Tables

List of Acronyms

ASL	-	Above Sea level
DEM	-	Digital Elevation Model
EML	-	Environmental Modelling Language
FEWS	-	Famine Early Warning System
GIS	-	Geographic Information System
GLOWS	-	Global Water for Sustainability Program
GPS	-	Global Positioning System
HOF	-	Horton Overland Flow
HRU	-	Hydrological Response Units
INF	-	Infinity
IPCC	-	Intergovernmental Panel on Climate Change
LAI	-	Leaf area Index
LEW	-	Lumped Elementary Watersheds Structure Management Options
MIKE-SHE	-	MIKE System Hydrologic European
MRB	-	Mara River Basin
MS TM	-	Multi Spectrum Thematic Mapper
MSE	-	Mean Square Error
NSE	-	Nash - Sutcliff Efficiency
PET	-	Potential Evapotranspiration
RFE	-	Rainfall estimates
SOF	-	Saturated Overland flow
SORTER	-	Digital Soil and Terrain Database of East Africa
SRTM	-	Shuttle Radar Topography Mission
STREAM	-	Spatial Tools for River Basins and Environment and Analysis of Management Options.
SWAT	-	Soil and Water Assessment Tool
TAC^D	-	Tracer Aided Catchment (Distributed)
TEMP	-	Temperature
USGS	-	United States Geological Survey.

1 INTRODUCTION

Water is an essential resource for all life on the planet. At present only about 0.8 % of the world's fresh water can be exploited by mankind in ever increasing demands for sanitation, drinking, manufacturing, leisure, agriculture and ecosystem functions.

Successful management of any resource requires accurate knowledge of the resource availability, its uses and competing demands and mechanisms that translate policy decisions into actions on the ground. For water as a resource this is particularly difficult since the flow of water is not restricted to national boundaries. This has resulted in a new focus for management of water resources which is the river basin.

According to the Helsinki rules a river basin is defined as "the geographical area determined by watershed limits of the system of waters, including surface and underground waters flowing into a common terminus" (International law association, 1967). Strong relations exist between ground water and surface water in river basins, between water quality and quantity and between land and water upstream and downstream. These relationships turn river basins from a geographical area into a coherent system.

The subject of this research is the Mara river basin that cuts across Kenya and Tanzania. Many livelihoods and ecosystems depend on the Mara River and yet it faces numerous interactions that require effective management to ensure sustainability of its water resources. The Mara river basin has over the last 40 years, undergone numerous changes as a result of increasing human population. This has manifested itself in the form of land use changes within the basin. These changes in the basin have resulted in changes in the Mara river hydrology, threatening the very existence of the Mara River. This research seeks to contribute to the sustainable management of the Mara river basin through generation of information that can be used in decision making.

1.1 Study Area Description

The Mara river basin is a transboundary basin cutting across Kenya and Tanzania. It lies between longitudes 33^{o} 47' E and 35^{o} 47' E and Latitudes 0^{o} 28' S and 1^{o} 52' S. It traverses the internationally acclaimed Maasai Mara and Serengeti game reserves in Kenya and Tanzania. The Mara is considered one of the least impacted catchments in the upper catchment of the Nile basin draining into Lake Victoria (Mati et al., 2008). The basin covers an area of 13,750 km^2 and the river runs 395 km long. It flows from an altitude of approximately 3000 m at the Napuiyapui swamp in the Mau forest in Kenya across different landscapes and drains into Lake Victoria at Musoma bay in Tanzania at an altitude of 1134 m above sea level. Major tributaries of the Mara are Nyangores, Amala, Talek, Sand and Engare Engito on the Kenyan side and the Bologonja River on the Tanzania side. 65% of the basin is in Kenya, while the remaining portion is in Tanzania. The Nyangores and Amala sub-basins form the upper part of the catchment and are the main sources of water in the Mara basin (See fig 1.1).

Figure 1.1 Location of Mara River Basin

Source: GLOWS, 2009

1.1.1 Climate

The Mara basin experiences bimodal rainfall, occurring in March to May and in October to December. The rainfall varies with altitude in the basin ranging from 1000 to 1750 mm/year in the Mau escarpment to 900-1000 mm/year in the middle rangelands and 700-850 mm/year in the lower Loita hills and around Lake Victoria (Mati et al., 2008).

1.1.2 Demographical and socio economic characteristics

The Mara River provides water for the basin inhabitants in both Kenya and Tanzania. In Kenya, the basin comprises the counties of Nakuru, Bomet, Transmara and Narok with a combined population of 573,883 as per the 1999 population census (See Fig 1.2). These are results aggregated from the sub locations that constitute the Mara basin (Barno, 2011). In Tanzania, the basin is constituted by the Districts of Serengeti, Rorya, Tarime and Musoma with a population of 231,614 people as of the 2002 population census (NBS, 2002). The basin supports the economic livelihoods of the basin inhabitants involved in several economic activities ranging from pastoralism, farming, fishery, tourism and some hunter gatherers in the forested regions (Mati et al., 2008).

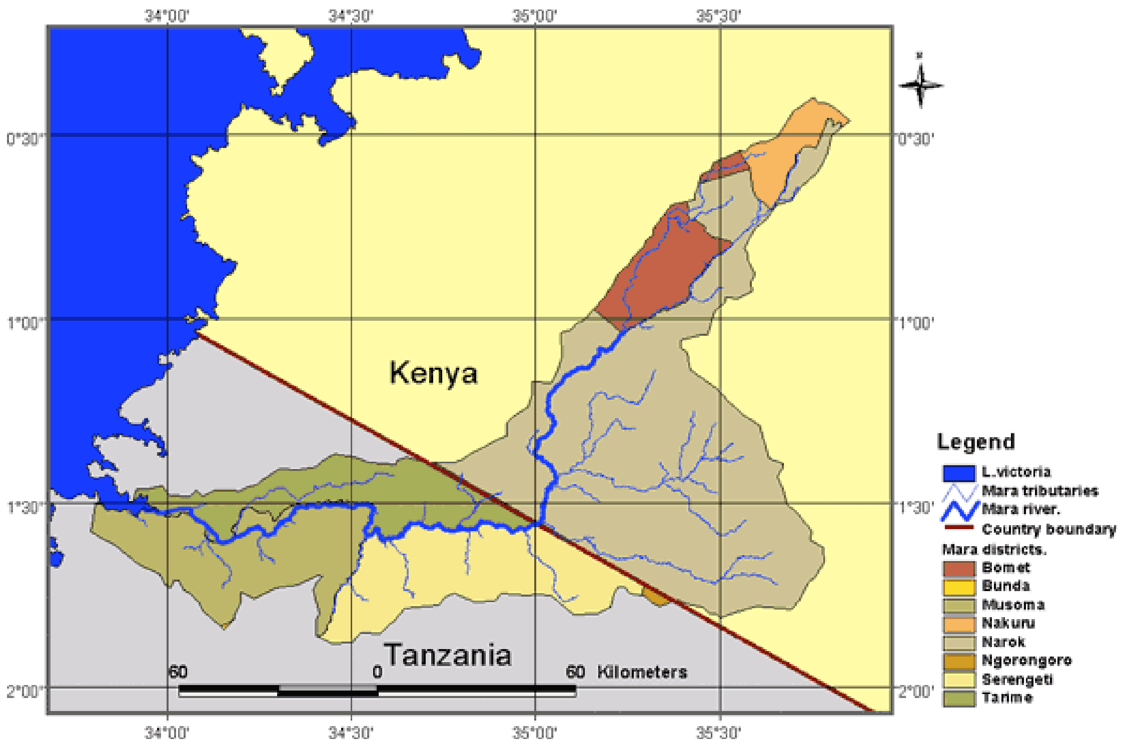

Figure 1.2 Mara Administrative Districts

Source: GLOWS, 2009

1.1.3 Soils, Land use and land cover

The Mara river basin is dominated by two types of soils Cambisols, found mainly in the middle and upper part of the basin, and Vertisols, which are characteristic of the lower part (Sombroek et al., 1982, as cited by Mati et al., 2008). These soil characteristics and two climatic zones in the Mara river basin result in predominance of savannah vegetation in the middle part of the basin, and dense tropical rainforests in the Mau escarpment in Kenya (Mati et al., 2008). The Mara river basin consists of four distinct land use sections as illustrated in Fig 1.3. The forested Mau escarpment, the middle section characterized by both small and large scale agricultural farms and the open savannah grassland protected by Maasai Mara National reserve and Serengeti National Park. In addition are the wetlands where the Mara River discharges into Lake Victoria.

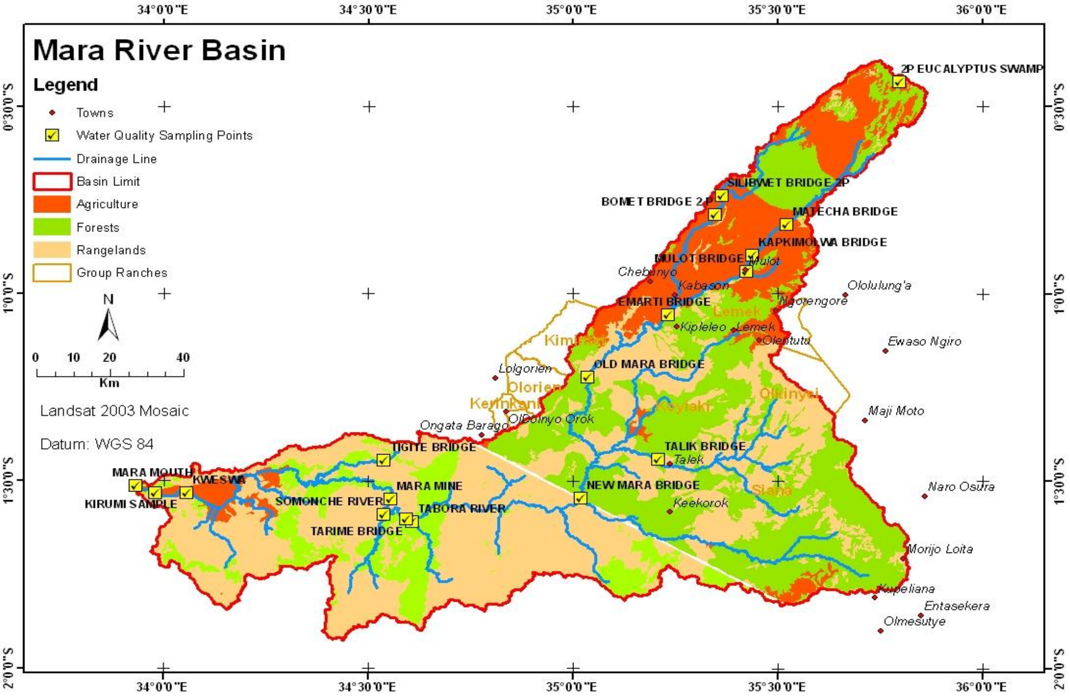

Figure 1.3 Mara Land use map, 2008

Source: GLOWS, 2009

1.2 Problem Definition

There is an urgent need to effectively manage the Mara River's resources to sustain both the ecosystem and economic activities that take place in the basin in the wake of the rapidly changing landscape. Understanding the hydrological response of watersheds to land use and climatic change is an important component of water resource planning and management (Mango et al., 2011). Effective planning for future situations requires understanding of impacts of changing conditions on river hydrology. The landscape of the Mara basin is changing and this has to be factored into any viable management options for the sustainability of the basin.

The Upper Mara contributes 75% of the Mara flows (LVBC & WWF-ESARPO, 2010). It is thus an important part of the basin and its hydrological integrity needs to be preserved to ensure sustainability of the Mara ecosystem (See Fig 1.4 and Fig 1.5).

Use of hydrological models is the best way to predict hydrological response to changing landscapes. However, hydrological modeling faces challenges in the case of lack of proper understanding of the hydrological processes in data scarce environments as is the case of the Mara basin.

A number of studies in the Mara Basin have been carried out to determine the hydrological response of the basin to changing conditions. Mango et al. (2011) has studied these changes using SWAT model, however their study did not result in satisfactory simulations of the Mara hydrograph. This has partly contributed to the limited hydrological information that may inform management decisions.

The challenge in the Mara basin is that of lack of sufficient information that creates an impetus for basin managers to make informed decisions and take appropriate actions with respect to anthropological activities, in management of the basin to ensure sustainability of this resource. This challenge can be partly addressed through modeling the hydrology of the upper reaches of the Mara river basin. The developed model can then be used to simulate changing land use conditions to generate scientific data that informs policy and management decisions in an environment of data scarcity as well as improve understanding of the underlying hydrological processes.

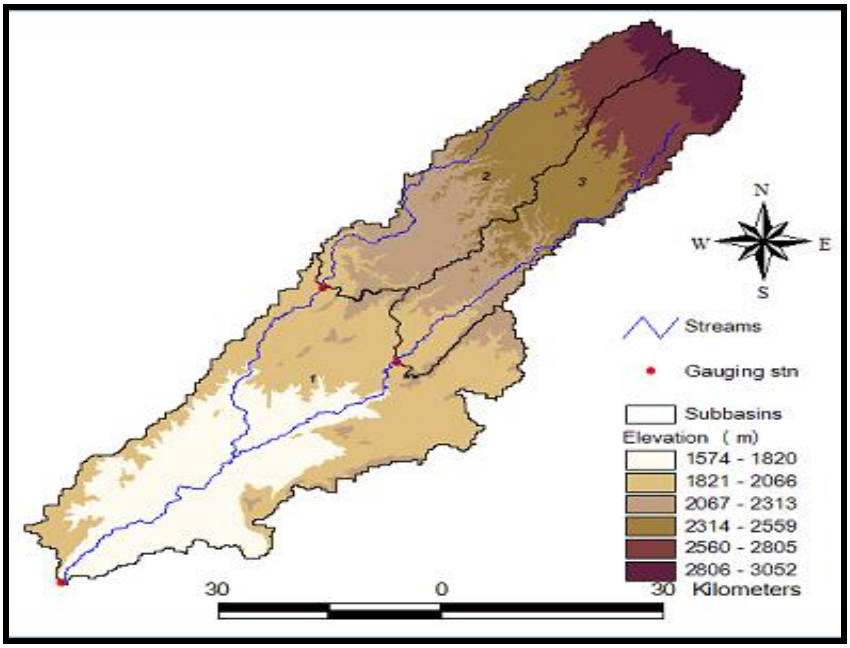

Figure 1.4 Upper Mara DEM

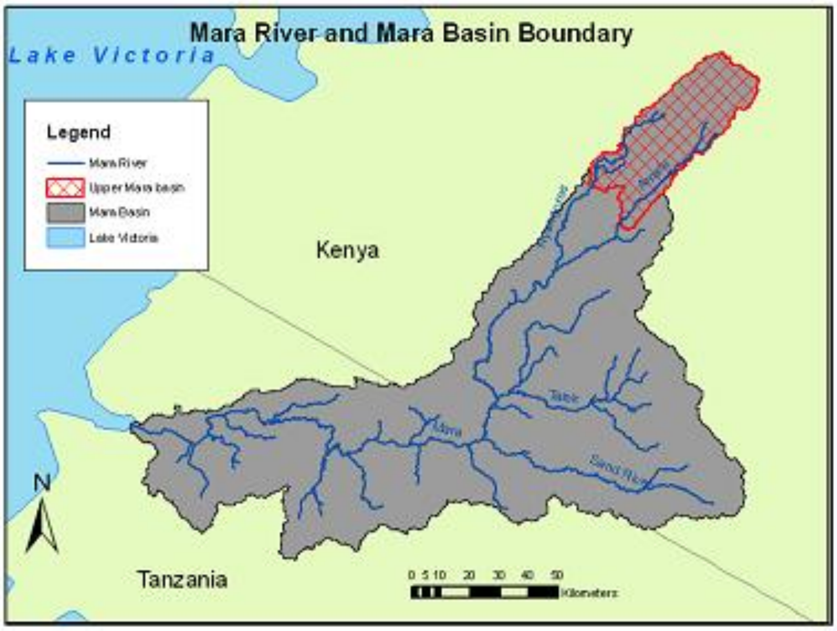

Figure 1.5 The Upper Mara -Nyangores and Amala Catchments
Source: (Mango et al., 2011)

1.3 Research Justification

The Mara basin supports an estimated population of 573,883 people on the Kenyan side and is home to the world famous Maasai Mara and Serengeti ecosystem (Barno, 2011). The upper reaches of the Mara are critical for the preservation of the Mara ecosystem. Over 75% of runoff in the basin is derived from the upper reaches (LVBC & WWF-ESARPO, 2010).

Over the past 40 years in conjunction with significant population growth, the area under agriculture has tripled and now covers approximately 3000 km^2 (Mati et al., 2008). The highest concentration of agriculture is distributed among small scale farmers on the slopes of the Mau escarpment and upper middle reaches of the basin. Increase in agricultural lands has occurred in tandem with a 34% decrease in rangelands and a 32% decrease in forested areas (Mati et al., 2008). This land use change is believed to have impacted on the hydrological processes and the water availability in the downstream part of the basin, there is need to mitigate the hydrological impacts of the changing land use as well as pre-emptive actions to forestall any impacts of climate change, to ensure sustainability of the Mara Ecosystem.

1.4 Research Objective

The main objective of this research is to improve the hydrological modeling effort in the Mara river basin, to improve prediction of the response of the Mara hydrograph to land use changes and generate scientific data that could inform policy and subsequent decision making in the management of the Upper Mara and inform future scientific work, in terms of better understanding of the Mara hydrological processes.

1.5 Research Questions

- How will the hydrograph of the upper Mara be affected by land use changes in the upper reaches of the basin?
- How does the performance of the developed model, in terms of simulating the Mara hydrograph compare to those of previous modeling efforts in the Mara basin?

1.6 Outcomes

The outcomes of this research are three fold:

- Development of a distributed hydrological model that simulates the Mara hydrograph
- Prediction of the response of the Mara hydrograph to changing conditions of land use.

1.7 Hypothesis

The performance of a simple, conceptual distributed hydrological model in modelling the river discharge of the upper Mara is comparable to that of complex models.

2 LITERATURE REVIEW

2.1 Water Availability

The Mara River has an average annual flow rate of 30 m^3/sec and has a seasonal flow regime (LVBC & WWF- ESARPO, 2010). It experiences two periods of high flow extending from March to May and October to December separated by periods of marked low flows. The river is reported never to have failed to flow in the more than 60 years that discharge has been monitored and there are no reports of the river having ever dried up. However the increasing anthropological activities in the basin marked by increased abstractions and land use change pose a serious threat to the continual flow of the river.

2.2 Climate and Land use change

It is no longer in doubt that climate change is taking place and it is up to society to develop strategies to adopt to changes it is likely to bring about. Science is looked upon as the source of answers on how to respond to the impacts of climate change. Applied science and research should inform choices in dealing with climate change and its subsequent impact on the landscape, this combined with the effects of anthropogenic activities.

2.2.1 Climate Change Assessments

The Intergovernmental Panel on Climate Change (IPCC) was established by the World Meteorological Organisation (WMO) and United Nations Environmental Programme (UNEP) to assess scientific, technical and socio economic information relevant for the understanding of climate change, its potential impacts and options for adaptation and mitigation. Its main activities are to assess the state of knowledge on climate change based on peer reviewed and published scientific literature and provide assessments of climate change at regular intervals. In this context IPCC defines climate change as "any change in climate over time, whether due to natural variability or as a result of human activity" (IPCC, 2007).

In the African tropics climate is highly variable in terms of rainfall, river flow regimes and lake levels, which have enormous socio economic impact. In most regions there is very limited time series meteorological and hydrological data both in space and time.

General Circulation Models (GCM) that predict long term trends in climate are often unsuitable for regional scale studies because of the coarse grid-size resolution. Consequently, there is need for tools that can be used to assess the likely effects of land use changes as well as climate variability on the hydrological cycle at catchment scale (Legesse et al., 2003). Despite this limitation, the IPCC predictions based on GCM offer the most reliable assessments on the state of the climate in the future. According to Wang et al. (2008) climate change affects the amount and distribution of regional precipitation and temperature, thereby influencing catchment runoff. It follows that the influence of climate change on catchment balance is of paramount importance in hydrological studies.

2.2.2 Land use Change

Due to the effects of changing climate as well as anthropogenic activities the landscape is rapidly changing in response to the needs of a growing human population. Multiple forces of change such as demographic trends, climate variability, national policies and macroeconomic activities result in extensive alterations in land cover and land use, which in turn impact the hydrologic system both at basin and regional scale (Legesse et al., 2003). The Mara basin has witnessed marked land use change in the recent past (See Fig 2.1 and table 2.1). Studies by Mati et al. (2008) indicate increased agricultural activities in the recent past have resulted in conversion of large forest plantations into farm land for cultivation of seasonal crops replacing natural perennial vegetation, this is further aggravated by a rapid population growth and migration rates.

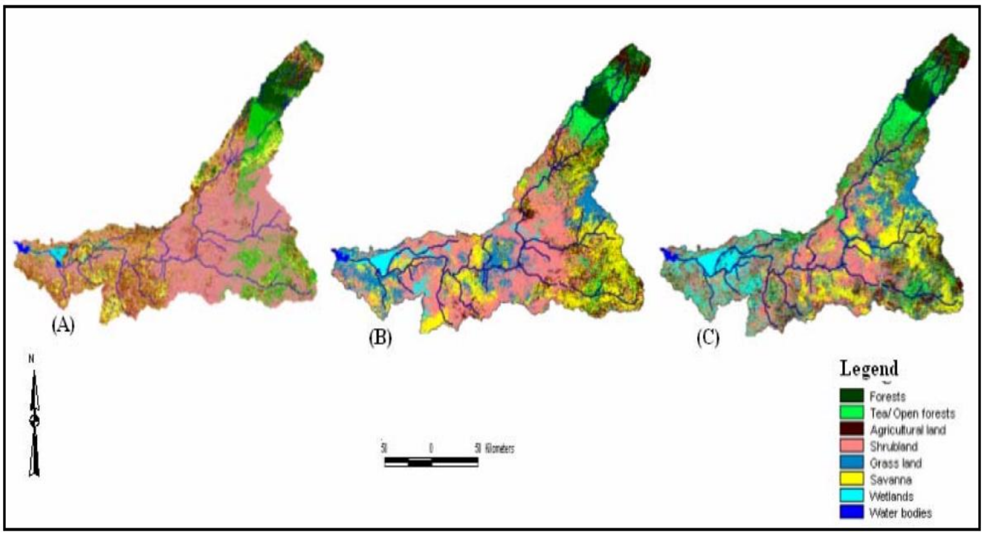

Figure 2.1 Land Cover Maps for 1973, 1986 and 2000

Source: Mutie et al., 2006

This has resulted in increased water demands. Sustainable management of the Mara resources requires accurate projections of the impacts of these changes on the hydrology of the Mara River hence this research.

Table 2.1 Land Use/cover area changes in the Mara River Basin, 1973-2000

Land cover type	1973(km^2)	1986(km^2)	2000(km^2)	Change (1973-2000) km^2	Change (%)
Forests	1008	893	689	-319	-32
Tea/open forests	621	1073	1948	+1327	+214
Agricultural Land	826	1617	2504	+1678	+203
Shrub land	5361	5105	3546	-1815	-34
Grassland	2465	1621	1345	-1120	-45
Savannah	3163	2867	2354	-809	-26
Wetlands	286	604	1394	+1108	+387
Water bodies	104	54	55	-49	-47

Source: Mati et al., 2008

2.2.3 Land Use change and Hydrology

Population growth triggers changes in land use, as an increasing population requires a habitat and increased demand for food. This invariably means that more land will be required to achieve basic needs resulting in change in land use. Changes in the landscape often have an effect on the hydrology of a river by alteration of infiltration rates and transpiration, and these have an impact on runoff generation and recharge and hence river flow characteristics.

Changes in hydrological regimes due to land management changes are often related to two partitioning points. The first partitioning point determines effective rainfall, which is rainfall minus evaporation and divides the flow into surface runoff and infiltration. The governing processes here are infiltration rate and interception which depend on vegetation and soil surface roughness. How much of interception occurs depends on land cover characteristics rainfall characteristics and evaporative demand. Interception can amount up to 15-50% of precipitation (Gerrits, 2010).

The second partitioning point separates the infiltrated water into transpiration, evaporation from the soil and deep percolation to groundwater. Vegetation cover and crop stage influence partitioning at this point. Transpiration rates are different for different vegetation covers and forests are known to transpire at higher rates than shrubs or agricultural land (Mul, 2009).

Several techniques have been used to assess the impacts of land use change on hydrology. Brown, (2005) used the paired catchment approach whereby response of two neighbouring catchments is compared for different land uses. The differences in response are attributed to the different land uses. Paired catchment studies can provide evidence of impacts of land use change on river hydrology, however they require long periods of study and cover small study areas. (Wang et al., 2008)

Hydrological modelling is often used to identify the possible changes to a hydrograph as a result of catchment response to changes in land use (Uhlenbrook et al., 2004).

In identifying the impacts of land use change Brown (2005) found that when forested catchments are compared to deforested catchments; the deforested catchment has a smaller time of concentration, peak flows are larger, whereas base flow is smaller due to reduced infiltration and reduced recharge of ground water.

.

2.3 Hydrological Modelling

2.3.1 Hydrological Modelling overview

Models are today routinely used for prediction of responses of a catchment or river basin to a rainfall event. They serve the purpose of providing decision makers with information to make decisions and inform policy for basin management as well as facilitating scientists to understand the underpinning hydrological processes.

There are many reasons why modelling of the rainfall runoff process takes place, key amongst these reasons is that we are limited in the measurement of hydrological processes and the most difficult being that most of the runoff generation processes take place below the ground surface presenting challenges of measurement, modelling thus offers a practical way of trying to understand these processes.

The construction and application of watershed models describing precipitation to stream flow processes has been a prime focus of hydrological research and investigations for many decades (Jakeman and Hornberger, 1993). According to Beven (1996), "Every hydrological model requires two essential components, one to determine how much of rainfall becomes part of the storm hydrograph and the other takes account of the distribution of the runoff in time, to form the shape of the hydrograph otherwise referred to as routing".

Hydrological models provide a framework to conceptualise and investigate the relationships between climate, land use changes and water resources. Hydrological models can basically be classified into three categories, empirical or black box, conceptual or grey box and physically based distributed or white -box models. Black box models do not explicitly consider governing physical laws of the process involved but only relate input to output series. The grey box models represent the effective response of an entire catchment without attempting to characterise the spatial variability of the response explicitly. The white box models are able to explicitly represent the spatial variability of some if not most of the important land surface characteristics such as topographic elevation, slope, aspect, vegetation, soil as well as climatic parameters including precipitation, temperature and evapotranspiration distribution (Legesse et al., 2003).

Physically based models are normally more feasible as research tools for process studies in the small scale where physical parameters are well under control and their variability small. Conceptual models are more basin oriented than physically based models. The parameters of a conceptual model thus represent an average over a large area and often integrate several processes and their variability. Conceptual models show their strength in limited data demand and thus great applicability in operational hydrology. Physically based models are often said to be superior to conceptual models as they demand less calibration or tuning of the parameters (Bergstrom, 1991).

Legesse et al. (2003) further argues that "because white box models relate model parameters directly to physically observable land surface characteristics, spatially distributed hydrological models have important applications to the interpretation and prediction of the effects of land use change and climate variability".

2.3.2 Hydrological modelling approaches

It is commonly argued that process based fully distributed models are best suited to simulate land use change effects, since in most cases only part of the land use within a basin changes, spatially distributed models depict these changes more precisely as compared to lumped modelling approaches. According to Breuer et al., (2009), the above mentioned reasoning has resulted in development of numerous complex models such as MIKE-SHE, TOPLATS and WASIM amongst others. They however, face difficulties of parameter estimation. An alternative is semi distributed models in whose class are SWAT, SWIM, SLURP and PRMs amongst others. These models simulate all hydrological processes within spatially non explicit Hydrological Response Units. The results of each HRU are lumped and routed downstream. This group of models require a considerable number of parameters that are difficult to acquire. Breuer et al. (2008) further argues that a further simplification is achieved if hydrological fluxes are simulated with the sub catchment scale as the smallest spatial unit. In this category are models such as HBV and LASCAM which have a coarse spatial resolution in their simulations. At the lower end of complexity are models such as IHACRES and NAM which are conceptual lumped models with a simple model structure and a small number of parameters. The vast majority of models used in Rainfall -runoff modelling are used in a deterministic way (Beven, 1996). It is argued that data limitations in many catchments limit the applicability of physically based models and that conceptual models provide a more appropriate alternative (Breuer et al., 2008). In a model inter comparison study carried out by Breuer et al. (2008) on modelling effects of land use change on hydrology it was found that the more conceptual models outperformed the physically based, fully distributed models.

2.3.3 Hydrological Modelling in Data Scarce Environments

In order to make appropriate management decisions regarding water resource management, there is need to have accurate projections under changing conditions. Modelling has long been used to predict future changes, but in order to do so there is need for accurate data. According to Mango et al. (2011),"Accurately modelling future runoff regimes is challenging in African catchments with limited current and historical runoff data, but an increasing number of model applications are possible" They further argue that using coarse datasets such as 90m resolution Shuttle Radar Topography Mission (SRTM) DEM, a 1:50,000 scale land use map, course soil data sets (1:1000000) and incomplete rainfall data, it is possible to simulate the hydrology of a river basin. They however caution that whereas several modelling efforts have been carried out using such coarse data sets, caution must be applied when interpreting and communicating such results as their value must been viewed from both heuristic and algorithmic terms.

Remote sensing offers options to dealing with data scarce environments, according to Winsemius et al. (2006). "Remotely sensed data offers a wealth of spatially distributed information which can be used to identify and parameterize relevant hydrological processes at smaller spatial scales".

2.3.4 Simplification in Hydrological modelling

Hydrology has witnessed an enormous growth in the twentieth century due to technological and methodological advances. The advent of powerful computers, remote sensors, Geographic

Information Systems (GIS) and worldwide web and networking have facilitated extensive data collection, better data sharing, development of sophisticated mathematical methods and development of highly complex models. However,there are also concerns that these advances indirectly contribute to additional problems in research as a result of development of complex models (Sivakumar, 2008b).

As a result of these technological advances the trend today is to "model everything" rather than modelling the essential processes only.

Sivakumar, (2008b) further observes that these advances raise two major concerns; they lead to complex models that have too many parameters and thus require more data than are actually needed; and secondly since these models are developed for specific situations, their extensions and generalizations to other situations is difficult.

According to Jakeman and Hornberger, (1993), the performance of such complex models can be worse than simple models because of the possible propagation and amplification of the deficiencies of the scientific concepts or system information or both, and these are manifest as over parameterization, and other related problems.

Climate acts as the unifying global force in the co evolution of landscapes and vegetation (Sivapalan, 2003). Hydrologic processes arise as a result of interactions between climate inputs and landscape characteristics that occur over a wide range of space and time (Sivakumar 2008a). Subsequently, the purpose of hydrological modelling is twofold, first to understand the processes and secondly for prediction to address specific problems. Current knowledge through general observations indicates that only few processes dominate hydrologic response in a given catchment and experience through modelling, parameter estimation and prediction indicates that simple models with only a few dominant parameters could capture the essential features of a given catchments response to hydrologic events (Sivakumar, 2008b). In addition, according to Jakeman and Hornberger (1993) several models of varying complexity when applied to a number of catchments result in the conclusion that simpler, less data intensive models provided as good or better predictions than a more physically based model.

As a result of the above observations, the "dominant processes concept" has emerged. It is premised on the idea of developing methods to identify the dominant processes that control hydrologic response in a catchment and then developing models to focus on these dominant processes (Sivakumar, 2008a). It is further suggested that one possible way to do this is through devising a procedure that starts with the simplest reliable situation and then move to the more complex potentially required solutions. Sivakumar, (2008b) argues that while advances in technology are attractive for developed countries, these are not an option for most developing countries where data are scarce and consequently developing more and more complex models that require more and more data is not an option.

This argument is further supported by Jakeman and Hornberger (1993), who argue that despite the fact that the physics governing the path of a drop of water through a catchment to the stream involves complex relationships, evidence suggests that the information content in a rainfall - runoff record is sufficient to support models of only very limited complexity.

It therefore follows that modelling efforts in data scarce environments like the Mara basin inherently involve dealing with limited data and hence simple models may be the most practical way to go. This can be done in the knowledge that "The development of more and more complex models that incorporate more and more details about processes, but which introduce more and more parameters that must be calibrated, does not appear to be the future. The future of modelling will have to place more emphasis on the value of data, carefully collected for specific purposes and on parametrically simple robust models, carefully designed for specific purposes" (Beven, 2002 as cited by Sivakumar, 2008b).

2.3.5 Evaluation of Hydrological Models

Advancements in the science of hydrology and hydrological modelling have resulted in the development of hundreds of hydrological models. It is inevitable today that current modelling efforts involve numerous comparisons, resulting in an obvious need for platforms to base such comparisons or evaluations. The process of evaluating the performance of models is important not only at the stages of model development and calibration but also during the communication of results to other researchers and stakeholders (Schaefli and Gupta, 2007).

According to Sivakumar (2008b) "Model comparisons are good provided the purpose is not simply to compare different models but to identify the specific advantages and limitations of each of the models. This would provide important clues towards possible integrations of the two or more models/concepts, making the most of their advantages and minimizing their limitations for better representation of hydrologic systems and processes".

Hydrographs are time series of hydrological variables. They are key in evaluation and comparison of hydrological modelling efforts. This is often done by comparison of simulated values against observed. According to Ewen (2011), "The gold standard in hydrograph comparison is manual inspection by hydrologists, because even the best available automatic methods are poor substitutes for the hydrologist's eye and brain, especially at spotting and interpreting patterns".

The Mean Square Error (MSE) and the related normalisation, the Nash - Sutcliff Efficiency (NSE) are the two criteria most widely used for calibration and evaluation of hydrological models with observed data (Gupta et al, 2009). According to Schaefli and Gupta (2007), "the NSE is a normalized measure (-inf to 1.0) that compares the mean square error generated by a particular model simulation to the variance of the target output sequence. In doing so it represents a form of noise to signal ratio comparing the average "size" (variability) of model residuals to the "size" (variability) of the target output. It is implicitly comparing the performance of the particular model to that perhaps the simplest imaginable model, one that uses as its prediction the (constant) mean value of the observed target.

$$NSE = 1 - \frac{\sum_t^N = 1[q_{obs}(t) - q_{sim}(t)] \quad ^2}{\sum_t^N = 1[q_{obs}(t) - q_{sim}(t)] \quad ^2}$$

Where $q_{obs}(t)$ = discharge at time step t
$q_{sim}(t)$ = simulated discharge
$\overline{q}_{obs}$ = mean observed discharge over simulation period N
(Nash and Sutcliffe, 1970)

This means that an NSE value = 1.0 indicates perfect model performance(the model perfectly simulates the target output) and NSE value = 0 indicates the model is on average performing only as good as the use mean target value as prediction and an NSE value of < 0.0 indicates an altogether questionable choice of model. It is therefore preferred to have NSE values larger than Zero (0) and approaching 1.0.This corresponds to an apparent normalisation because the implicit reference model has different implications for different case studies." It therefore follows that the NSE does not measure how good a model is in absolute terms. (Nash and Sutcliffe, 1970).

However according to Ewen (2011), one of the strengths of the NSE, as a result of being based on the square, is that it is sensitive to differences for peaks. One of its weaknesses is that it is also quite sensitive to differences in timing. Handling and interpreting differences in timing is one of the most difficult problems faced when computing hydrographs. The sensitivity of NSE to timing arises because even quite small misalignments in the timings of peaks can give rise to large differences in amplitude between the hydrographs. This sensitivity can result in poor value for NSE even when the size and shapes of the peaks in the two hydrographs are very similar. However it is common practice for modellers to show hydrograph time series plots in which model simulation" goes up and down" in a similar fashion as that which is measured as an indication of modelling success (Schaefli and Gupta, 2007).

2.4 Modelling efforts in the Mara Basin.

Currently limited historical hydrological and hydro chemical data exists for the Mara river basin (Mati et al., 2008). Consequently there is limited understanding of the hydrology of the Mara Basin, which was once fairly well gauged but most of the stations are no longer working with either staff gauges washed away or no data recording taking place now.

2.4.1. Approaches

Hydrological studies in the Mara have mainly focussed on modelling the impacts of land use changes such as those by Mati et al. (2008) and Mango et al, (2011). These studies have mainly focussed on the upper Mara consisting of either the Nyangores, Amala or both major tributaries of the Mara. Mati et al. (2008) studied the impacts of land use/cover changes on the hydrology of the transboundary Mara River, in Kenya and Tanzania. Modelling changes in the flow regimes between 1973 and 2000 was established with the geospatial stream flow Model, GeoSFM, a physically based semi distributed geospatial hydrological model by Mutie et al. (2006). The model used remotely-sensed data, numerical weather forecast data, ground observation and geographical datasets that describe the soils and land surface to calculate different parameters of basin hydrology.

On the other hand Mango et al. (2011) investigated the land use and climate impacts on the hydrology of the upper Mara river basin. Mango used the SWAT model, which is a semi distributed model. The SWAT model requires the use of spatially explicit datasets for land topography, land use or land cover, soil parameters for hydrological characteristics and climate and hydrological characteristics, and climate and hydrological data on daily time steps (Mango et al., 2011).

2.4.1.1 Modelling Mara using SWAT

The Soil and Water Assessment Tool (SWAT) is a physically based continuous event hydrological model developed to predict the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time (Arnold et al., 1998). For simulation purposes, a watershed is divided into a number of homogeneous sub basins (Generally referred to as hydrologic response units or HRUs) having unique soil and land use properties. The input information for each sub basin is grouped into categories of weather, unique areas of land cover, soil and management within the sub-basin, ponds/reservoirs; groundwater; and the main channel or reach draining the sub basin. The loading and movement of runoff, sediment, nutrient and pesticide loadings to the main channel is simulated considering the effects of several physical processes that influence the hydrology.

Hydrological processes in SWAT are modelled in six categories;

- The Evaporation model in SWAT uses canopy interception based on leaf area index (LAI), the potential evaporation has three options, Penman-Monteith method, Priestly Taylor and Hargreaves method, the actual evaporation calculated from interception, transpiration from plants and evaporation from the soil.
- The surface flow is modelled using surface runoff volume calculated using the SCS curve number method and the Green and ampt infiltration method.
- Surface routing is calculated based on empirical formulae incorporating time of concentration, runoff lag and transmission losses amongst others.
- The unsaturated zone is modelled using lateral flow layers modelled using kinematic approximations.
- The saturated zone flow in SWAT model uses a two aquifer system, confined and unconfined defined per sub-basin.
- River flow is modelled using uniform flow formula to compute velocity and discharge based on a trapezoidal channel cross section and routed using variable storage and Muskinghum methods.

The main input data for SWAT are the climatic data, soils data, a Digital elevation model and land use data (www.brc.tamus.edu).

2.4.1.2 Modelling Mara using GeoSFM

Mati et al. (2008) investigated land cover changes in the Mara using the distributed United States Geological Survey (USGS) geo spatial stream flow model. This model is a physically based semi-distributed geospatial hydrologic model. It operates as an extension within Arc view and uses data in GIS formats, the model consists of two parts, a GIS based module used for model data input and preparation and a simulation module for rainfall runoff. The model uses remotely sensed input data such as soil, land cover, rainfall and evaporation data. The topography of the land was obtained from SRTM of the year 2004. The soils data was obtained from Digital Soil and Terrain Database of East Africa (SOTER) soil map from FAO, 1997 and land use data for the years 1973, 1986 and 2000 respectively was obtained from Land Sat MS TM. The Hydro meteorological data was obtained from both Kenya and Tanzanian meteorological departments.

2.4.2 Outcomes

Mati et al. (2008) concluded from the simulation studies that significant changes have occurred in the flow regime of the Mara between the periods 1973 and 2000 resulting in increasing and earlier occurrences of high flows. This was corroborated with observed changes in the basin, such as reducing vegetation land cover and the expanding wetland at the mouth of the river resulting from increased erosion. Mango et al. (2011) on the other hand modelled the Nyangores tributary hydrograph using the SWAT model and developed land use and climate change scenarios. The climate change scenarios were developed from projected climate change as indicated in the IPCC fourth assessment report (IPCC, 2007), the variables usually considered from these GCMs are precipitation and temperature.

2.4.3 Research Gaps

In the case of modelling the Mara hydrograph by Mango et al. (2011), the results of modelling were modest in nature and the developed SWAT model did not simulate the hydrograph to satisfactory levels, with model evaluations of an NSE of - 0.53 and R^2 of 0.09 for the calibration period and NSE values of -0.06 and R^2 of 0.32 for the validation period. These figures apply to Nyangores for simulations using rain gauge data. There was consistent overestimation of the discharge when visually comparing the observed against the simulated for Amala and underestimations for Nyangores (See Fig 2.2 and 2.3).

According to Mango et al. (2011) "the modest performance of the model does not justify detailed analysis of differences between land use change scenarios even though trends and relative magnitudes of impacts are evident." This modelling effort was thus considered "an exploratory analysis and evaluation of trends describing the response of the Mara River Basin to future land use and climate change scenarios." The poor model performance illustrates the challenges of modelling in data scarce environments which Mango et al. (2011), mainly attributes to using data from limited rain gauges due to very coarse distribution of climate stations in the catchment. This happened despite efforts at augmenting rainfall data using remotely sensed data from FEWS network. The results of these modelling efforts illustrate the need for more efforts at modelling the Mara hydrograph to reliable levels, without losing focus on the need to develop modelling approaches that work better in data scarce environments and that promote better understanding of the underlying hydrological processes.

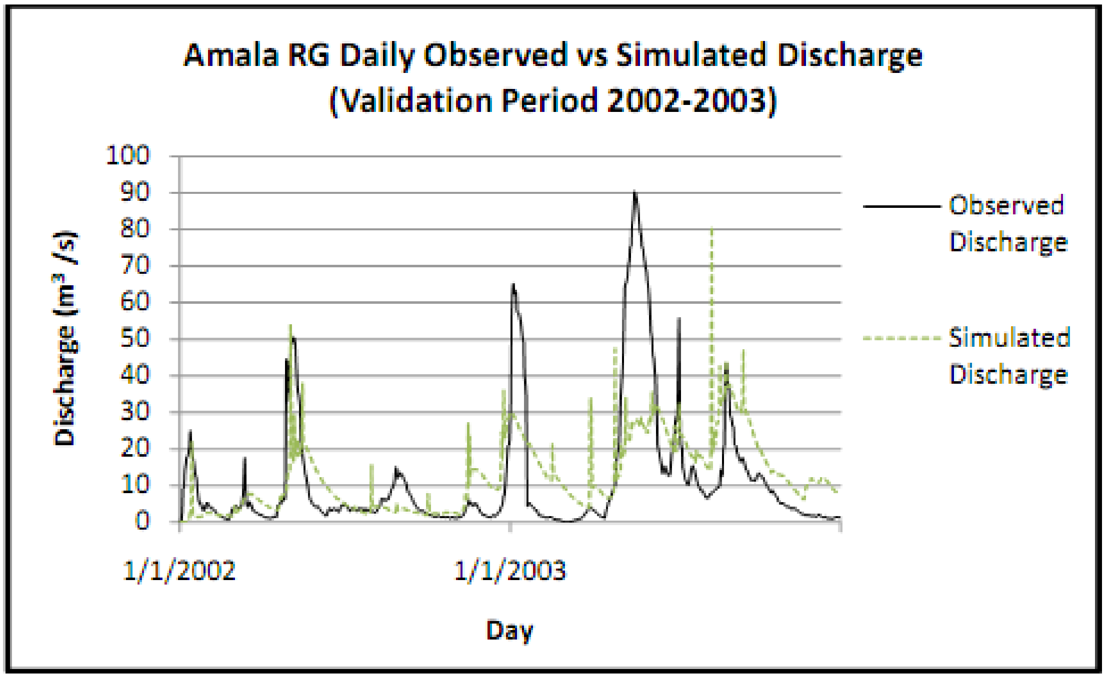

Figure 2.2 Amala Observed vs Simulated Discharge

Source: Mango et al., 2011

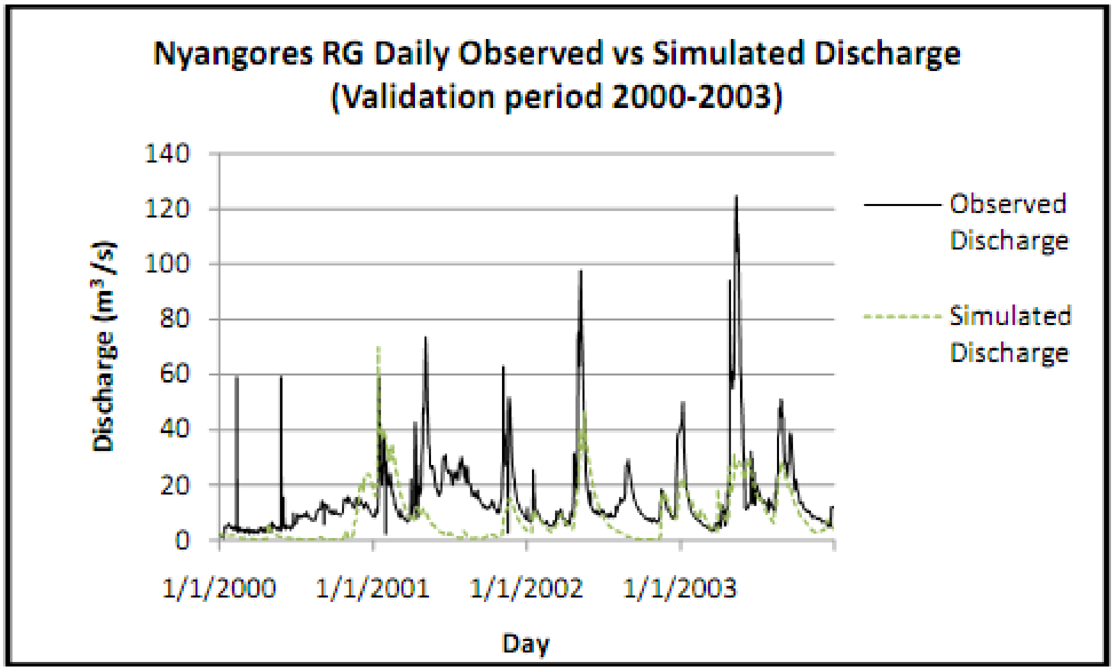

Figure 2.3 Nyangores Observed vs Simulated Discharge

Source: Mango et al., 2011

In the case of modelling using the USGS geo stream flow model, the model was calibrated for a period of six years from 1983 - 1988. The R^2 values obtained for Amala, Nyangores and Mara mines gauging stations were 0.76, 0.74 and 0.83 respectively. Validation was done for the period 1989 to 1991 with R^2 values of 0.72, 0.69, and 0.87 for Amala, Nyangores and Mara mines respectively.

Findings of this study indicate that the model performed better for bigger catchment areas thus posting better R^2 values at Mara mines gauging station. In addition the model accurately simulated the hydrograph rising limbs and peaks, but was unable simulate accurately the recession limbs and low flows.

From the two models discussed above, it is evident that, the success of modelling the Mara has had mixed model performance with SWAT performing poorly. The GeoSFM model however, posts better results than SWAT, but its temporal scale does not promote understanding of the underpinning hydrological processes, and is more suited to understanding trends.

There is need for a model that simulates the Mara hydrograph at acceptable performance levels, and promotes understanding of the hydrological processes. This would in turn generate information needed for basin management. Relevant, accurate and timely management information would serve as an impetus to basin managers to take appropriate measures in management of the resources of the Mara basin.

2.5 Review of hydrological model structures

There are numerous model structures that have been developed over time. This research has adopted from existing conceptual model structures,

From literature three (3) potential model structures were identified and considered these were:

- The Spatial Tools for River Basins and Environment and Analysis of Management Options(STREAM) structure, (Aerts et al., 1998)
- Lumped Elementary Watersheds structure (LEW), (Winsemius et al., 2006)
- TAC^D structure (Uhlenbrook et al., 2004)

The three model structures were evaluated based on the following considerations, the complexity of the model structure, data requirements, flexibility for adaptation, suitability for modelling land use change and previous modelling success in tropical environments.

The TAC^D structure is a complex structure based on using data derived from tracer studies aimed at understanding hydrological processes whereas the LEW model structure is a custom structure adopted specifically for the Zambezi basin. STREAM is a simple model structure, requiring a minimum set of data inputs and easily adaptable. Both the STREAM structure and the LEW structures have been tested in tropical environments. All the three structures are distributed model structures and are thus suited to modelling land use changes.

2.5.1 Modeling Philosophy

The key factors informing choice of model structure are simplicity and robustness of the model. Consideration was made for a distributed conceptual model. A conceptual model is preffered to reduce the challenges associated with model complexity such as equifinality, parameter identification and calibration amongst others. To prevent equifinality (Savenije, 2001) the number of calibration parameters need to be kept to a minimum. A distributed model is recommended more so when the eventual intension of model development is to project the impact of land use change.

The general structure of a conceptual model for rainfall - runoff is illustrated in Fig 2.4 below:

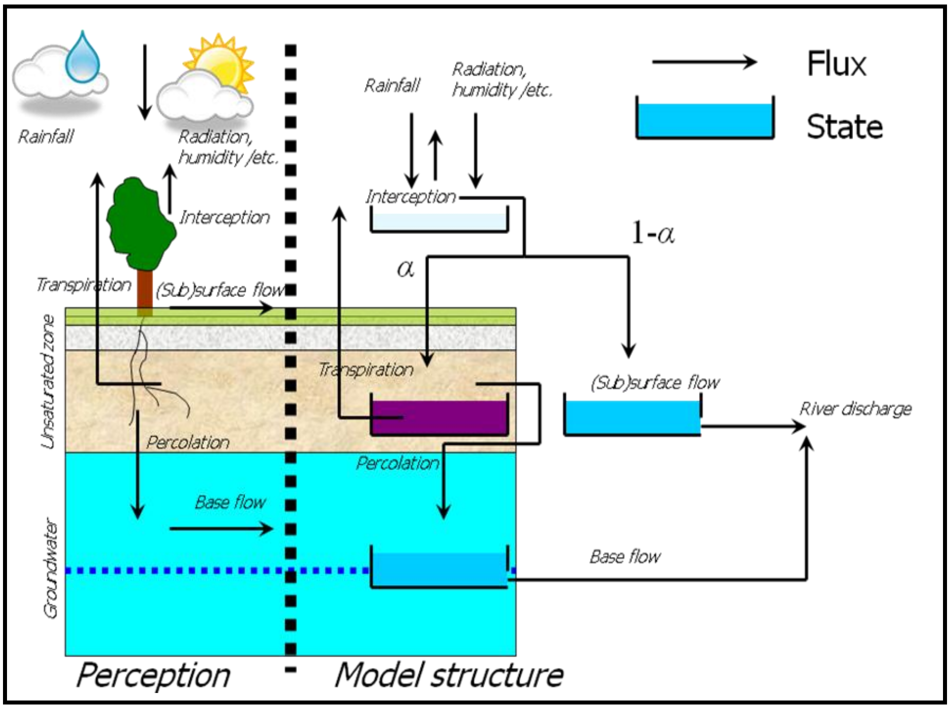

Figure 2.4 Conceptual Rainfall-Runoff Model structure.

Source: Deltares

2.5.2 Choice of Module Structure

Preference of model structure for adoption for this research was based on consideration of its simplicity, parameterisation, especially parameters whose values are derived from calibration and previous modelling success. In addition the structure should satisfy the following basic criteria:

- The amount of parameters needed should be limited and should be derived as much as possible from the available data.
- The conceptual structure should be as representative as possible of the Mara river basin
- The structure should have calibration parameters that are identifiable.

2.5.3 Modelling Environment

PCRaster belongs to a group of software referred to as Environmental Modelling languages (EML).

These are a group of languages that run inside a GIS to benefit from powerful database and visualisation functions that a GIS environment offers.

PCRaster is a collection of software targeted at the development and deployment of a spatio temporal environment. It is a scripting model development environment that allows users to develop their own simulation models, such as rainfall-runoff models, vegetation competition models amongst others. It consists of a rich set of model building blocks and analytical functions for manipulating Raster GIS maps. The language consists of 125 generic functions. These include mathematically defined non-spatial (Point) operations, network creation, transport and flow operations and window operations as well as spatio - temporal time operations for reading and writing temporal data. The program contains a GIS database and is integrated with visualization software (Karssenberg, 2001)

The model development cycle in PCRaster consists of two (2) steps

1. The conversion of the conceptual model structure into a numerical expression.
2. The assimilation or calibration of the model with observed data

2.5.3.1 Modeling Environment justification

PCRaster environmental modeling language, according to Karssenberg (2002) is based on the following approaches:

- Provides a set of operators operating on spatio- temporal data in which widely accepted generic hydrological processes have been coded using clearly understood algorithms.
- Provides operators in a suitable way that, they can be glued together in a model by a hydrologist using his or her hydrological understanding, rather than computer expertise.
- It embeds a set of tools for model construction in a GIS-like software environment providing database management and generic visualization routines for the spacio-temporal data read and written by the model.
- It provides standard interfaces to other programming languages so that new or alternative operators can be added by the user in ways that are fully compatible with the EML.

PCRaster uses a raster based approach and as such is powerful in distributed modeling of hydrological processes. In addition it offers a dynamic programming environment. Karrensberg (2002) further states "the environmental modeling languages will be used in the future mainly for development of new models that can be tailored to modeling aims and the field data available."

It is these properties that make PCRaster a suitable environment for development of a model aimed at modeling in data scarce environments and subsequent projection of possible impacts of spatial changes (land use change).

3 METHODOLOGY

To model the hydrograph of the upper Mara, a distributed conceptual model was used. A conceptual model was chosen to reduce the challenges associated with model complexity such as equifinality, parameter identification and calibration amongst others. To prevent equifinality the numbers of calibration parameters are kept to a minimum. A distributed model is preferred more so when the eventual intension of model development is to project the impact of land use change.

3.1 Hydrological model description

The model structure chosen for modelling the Mara hydrograph is a distributed conceptual model, where the non linear behaviour of the river basins is explained by a combination of thresholds and linear reservoirs. The model structure adopted is the STREAM model. The idea behind the STREAM is to model the hydrology of the basins in a simplified way but which at the same time allows sufficient insight into the major hydrological processes which take place in the basin (Aerts et al., 1999). The model is based on a raster GIS which calculates the water balance of each grid cell and routes this through a stream channel network which is based on the digital elevation model (DEM).

3.1.1 STREAM Structure

The structure of the STREAM model adopted is illustrated below; it is adopted from the script used to model the Zambezi basin (Winsemius et al., 2006).

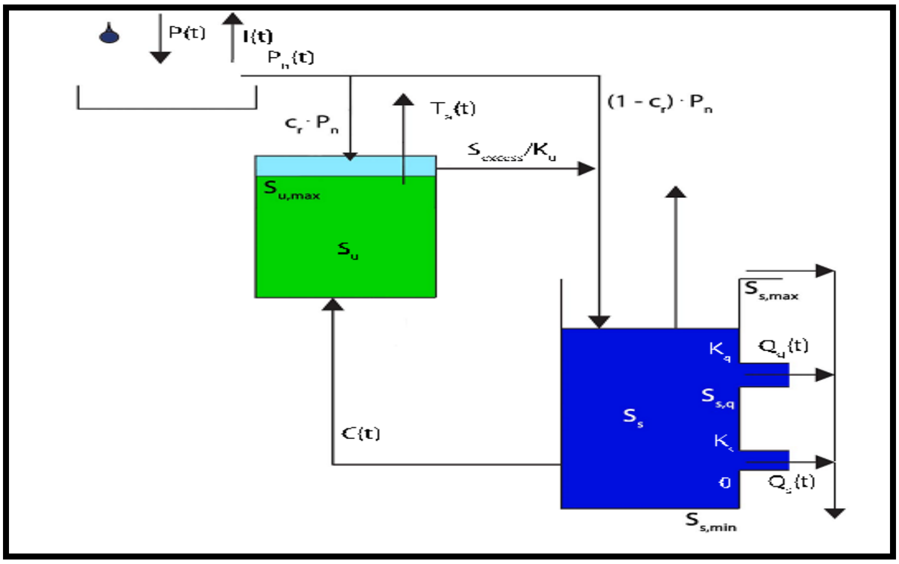

Figure 3.1 STREAM model structure

Source: Winsemius et al., 2006

3.1.2 Interception

The STREAM model structure is illustrated by Fig 3.1. When precipitation occurs over a landscape, not all of it infiltrates into the sub surface; some is lost to the atmosphere in the process of moisture recycling. This part of moisture recycling consists of several components (Savenije, 1977 as quoted by Gerrits, 2005).These are:

- Canopy interception
- Shallow soil interception
- Evaporation from temporary surface storage(pools)
- Immediate transpiration (within a month)

In this structure fast evaporation consists of interception and transpiration from shallow rooting vegetation which is generally referred to as interception. All these Interception processes are modelled with one simple threshold value *D*. By subtracting the threshold *D* from the precipitation *(P)* the net precipitation *(P_n)* is obtained. Since it is not possible to evaporate more than precipitation, net precipitation is minimally set as zero (See formulae 1, 2, and 3). The threshold value *D* is a calibration parameter.

$$P_n = \max\,(P - D,\ 0) \tag{1}$$

$$I = \min\,(P,\ D,\ E_p) \tag{2}$$

$$P_n = P - I \tag{3}$$

Where

I = Interception [mm month $^{-1}$]

P = Precipitation [mm month $^{-1}$]

E_p = Potential evaporation [mm month $^{-1}$]

P_n = Net precipitation [mm month $^{-1}$]

D = Interception threshold [mm month $^{-1}$]

After the interception process, the water infiltrating into the subsurface is divided into two components, that which goes into the saturated zone, this fast component flows through macro pores and cracks and the other component flows into the unsaturated component. Thus net precipitation P_n is separated over the unsaturated and saturated zone using a separation coefficient C_r [-].

3.1.3 Transpiration

The water that flows into the unsaturated zone is available for transpiration. Transpiration is a process of water uptake by plant roots, its transportation through the plant and its loss through stomata in the leaves. However not all water that infiltrates into the unsaturated zone is available for transpiration. This is limited by how much water the Soil is able to hold, referred to as the field capacity and subsequently by the available moisture content, the amount that plants can easily extract. Actual transpiration (*T_a*) is defined by the following equation as calculated by Winsemius et al. (2006).

$$T_a = min\left(\frac{1}{0.5 \cdot Su, max} \cdot T_p \cdot S_{u,}, T_P\right) \qquad \boxed{4}$$

Where

T_a = Actual transpiration [mm month^{-1}]
$S_{u,max}$ = Field capacity [mm]
T_p = Potential transpiration [mm]
S_u = Storage in unsaturated zone [mm]

The field capacity ($S_{u,max}$) determines the amount of water that will percolate into ground water and which part stays in the unsaturated zone. The amount of water stored in the soil in the unsaturated zone is referred to as soil moisture (S_u). If there is enough moisture in the soil then actual transpiration can reach potential transpiration.

The figure below illustrates the relation between soil moisture and actual transpiration.

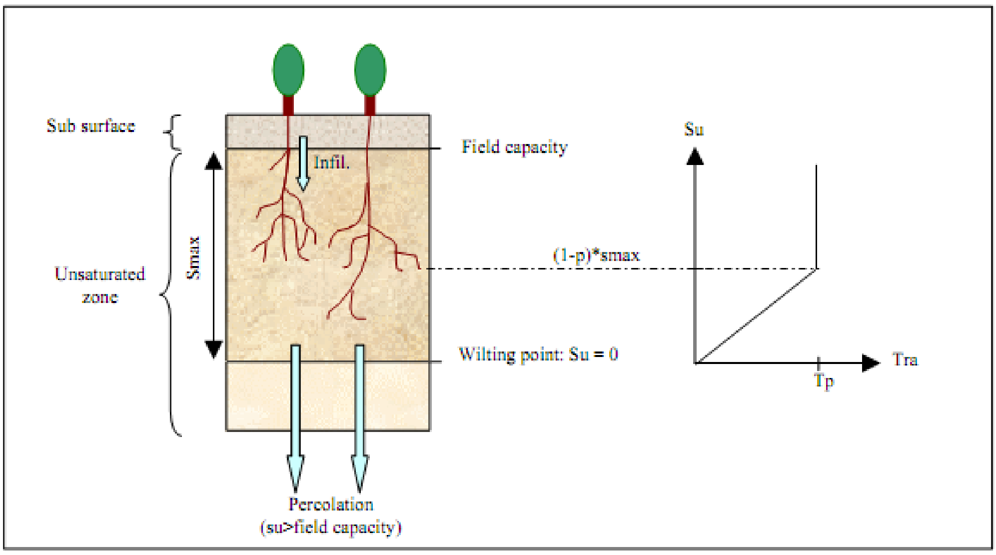

Figure 3.2 Relation between soil moisture and actual transpiration

(Gerrits, 2005)

Calculation with the above formula assumes that potential evaporation accurately accounts for interception. In this model potential evaporation is calculated using the Hargreaves method which is solely based on temperature. There are three possibilities of relating potential transpiration to potential evaporation (Seyam, 2002 as quoted by Gerrits, 2005)

These are:

1. Potential evaporation *(E_p)* does not account for all the interception *(Int).* This is because potential evaporation is based on air temperatures whereas interception is driven by factors such as wind and air temperature. A scenario where potential transpiration can equal potential evaporation is in dense forests.
2. Potential evaporation does account for interception. This happens where interception is mainly evaporated from bare soil. The potential transpiration is then equal to potential evaporation less the interception.
3. Potential evaporation does partly account for interception. This scenario occurs between option 1 and 2 and is dependent on land cover.

The figure below illustrates the three possibilities for calculation of potential transpiration.

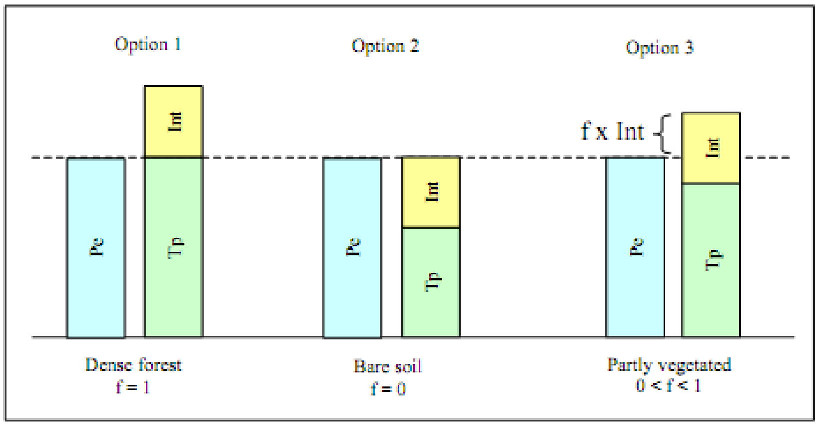

Figure 3.3 Possibilities of potential evaporation (Gerrits, 2005)

These possibilities are incorporated into the model by use of a treetop factor *(f)*. If the basin is covered in forest the *f* factor is set as one (1) assuming interception is not included in potential evaporation. For bare soil the factor is set at zero (0) assuming interception is included in potential evaporation.

In this model this is calculated as follows:

$$T_p = Ep - (1-f) * Int \ldots\ldots\ldots\ldots\ldots\ldots \quad \boxed{5}$$

Where:

T_p = Monthly potential transpiration [mm] month^{-1}

Ep = Monthly potential evaporation [mm] month^{-1} *

f = treetop factor (between 0 and 1) based on land cover map.

Int = Monthly interception [mm] month^{-1}

* (*Ep* is also referred to *as Pe* in the schematic)

3.1.4 Ground water flow

The ground water is recharged with the component after separation using the separation coefficient *(Cr)* herein referred to as the fast component *($Cr * P_n$).* This water flows through macro pores and cracks and the slow flow component is S_{excess}. S_{excess} is the water that percolates from the unsaturated zone if the field capacity (S_{umax}) is exceeded. It is routed to ground water using a time scale K_u. The saturated zone water depletion is separated into three components: The saturation overland flow *(saof)* ,the quick flow(*qflo*) and the slow flow (*sflo*).

The *saof* is the amount of water that flows over the surface because the soil is saturated otherwise referred to as Horton type overland flow(HOF). At this point it is considered that the ground water table has reached the surface due to saturated soil. This is determined by the Digital elevation model (DEM) and the bottom of the draining river (GWS_0).The distance between the surface and bottom is called GWS_{dem} (See figure 3.4).

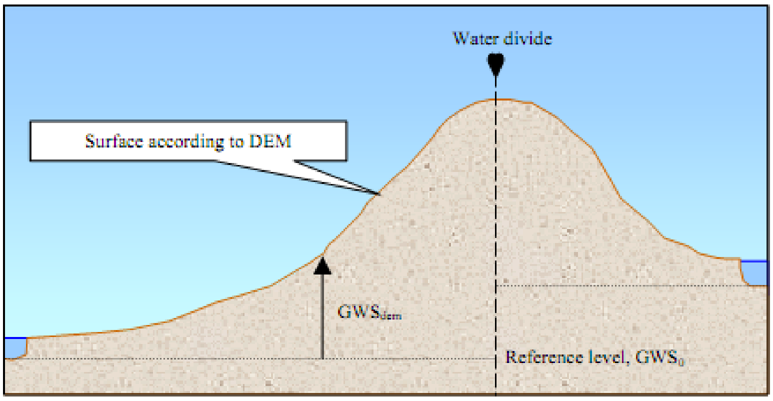

Figure 3.4 Schematic of GWS_{dem}

(Gerrits, 2005)

According to Gerrits (2005), the space GWS_{dem} is not completely available for water, due to the porosity and compression of the soil. To give the relationship between GWS_{dem} and GWS_{max} the equation as illustrated in Fig 3.5 below is used. When ground water table exceeds the GWS_{max} then the water will directly drain into the river.

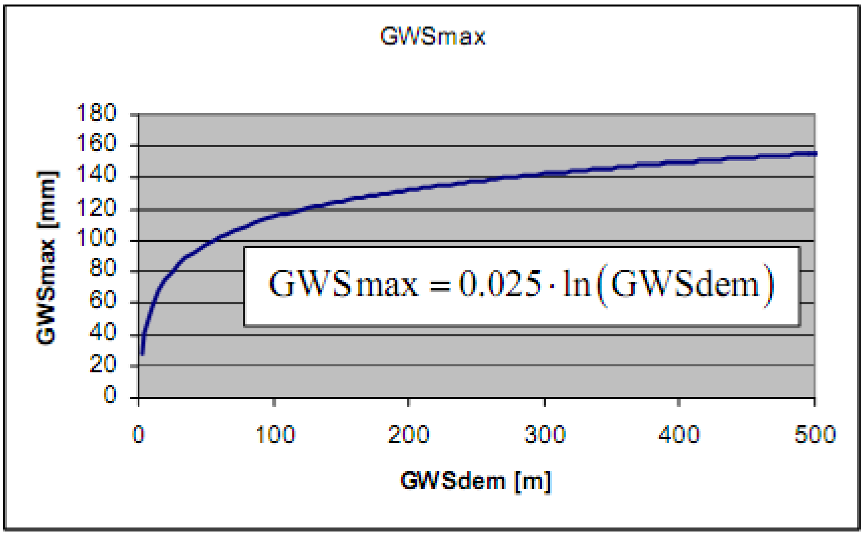

Figure 3.5 Relation between GWS_{dem} and GWS_{max}.

(Gerrits, 2005)

The saturated zone consists of two linear reservoirs, which are separated by a threshold S_{sq} and bounded by S_{smax}. Slow and quick recession coefficients generate runoff from the reservoir. S_{sq} represents overtopping of the saturated zone representing rapid subsurface flow.

The second ground water flow component quick ground water flow through macro pores and cracks (*qflo*) is assumed linearly depends on the ground water level minus threshold level (GWS_{quick}).this level is calculated by multiplying the GWS_{max} with a calibration factor *qc*. The flow is calculated using the equation below, *rtq* is the recession coefficient for quick flow determined from the recession curves:

$$Qflo = \frac{max(GWS - GWSmax, 0)}{rtq} \quad \dots\dots\dots\dots\dots\dots\dots\dots \quad \boxed{6}$$

The third ground water component is the slow ground flow water *(sflo).* In STREAM model the calculation of ground water is given as:

$$Sflo = \frac{max\,(GWS, -15)}{rts} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \boxed{7}$$

To model the situation where the river gives its water to the surrounding cells during a recession a negative ground water level is used. The slow flow recession constant is *rts*. q_C is a calibration coefficient (quick reacting component) of values between 0 and 1 . K_S and K_q are determined from recession curve analysis (Winsemius et al., 2006).

Usually from ground water some of the water flows back to the unsaturated zone. This process is called capillary rise. This is the process of upward movement of water through narrow pores by cohesive forces. Capillary rise takes place if the ground water table is not too low. Capillary rise causes an increase in transpiration if the capillary zone is in the root zone of the plant.

The saturated zone consists of a dead storage zone below 0, from which capillary rise *C* is possible, this is described as:

$C \quad = \quad C_{min} \qquad Ss \quad \leq \quad S_{s,min}$ …………………………………………………… $\boxed{8}$

or

$C \quad = \quad C_{max} \qquad Ss \quad > \quad S_{s,min}$ …………………………………………………… $\boxed{9}$

Where C_{min} [mm month $^{-1}$] is the minimal capillary rise equal to 2mm month^{-1} and C_{max} [mm month^{-1}] is a calibration parameter, larger than C_{min}.

3.1.5 STREAM variables summary

The following variables constitute the STREAM model:

I = Interception [mm month $^{-1}$]
P = Precipitation [mm month $^{-1}$]
E_p = Potential evaporation [mm month $^{-1}$]
P_n = Net precipitation [mm month $^{-1}$]
S_{umax} = Field Capacity [mm]
S_{excess} = Excess of moisture over S_{umax}[mm]
T_a = Actual transpiration [mm month^{-1}]
T_p = Potential transpiration [mm]
S_u = Storage in unsaturated Zone [mm]

3.1.6 STREAM parameters

Table 3.1 STREAM parameters

Parameter	Determination method
$S_{s,max}$ - Saturated zone storage [mm]	Relation with land use Map
q_C - Quick reacting component [-]	Calibration
K_q - quick flow recession constant [-]	Recession curve Analysis
K_s - slow flow recession constant [-]	Recession curve Analysis
D - Interception threshold [mm]	Calibration
C_r - Unsaturated/saturated zone separation coefficient [-]	Calibration
C_{max} - Capillary rise [mm]	Calibration
$S_{s,min}$ - Saturated zone storage	Calibration
K_u - Routing time scale [months]	Calibration
S_{sq} - Separation threshold [-]	Calibration

Source: Winsemius et al., 2006

3.2 Data Collection and field observations

The general data used for the modeling effort are classified in the table 3.2 together with the associated sources.

Table 3.2 Input Data and Sources

Data Category	Description of Data	Data Source
River FLOW data	Time series flow data	Water Resources Management Authority. Lake Victoria South Water Board, Kisumu
Meteorological Data	Climate Change Data	GCM Data/IPCC fourth assessment report
	Rainfall Data	Kenya Meteorological Department/Lake Victoria South Water resources management authority Climate Research Unit monthly grids
	Evaporation data	
	Wind speed	
	Temperature	
	Humidity	
	Solar radiation	
Basin Maps	Land Use /Land cover Map	USGS, Government Reports on land use
	DEM	Shuttle Radar Topographic Mission
	Soil Map	Kenya Soil Survey/Soil Terrain Database of East Africa
Observations	Soil profiles Vegetation Cover Surface flow patterns Land Formations Ground water levels Land Changes i.e afforestation	Mara basin field observations

3.3 Hydrological model Spatial input data

The spatial data input necessary for running the STREAM model consist of five input layers. These are:

- A land use map
- A soils map
- Monthly precipitation map
- Monthly temperature map
- A digital elevation Model(DEM)

3.3.1 The digital elevation model.

The digital elevation model is obtained from the Shuttle Radar Topographic Mission (SRTM). This is a 90 m DEM (Fig 3.6). The DEM was processed using ARC GIS in terms of delineating the upper Mara catchment. The catchment boundary adopted is similar to that which has been used in previous modeling efforts in the Mara. This was deliberately done so as to provide a similar platform for comparison purposes. The DEM was imported into the PCRaster Modeling environment and all the sinks filled and a local drain direction map generated.

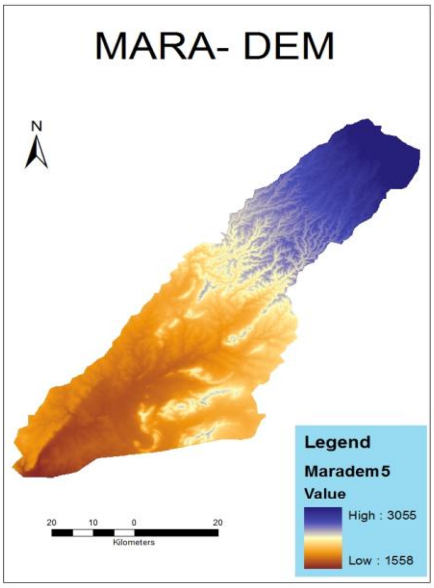

Figure 3.6 Mara DEM

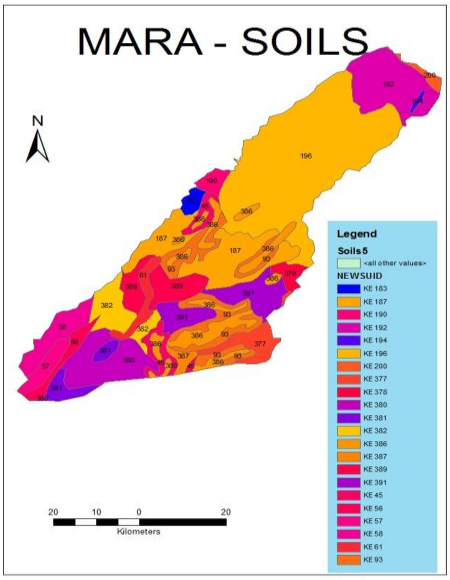

Figure 3.7 Mara Soils

The Upper Mara landscape elevation varies from a low of 1558 m a.s.l to a high of 3055 m a.s.l. (See Fig 3.6).

3.3.2 Mara Soils

The soils data used entailed soil classifications based on the FAO 1988 system as well as the local Kenyan classification with a prefix of KE. The soil data used are extracted from soils data derived from SORTER studies of carbon stocks and change in Kenya (Fig 3.7). Table 3.3 below gives parameter estimates for the soils. The soils in Mara are basically classified into two dominant classes known by their common names as Clays and Sandy loams.

Table 3.3 Mara soil parameter estimates

			MASS %	MASS %	MASS %	Available water(%W/V)	
		FAO	SDTO	STPC	CLPC	TAWC	
No	CODE	Classification	SAND	SILT	CLAY		
1	KE45	PHI	9	24	67	11	Clay
2	KE56	RGd	60	19	21	15	Sandy clay Loam
3	KE57	Cmu	35	20	45	19	Clay
4	KE58	GRh	32	28	40	-1	Clay
5	KE61	Vre	20	30	50	12	Clay
6	KE93	PHh	15	70	15	12	Silt Loam
7	KE183	NTu	30	26	44	0.89	Clay
8	KE187	LVv	38	35	27	15	Loam
9	KE 190	NTu	30	26	44	0.89	Clay
10	KE192	ANm	20	48	32	24	Clay loam
11	KE194	Plu	22	50	28	18	Clay loam
12	KE196	ANm	20	48	32	24	Clay loam
13	KE200	ANm	20	48	32	24	Clay loam
14	KE 377	Ple	58	26	16	18	Sandy loam
15	KE378	PHh	15	70	15	12	Silt Loam
16	KE 380	VRe	58	26	16	18	Sandy loam
17	KE 381	PLe	58	26	16	18	Sandy loam
18	KE382	GRh	32	28	40	-1	Clay
19	KE386	CMu	35	20	45	19	Clay
20	KE 387	PLe	58	26	16	18	Sandy loam
21	KE 389	PLe	58	26	16	18	Sandy loam
22	KE 391	PLe	58	26	16	18	Sandy loam

3.3.3 Mara land use

The land use as described earlier is predominantly agricultural with significant areas under forest cover (See Fig 3.8). This land use information is based on land use data of the year 2000. Obtained from satellite Imagery data of the Landsat thematic mapper.

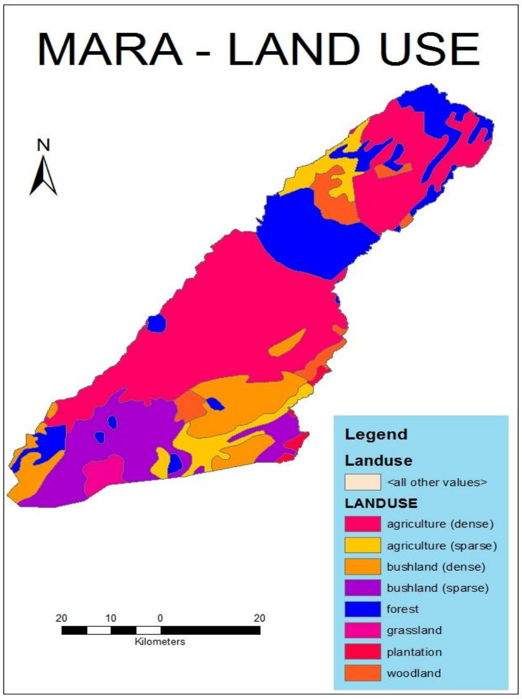

Figure 3.8 Mara land use

3.3.4 Mara Hydro met stations

Four (4) stations were chosen for rainfall data. This choice was based on the completeness of the data as well as the period required to match previous modeling efforts in the Mara. Three of these stations are within the basin whereas Narok station is located outside the basin. Narok station is located approximately 35 km away from the basin (See Fig 3.10).

The four stations with their associated location data are as enumerated in Table 3.4 for rain gauges and table 3.5 for hydromet stations. Most of the precipitation data stations consist of a simple rain gauge as illustrated in Fig 3.9.

Table 3.4 Location of rain gauge stations

No	STATION NAME	LATITUDE	LONGITUDE
1	Nyangores	35.433	-0.700
2	Bomet water supply	35.350	-0.783
3	Kiptunga forest station	35.800	-0.450
4	Narok	35.833	-1.330

Table 3.5 Location of Metrological stations

No	STATION NAME	LATITUDE	LONGITUDE
1	Narok	35.833	-1.330
2	Kisii	34.783	-0.667
3	Kericho	35.267	-0.367

Figure 3.9 Bomet rain gauge station

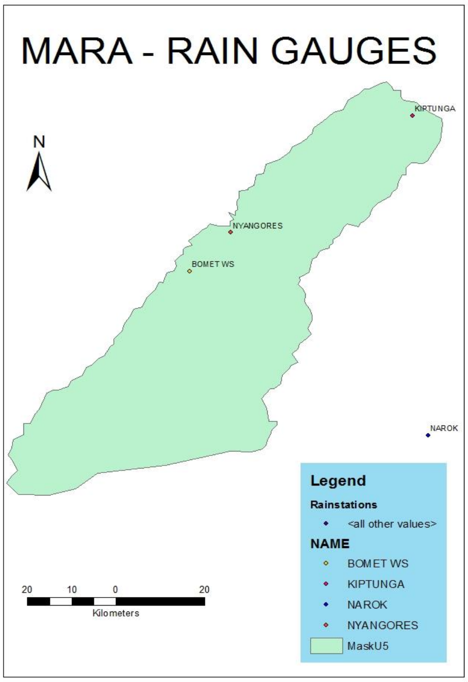

Figure 3.10 Rain gauges Location

Figure 3.11 Meteorological stations location

3.3.5 Mara flow gauging and abstraction points

Two discharge stations were selected to be used for this research as they represent discharges from the upper Mara and in addition are the only stations within the Upper Mara with data that spans the research period. These two Stations are Nyangores and Amala gauging stations. They are staff gauge stations which are located on bridges crossing the two tributaries of the Mara river (See Fig 3.14 and Fig 3.15). Because of the rocky foundations they are quite stable. Their location data is enumerated in Table 3.6 and illustrated in Fig 3.12.

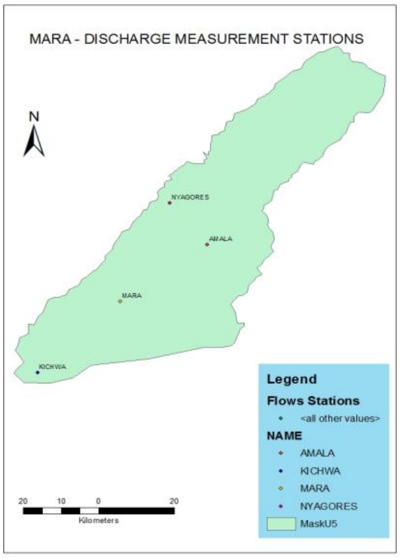

Figure 3.12 Discharge measurement stations
(Source: WRMA , Kisumu)

Table 3.6 Discharge measurement stations locations

NO	STATION	STATION CODE	LATITUDE	LONGITUDE
1	Amala	1LB02	35.437	-0.899
2	Nyangores	1LA03	35.346	-0.789
3	Mara	1LA05	35.019	-1.546
4	Kichwa Tembo	1LA04	35.03611	-01.233

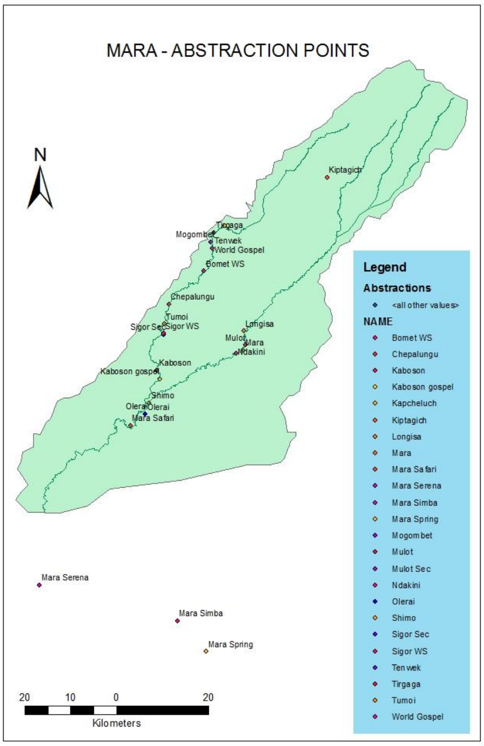

Figure 3.13 Abstraction points

Source (WRMA, Kisumu)

Fig 3.13 above illustrates the location of the major abstraction points in the basin. This data relating to their spatial locations was obtained from the Water Resources Management Authority Offices in Kisumu. These are the licensed abstractors. As can be noticed most of the abstractions are located downstream of the gauging stations of Amala and Nyangores.

Figure 3.14 Nyangores gauging station

Figure 3.15 Amala gauging station

3.3.6 Input data processing

The input data collected were processed to aggregate daily values to monthly values, check for consistency as well as fill missing data gaps for precipitation, discharge and temperature datasets. The main approach was use of correlations between station data. Multiple linear regressions were used to develop relationship equations which were then used to fill the missing data gaps.

3.3.6.1 Rainfall data.

The rainfall data used for this modelling effort were selected for the period 1999 to 2007, for the stations described herein.

The rainfall data were filled at daily level, but with a threshold of 15days missing data maximum. If more than 15 days continuous missing data it was considered that the station did not have data for that month and the missing month was subsequently filled using multiple linear regressions at aggregated monthly data (See fig 3.18). Below is a summary of this data filling exercise, the R^2 values stated here are those used to fill the gaps using multiple linear regressions:

Table 3.7 Summary of gap filling rainfall data

NO	STATION	GAPS DAILY DATA NO	R^2 DAILY DATA	GAPS MONTHLY DATA NO	R^2 MONTHLY DATA	REMARKS
1	Narok	none	-	none	-	-
2	Nyangores	172	0.33	4	0.85	5.2 %
3	Bomet	75	0.31	2	0.87	2.3%
4	Kiptunga	215	0.18	6	0.65	6.5%

Double mass curves were plotted to check the consistency of the data (See fig 3.16 and 3.17), and to identify whether there have been any shifts in the trends of the data.

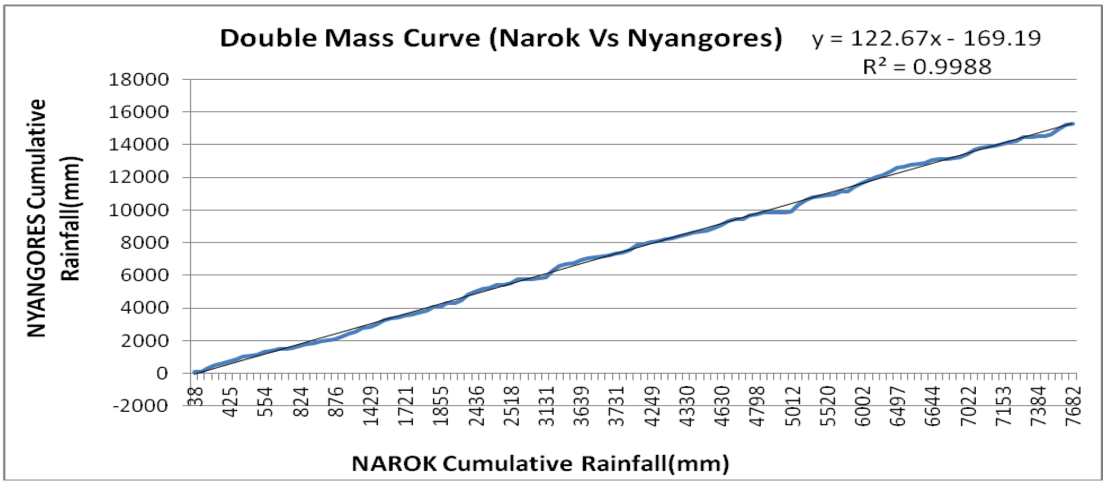

Figure 3.16 Double mass curve of Narok against Nyangores

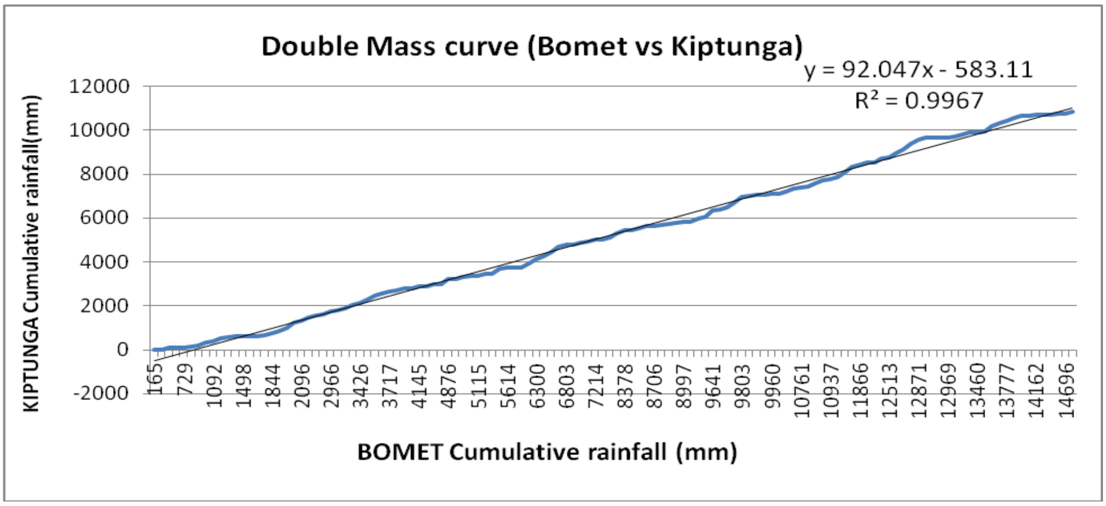

Figure 3.17 Double mass curve Bomet against Kiptunga

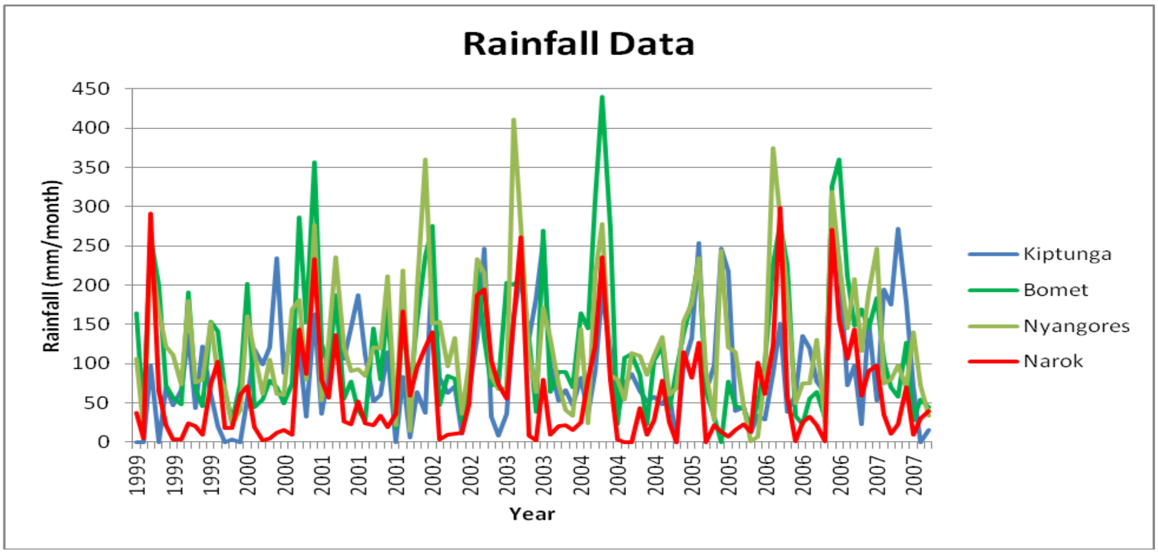

Figure 3.18 Monthly rainfall data for the four rain gauge stations

3.3.6.2 Temperature data

The temperature data used were from the three hydromet stations namely Narok, Kericho and Kisii. These stations have data with several gaps. Temperature was the only parameter with some level of consistency in terms of availability of data. The temperature data from the three stations were first checked for any abnormal readings (outliers), the daily data were put through a gap filling process using multiple linear regressions. The daily data were aggregated to monthly level (See Figures 3.19, 3.20 and 3.21). The data consisted of two sets, a daily maximum temperature and a daily minimum temperature.

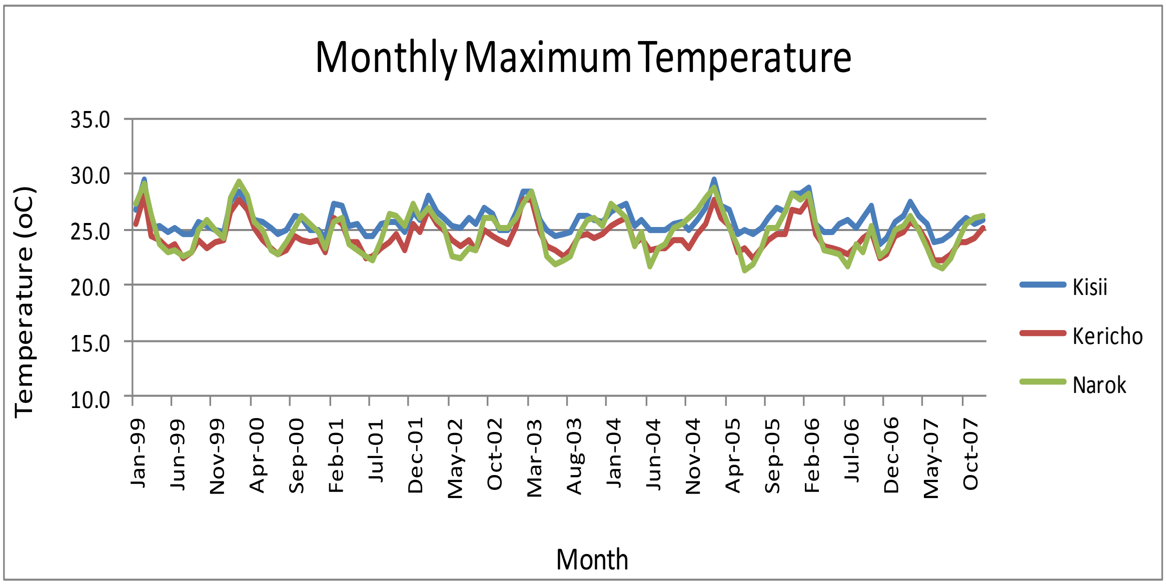

Figure 3.19 Monthly mean maximum temperature.

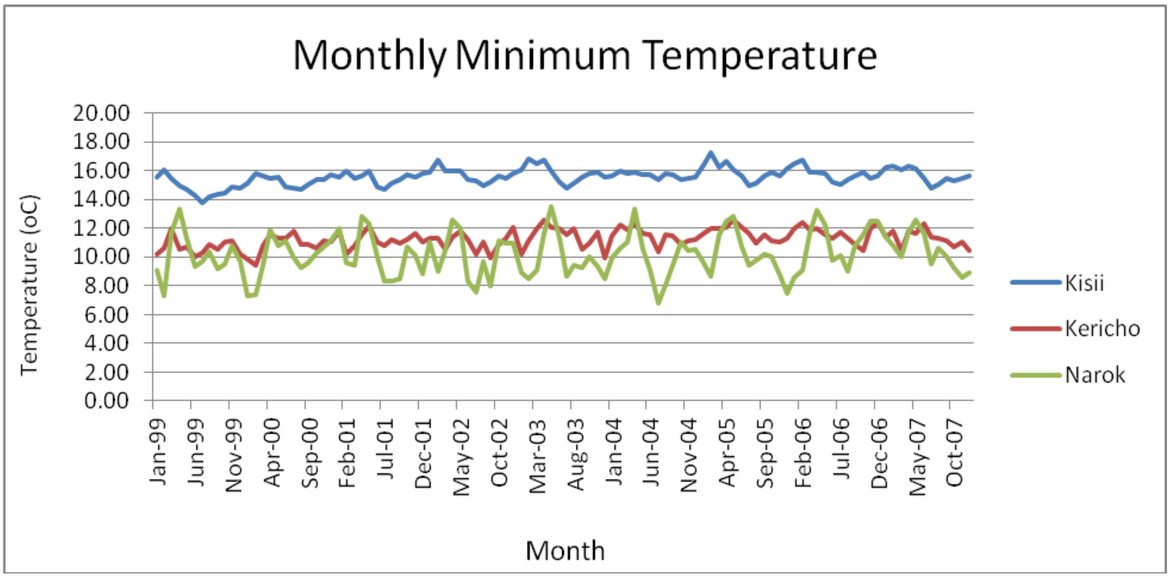

Figure 3.20 Monthly mean minimum temperature.

The coefficients of correlation for the data gap filling were as enumerated in Table 3.8

Table 3.8 Coefficient of correlation ($R^{2)}$ values for multiple linear regressions

NO	STATION	Gaps Max Temperature DATA NO	R^2 MAX Temp DATA	Gaps Minimum Temperature DATA NO	R^2 MIN Temp DATA	REMARKS
1	Narok	5	0.77	3	0.28	4.6%
2	Kisii	13	0.89	12	0.17	12%
3	Kericho	5	0.93	none	-	4.6%

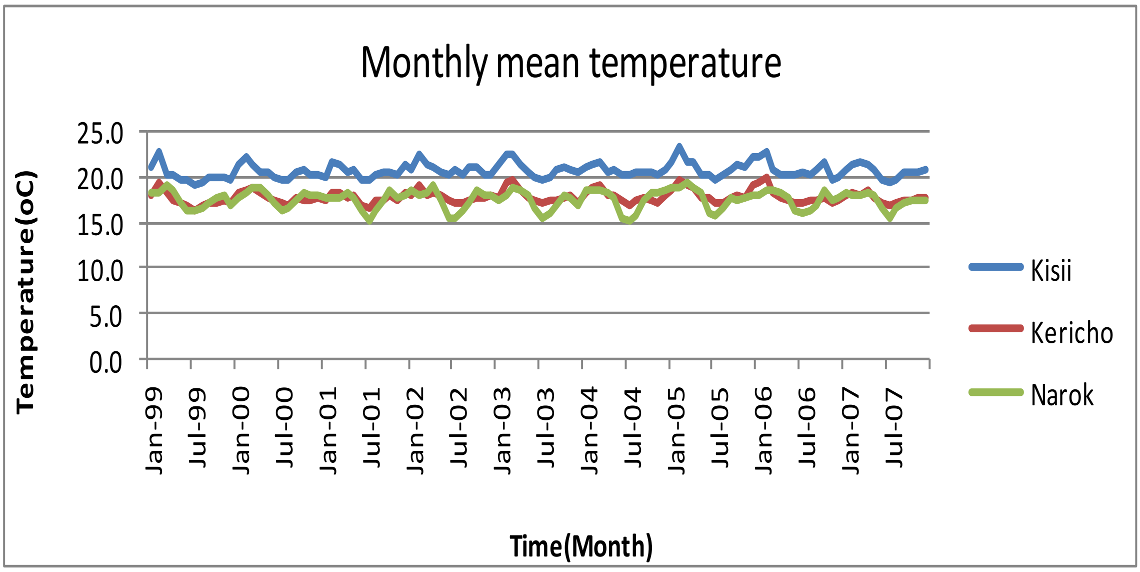

Figure 3.21 Monthly mean temperature

3.3.6.3 Calculation of Potential Evapotranspiration (PET)

The temperature data were used to calculate the potential evapotranspiration using the Hargreaves method (Hargreaves, 2003).

The Hargreaves method uses the following formula

$$ET = 0.0023(T_{mean} + 17.8)(T_{max} - T_{min})^{0.5} * Ra$$

Where:

ET = Evaporation in mm/day
T_{mean} = mean air temperature °C
T_{min} = Minimum air temperature in °C
T_{max} = Maximum air temperature in °C
Ra = Extraterrestrial radiation in mm/day.

The results obtained are as illustrated in Fig 3.22 and Fig 3.23.

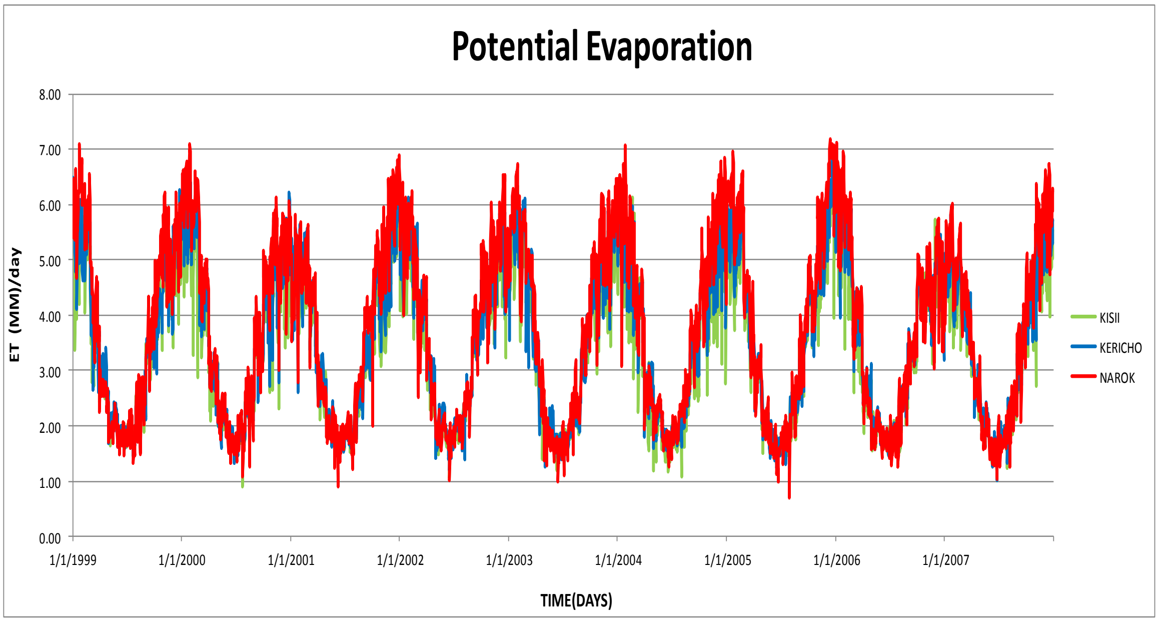

Figure 3.22 Daily potential evapotranspiration

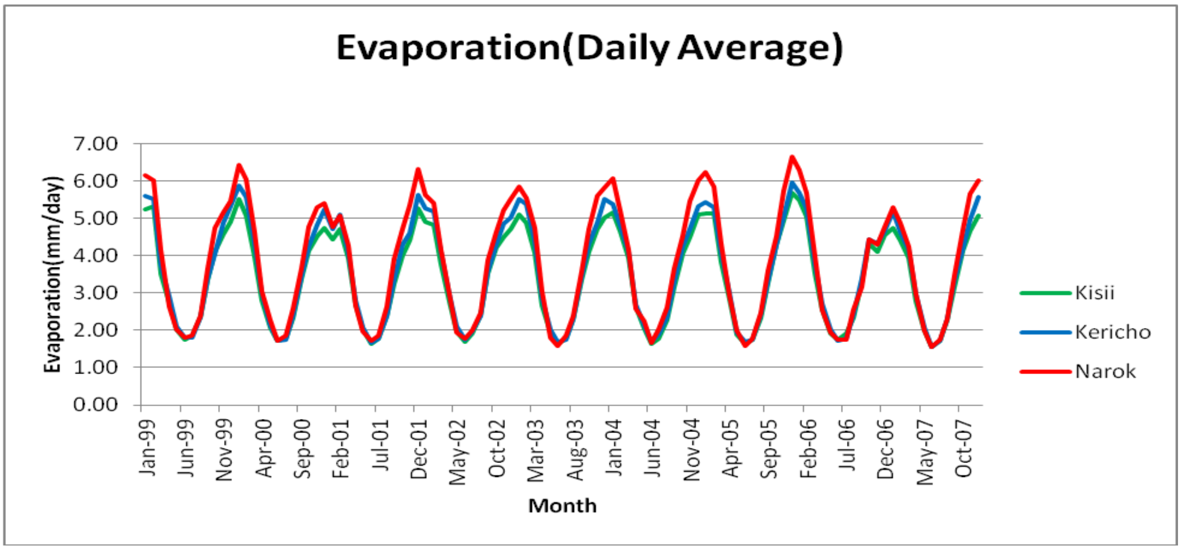

Figure 3.23 Mean Monthly potential evaporation data

3.3.6.3 River Discharge Data

The discharge data were calculated from gauge height readings using the respective rating curves for Amala and Nyangores (See Annex 1).

The location data for the two stations is indicated in table 3.6. The raw data were fed in an analysis program, Indicators of Hydraulic Alteration (IHA) for computation of discharge statistics and filling of gaps. The program fills discharge data gaps through a simple linear interpolation process. The results of this process are presented in Fig 3.24 and 3.25. Figure 3.26 shows a comparison of discharge data against the incident precipitation.

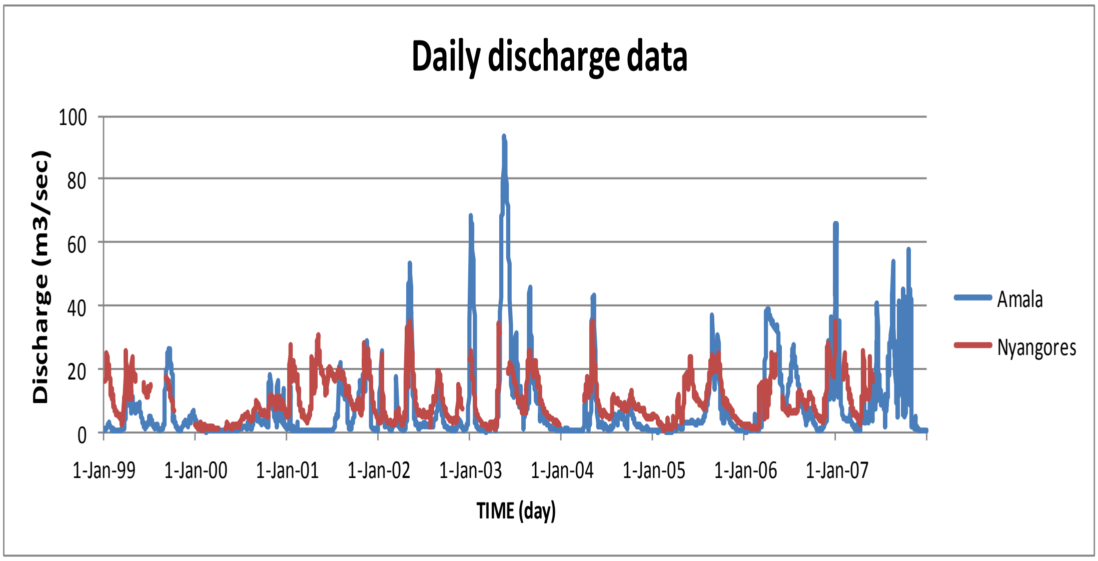

Figure 3.24 Daily discharge data Amala and Nyangores gauging stations

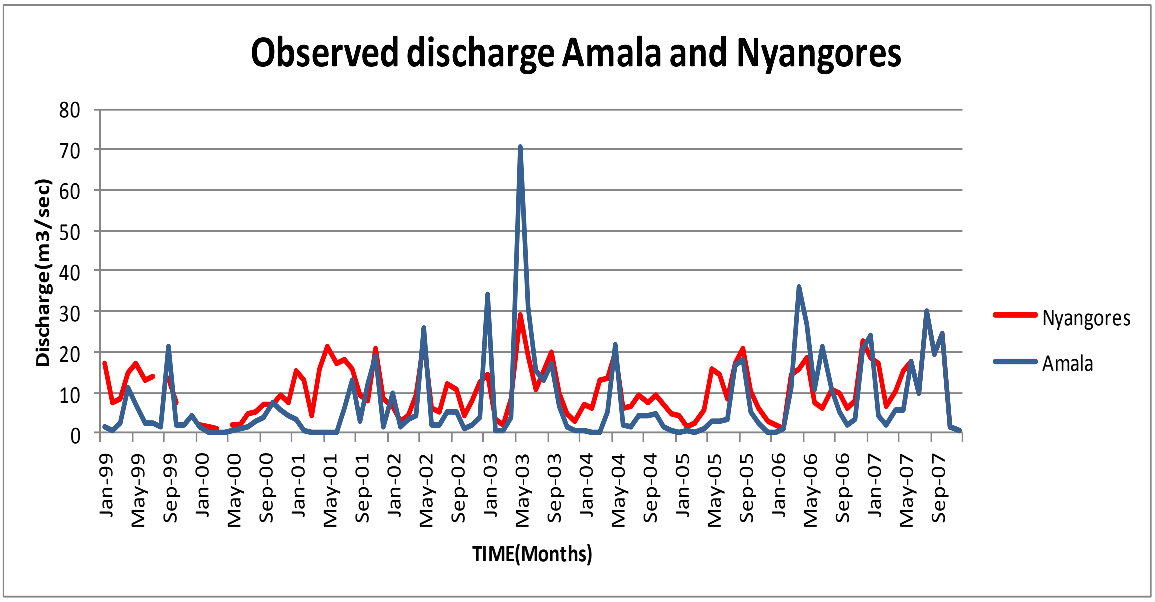

Figure 3.25 Monthly discharge data Amala and Nyangores gauging stations

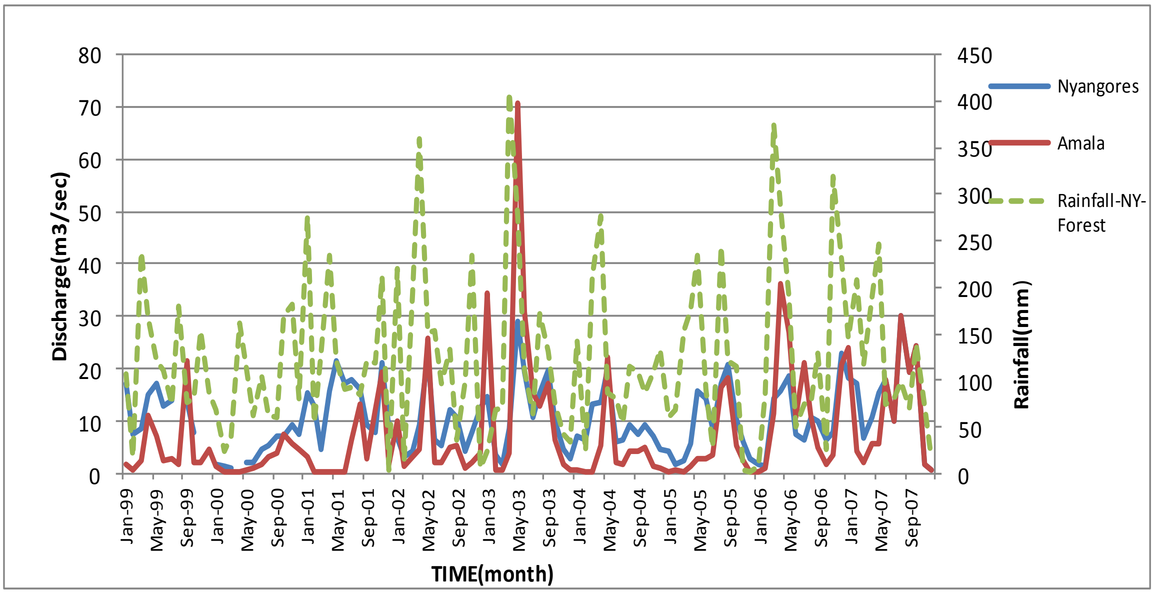

Figure 3.26 Comparison of discharge with rainfall at Nyangores forest station

3.4 Modelling procedure

Modelling in PCRaster involves several steps that can be categorised into four broad categories (See Fig 3.27 for schematic), preparation of data for the model run, running the model, calibration and validation.
Data preparation involved the following:

- Creation of time series data, this involved preparation of time series files for rainfall, potential evaporation and discharge data.

- Preparation of the digital elevation model. Terrain processing for the DEM involved extracting the Mara basin from the DEM tiles and confirming the boundaries. In this case the well established boundaries for the Mara basin were used to mask out the basin.
- The second step involved DEM reconditioning which involved filling in the sinks and burning in the existing stream network. The third step involved creation of a local drain direction (LDD) map. The PCRaster program uses the LDD to accumulate flows over the catchment area.
- Land use and soil data maps were processed in Arc View and subsequently transformed into a PCRaster acceptable format.
- Using PCRaster commands input data were regionalised to pixel level. In this modelling effort a 90m x 90 m pixel size is used, based on the DEM

The second category of activities involved the actual writing of the modelling code in PCRaster language. The modelling philosophy of using buckets with reservoirs, thresholds and recession constants was used here. Subsequently the input data files are read into the program during the model run phase of the procedure.

The third category of activities involved calibrating the model. The model was calibrated manually. The discharge data for Nyangores and Amala stations were used to calibrate the model. The model calibration was carried out for the period 1999 to 2003 and validated for the period 2004 to 2007. An evaluation of the model performance was carried out. Subsequently after successful calibration the model was validated. However during computation of model performance long periods of missing data were excluded.

The last category of activities involved applying the developed land use change scenarios and evaluating the impact on the Mara hydrograph. The land use maps for the various scenarios were used as input to run the model for each scenario and the results recorded.

The resultant simulated discharge data was fed into Indicators of Hydraulic Alteration (IHA) software for analysis and computation of hydrograph statistics. This is software designed to analyse daily river flow data for indications of alteration in the hydrograph. For this research the analysis parameters adopted are the 7, 30 and 90 day moving averages, and accompanying parametric statics such as means and coefficient of variance. A moving average is commonly used with time series data to smooth out short-term fluctuations and highlight longer-term trends or cycles.

For this research the values for 7 day, 30 day and 90 minimum and maximum flows are calculated. In addition, a base flow index is computed. This is worked out as the 7 day minimum divided by the mean annual flow of the river. The seven day minimum is regarded as an indicator of the river's base flow and the base flow index indicates the proportion of the river flow that is contributed to by base flow.

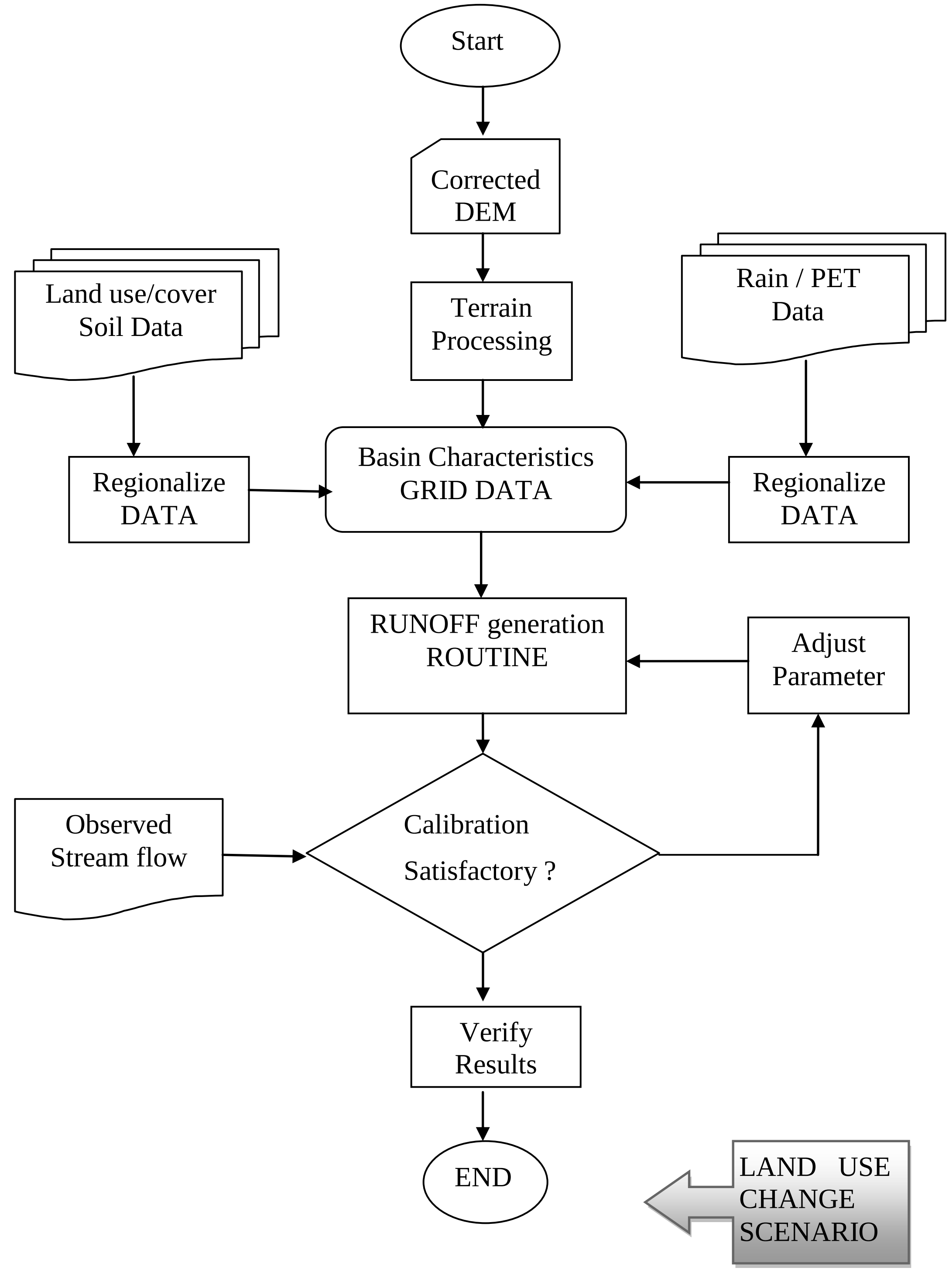

Figure 3.27 Modeling procedure for the MARA basin in PCRaster

3.5 Land use change Scenarios

The land use change scenarios were adopted from a previous modelling effort in the Mara (Mango et al., 2011).In addition to this they are also based on trends in the basin both past and current. Table 2.1 indicates the historical land use changes in the basin. These have informed the formulation of scenarios coupled with more recent information such as afforestation. Three land use change scenarios were adopted, these are mainly derived from trends on land use in the basin over the last 40 years. The upper Mara comprises of a total of 3,462 km^2 which is distributed in various land use categories as illustrated in Fig 3.28 below.

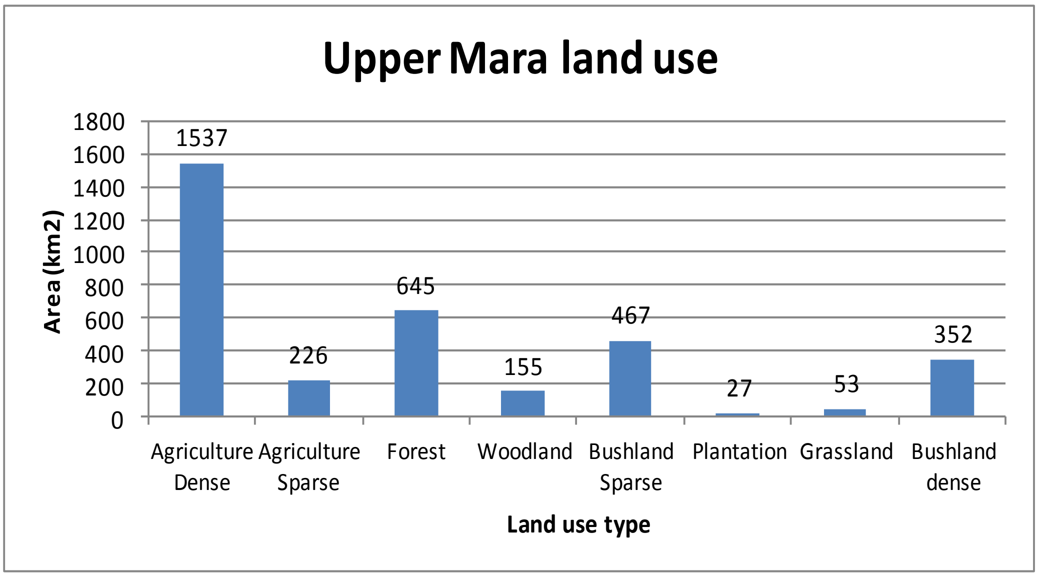

Figure 3.28 Upper Mara land use distribution

On the other hand Amala and Nyangores catchments comprise of, a total of 1312 km^2 distributed as indicated in Fig 3.29.The dominant land use is dense agriculture followed by forest cover.

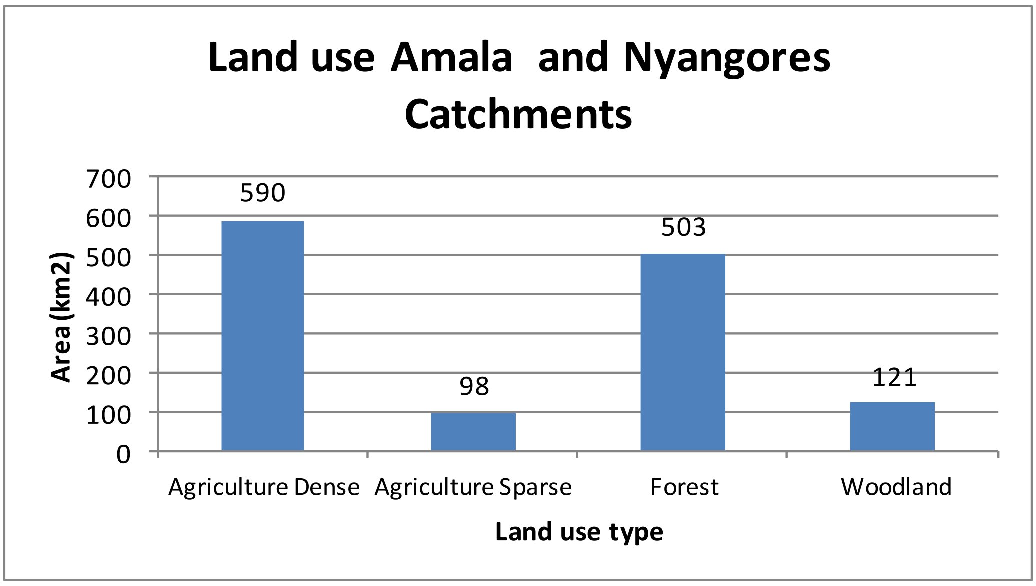

Figure 3.29 Land use distribution in Amala and Nyangores catchments

3.5.1 Partial deforestation scenario

This scenario is indicative of trends in the basin where excision of forest land has been taking place over the last 40 years. This scenario seeks to understand what the impacts of deforestation are and subsequently identify suitable mitigation measures. It envisages a partial loss of up to 25% of the current forest cover. The lost forest is modelled as shrub land. This makes the percentage of the area covered by forest to shift from 38% in the base scenario to 29%, whereas bush land then occupies 10% of the total catchment area of Amala and Nyangores (See Fig 3.30 and 3.31).

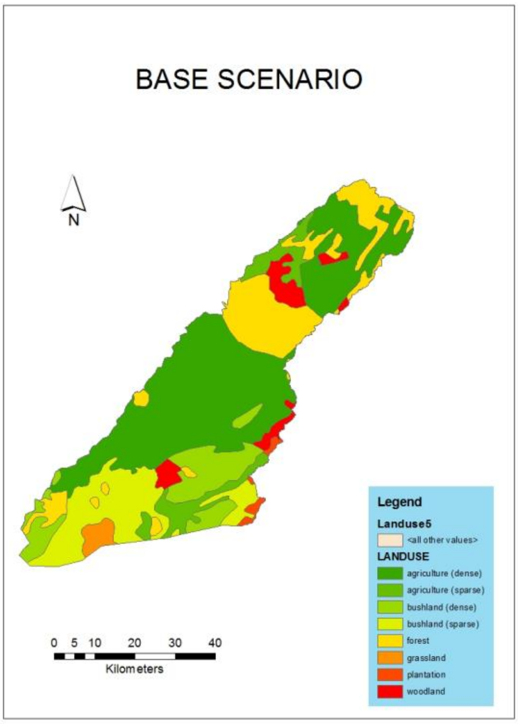

Figure 3.30 Base scenario

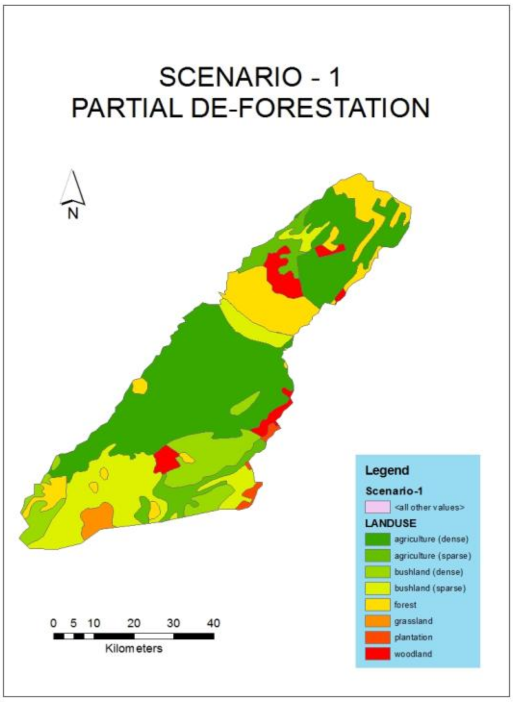

Figure 3.31 Partial deforestation scenario

3.5.2 Replacement of forest with Agriculture.

Most of the forest excisions in the Mara forest are for agricultural purposes, replacement of forest with agriculture models this occurrence. The percentage of agricultural land in the catchment then shifts from 45% to 85% of the total catchment area of Amala and Nyangores, whereas forest land is reduced to zero percent (0%).

3.5.3 Aforestation

In addition to the Scenarios adopted from previous modelling efforts, the afforestation scenario is developed. This is informed by the current ongoing reforestation efforts in the basin. In the basin evictions are taking place in areas settled by people but which were previously gazetted forest land. Subsequently afforestation is taking place in these repossessed lands. This scenario envisages an increase of 26 % of the current forest cover. This makes the total percentage of forest cover in the catchment to increase from 45% to 48%. Figure 3.31, 3.32 and 3.33 illustrate these land use change scenarios. Table 3.9 gives the land use statistics for all the scenarios with regard to the area of every land use and its percentage coverage of the total catchment of Amala and Nyangores.

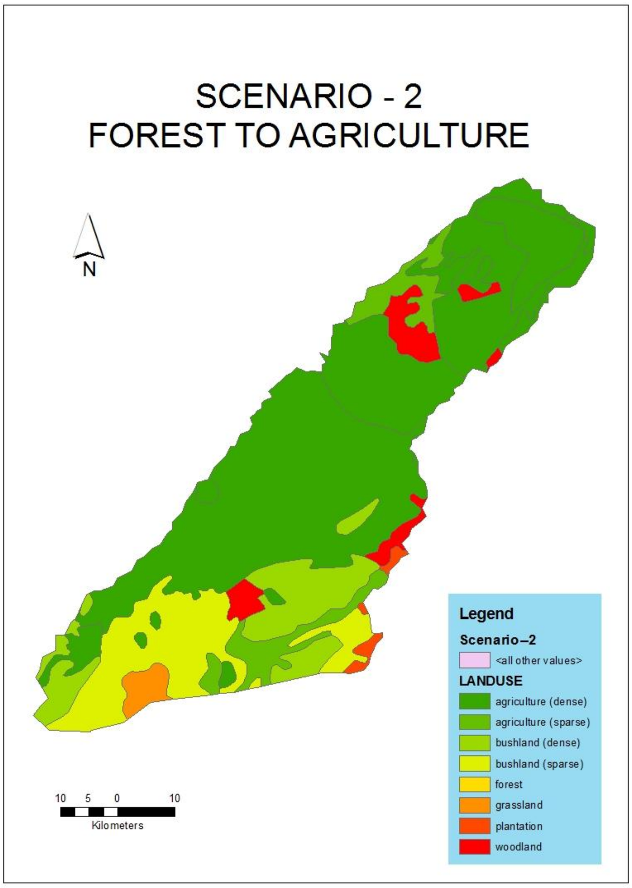

Figure 3.32 Conversion of forest to Agriculture scenario

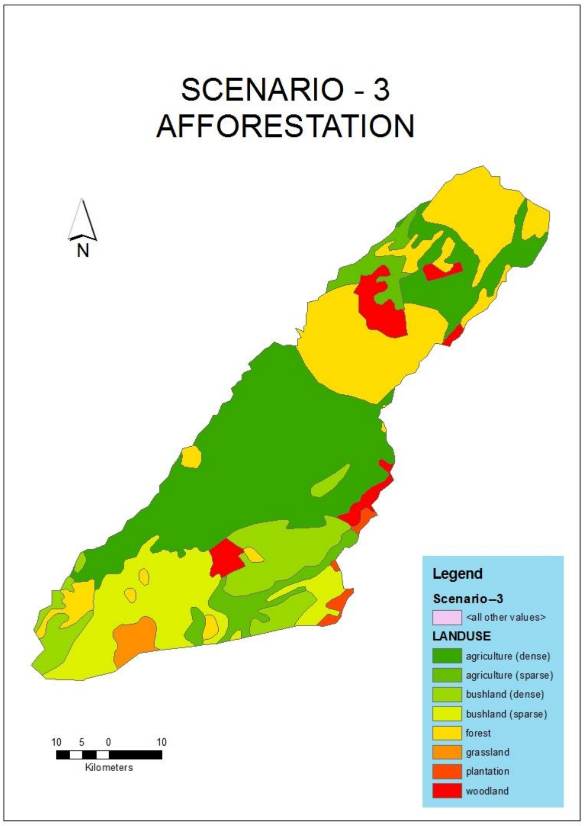

Figure 3.33 Afforestation scenario

Table 3.9 Land use change scenario statistics

NO	**LAND USE CATEGORY**	**UPPER MARA (Km^2)**	**AMALA AND NYANGORES CATCHMENT Area (Km^2)**			
			Base Scenario	**Scenario 1 Deforestation**	**Scenario 2 Conversion**	**Scenario 3 Afforestation**
1	Agriculture dense	1536.86 (44.41%)	590 (44.95%)	590 (44.95%)	1092.91 (83.26%)	472.70 (36.04%)
2	Agriculture Sparse	225.56 (6.52%)	98.20 (7.48%)	98.20 (7.48%)	98.20 (7.48%)	85.20 (6.5%)
3	Forest	644.58 (18.63%)	502.91 (38.31%)	377 (28.72%)	0	633.21 (48.2%)
4	Wood land	155.18 (4.48%)	121.49 (9.26%)	121.49 (9.26%)	121.49 (9.26%)	121.49 (9.26%)
5	Bushland Sparse	467.27 (13.5%)	0	125.90 (9.59%)	0	0
6	Plantation	26.77 (0.77%)	0	0	0	0
7	Grassland	52.73 (1.52%)	0	0	0	0
	Total	**3460.47**	**1312.60**	**1312.60**	**1312.60**	**1312.60**

Figure 3.34 below illustrates the various land use percentages for each scenario as enumerated in Table 3.9.

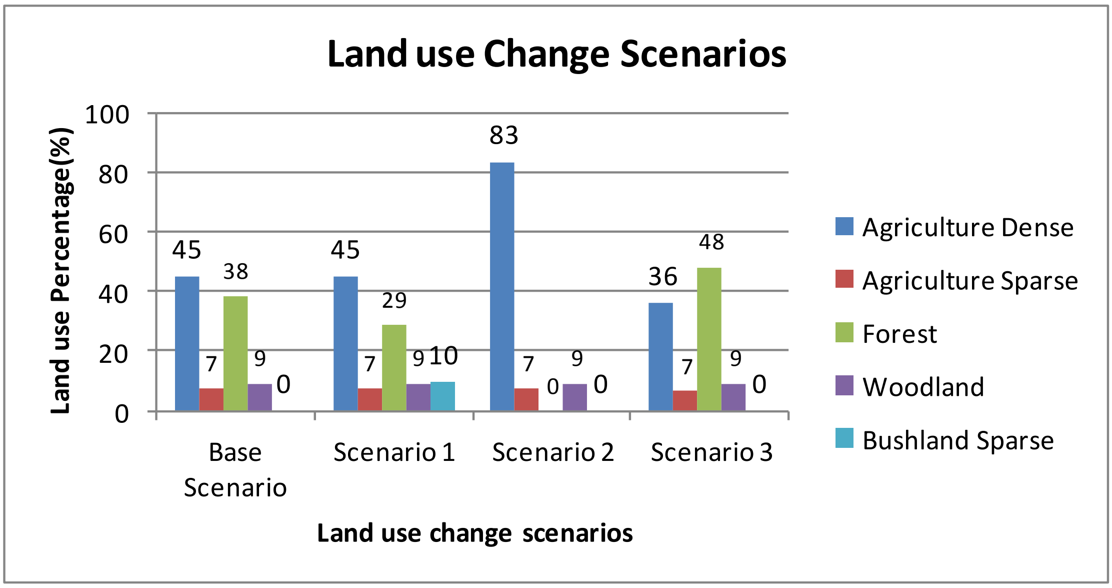

Figure 3.34 Percentage land use under various land use change scenarios.

4.0 RESULTS AND DISCUSSIONS

This chapter presents the results obtained after the model was successfully run and subsequently applied to the various land use change scenarios.

4.1 Model Run

The model was run in dynamic mode to simulate a combined period of nine (9) years translating to a total of 3287 time steps. The time steps were on daily basis and the results were then aggregated to monthly basis.

4.2 Model Calibration

Model calibration involved varying the following parameters (see Table 4.1) to calibrate the model to observed data. The calibration method used was manual calibration. During calibration as one parameter is varied, the rest are held constant till a suitable result is achieved. This was repeated for all the parameters, however not all parameters responded equally as some were more sensitive than others. The most sensitive parameters were those relating to the size of the saturated and unsaturated zone and the quick flow coefficient. A sensitivity analysis was not carried out for this modelling effort.

Table 4.1 Model calibration parameters

Parameter	Calibration Value range
$S_{s,max}$ - Unsaturated zone storage	0 - 330 mm
q_C - Quick reacting component	0 -100
K_q - quick flow coefficient	0 - 1 month
K_s - slow flow coefficient	0 - 26 months
D - Fast evaporation threshold	2 - 4 months
C_r - Unsaturated/saturated zone separation coefficient	0 -100
C_{max} - Capillary rise	0 - 5 mm/day
$S_{s,min}$ - Saturated zone storage	100 - 200 mm

Several parameter sets were used in the model. The unsaturated zone storage $S_{s,max}$ was modelled using parameter values adopted from modelling of the Zambezi basin (Gerrits,2005).The value $S_{s,max}$ was based on land cover. Table 4.2 gives the adopted values (See Annex 2 for calibration values).

Table 4.2 Parameter values for the unsaturated zone

Land cover	S_{smax}(mm)
Forests	270
Closed shrub lands	180
Open shrub lands	130
Woody savannas	280
Savannas	320
Grassland	150
Croplands	250
Cropland and natural vegetation mosaic	330
Barren or sparsely vegetated	110
Water	0

Nyangores Calibration

Nyangores was calibrated for the period July 2001 to November 2003.This is a shortened calibration period to avoid periods of missing data. The efficiency during calibration was an NSE of 0.59 and an R^2 of 0.80 at monthly level. (See Fig 4.1)

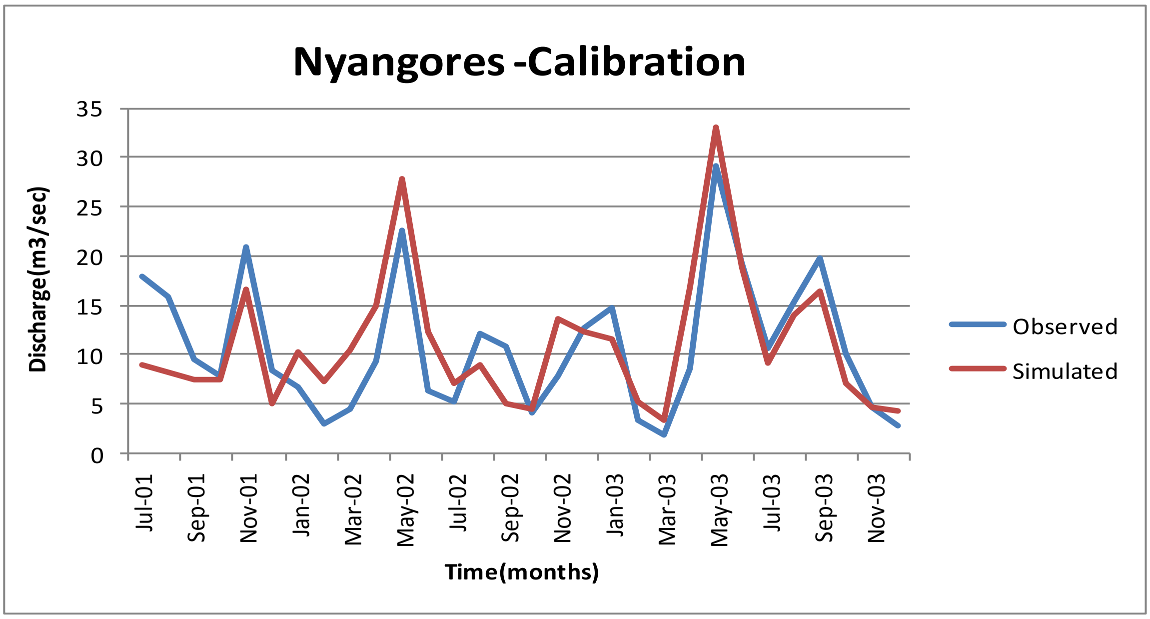

Figure 4.1 Nyangores Calibration

Amala Calibration

Amala was calibrated for the period July 2001 to November 2003. The efficiency during calibration was a NSE of 0.56 and an R^2 of 0.81 at monthly level.(See Fig 4.2)

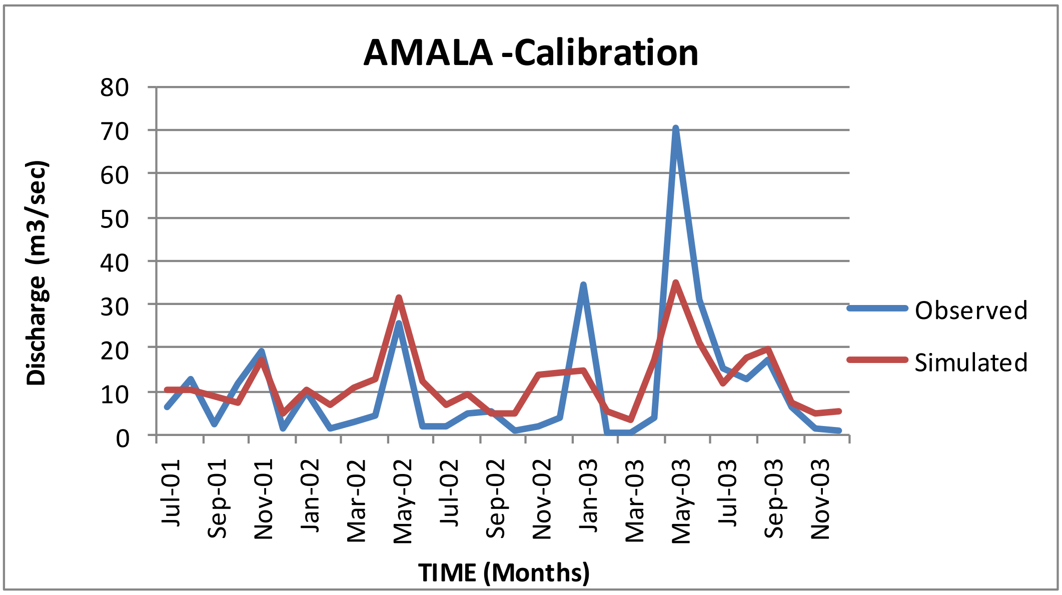

Figure 4.2 Amala calibration

4.3 Model validation

To validate the model, the calibrated model was applied to a period of four (4) years. A total of 1460 time steps, translating to 4 years. The data used for validation was for the period January 2004 to December 2007. Again this was in actual practice applied in shortened periods to avoid periods of filled data.

Nyangores validation

Nyangores was validated for the period July 2004 to July 2007 a period of three years. The validation efficiency obtained was a NSE of 0.52 and an R^2 of 0.79 at monthly level.(See Fig 4.3)

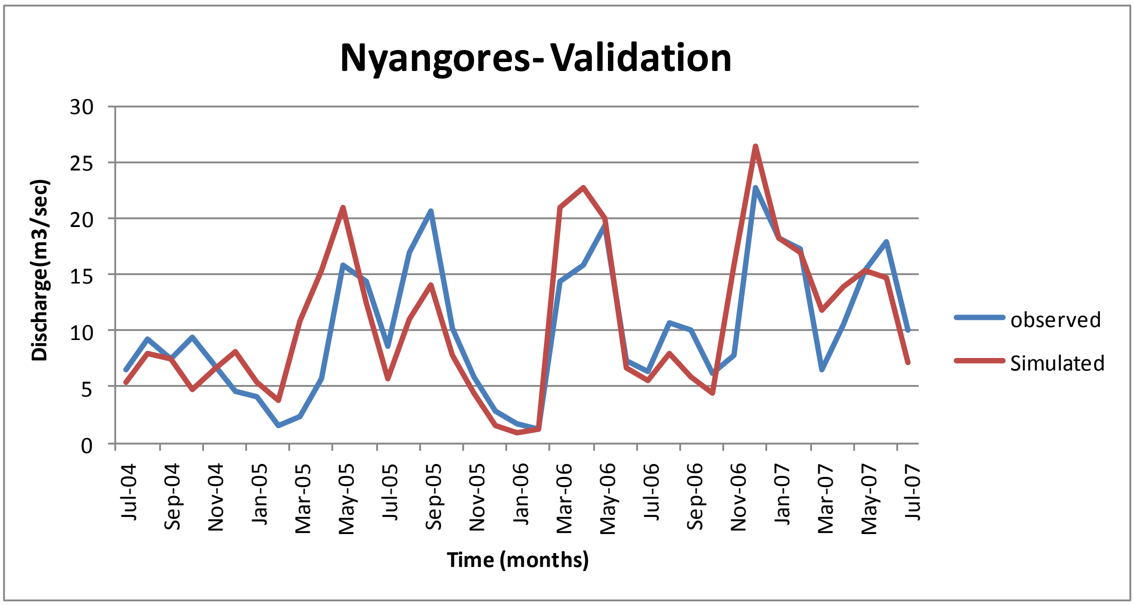

Figure 4.3 Nyangores Validation

Amala validation

Amala was validated for the period July 2005 to July 2007 a period of three years. The validation efficiency obtained was a NSE of 0.35 and an R^2 of 0.66 at monthly level.(See Fig 4.4)

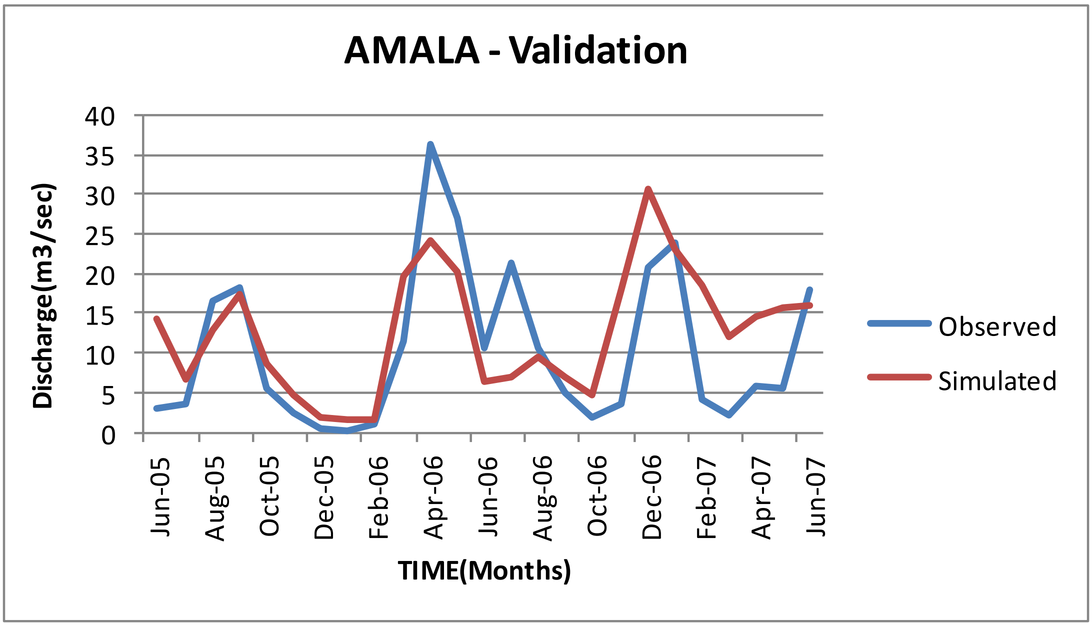

Figure 4.4 Amala Validation

4.4 Model evaluation

The table 4.3 and 4.4 below are a summary of the results of model evaluation.

Table 4.3 Summary model evaluation statistics Amala and Nyangores(Monthly basis)

Statistic	RIVER			
	Amala		Nyangores	
	Calibration	Validation	Calibration	Validation
NSE	0.56	0.35	0.59	0.52
R^2	0.81	0.66	0.80	0.79

In addition to using the indices, model performance was also evaluated by visual inspection of the hydrograph. Visual inspection of the hydrograph involved looking at three aspects, the accuracy in simulating the peaks, low flows and the timing.

Table 4.4 Summary model evaluation statistics Amala and Nyangores (Daily basis)

Statistic	RIVER			
	Amala		Nyangores	
	Calibration	Validation	Calibration	Validation
NSE	0.46	0.14	0.33	0.29

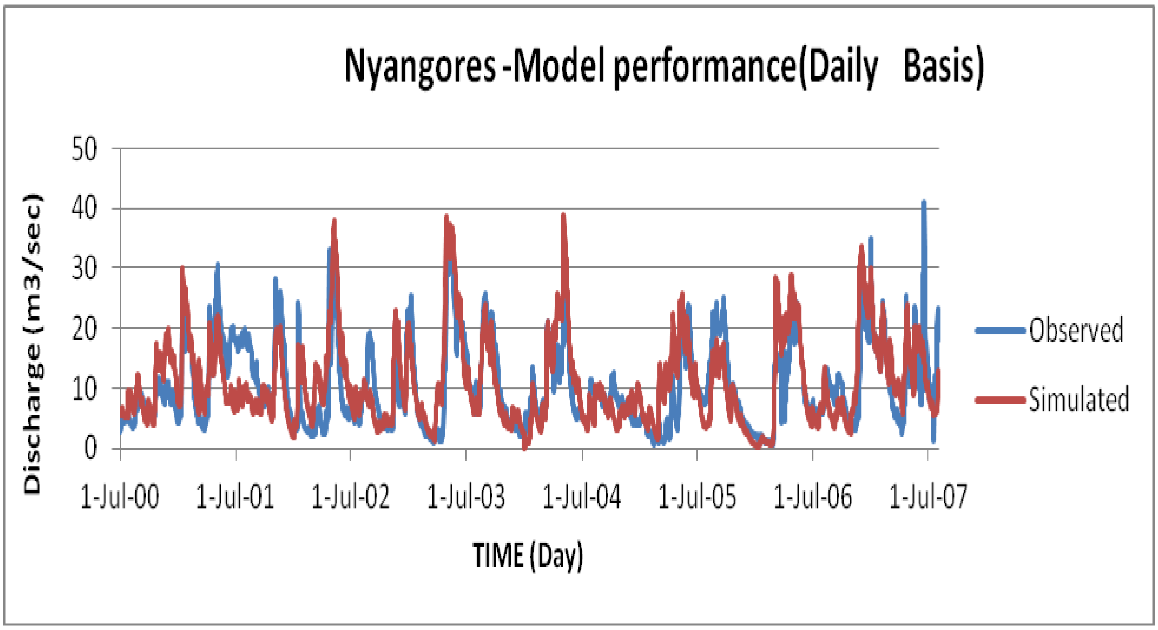

Figure 4.5 Nyangores model performance (daily basis)

Figures 4.5.and 4.6 illustrate the performance of simulation of Nyangores river on daily basis and aggregated to monthly basis. From visual inspection of the hydrograph it can be seen that the model largely simulates the pattern of flow of the river. The model consistently marginally overestimates the peaks throughout the simulation period. The timings are in tandem, the simulated and observed. On the overall the simulation of Nyangores meets the acceptance threshold of a validation NSE of 0.50.

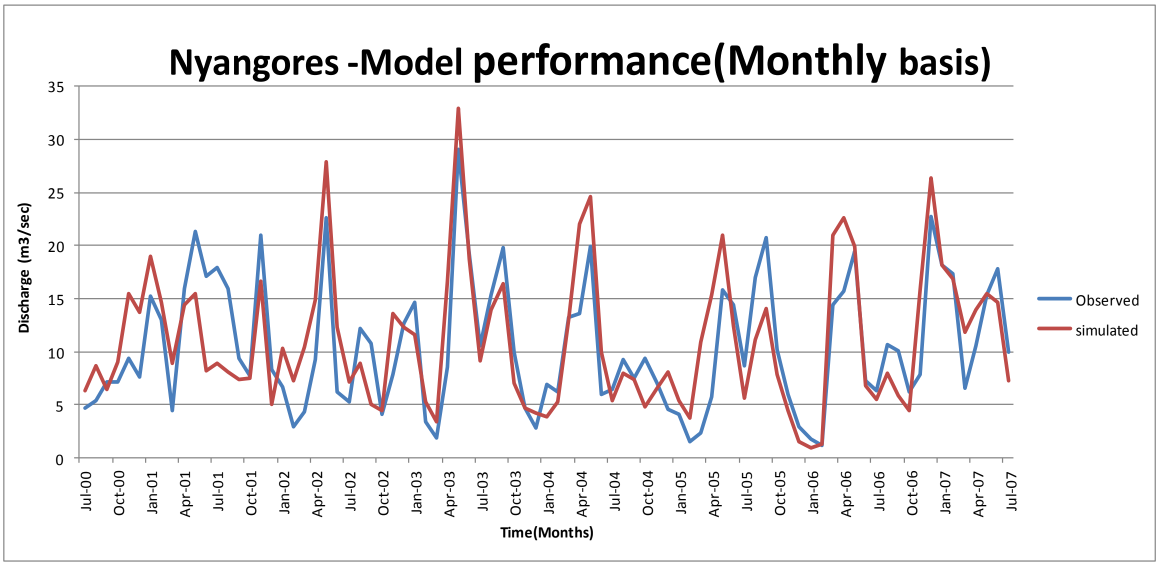

Figure 4.6 Nyangores model performance (monthly basis)

The performance of the simulation of Amala is below satisfactory with a validation NSE of 0.35. Visual examination of the hydrograph (Fig 4.7 and Fig 4.8) indicates the model largely overestimates the base flows, and has a mixed performance as regards the peaks. During the period of April 2005, the model indicates a peak which does not exist on the observed discharges. However over the same period such a peak occurs on the Nyangores hydrograph. The timings of the pattern of flow are satisfactory.

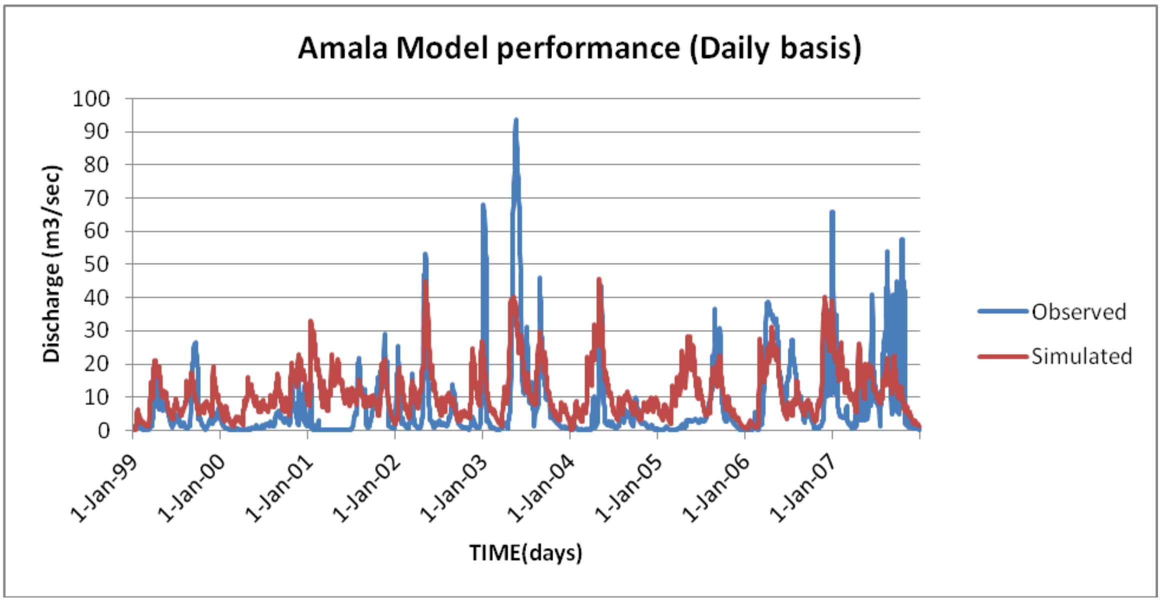

Figure 4.7 Amala model performance

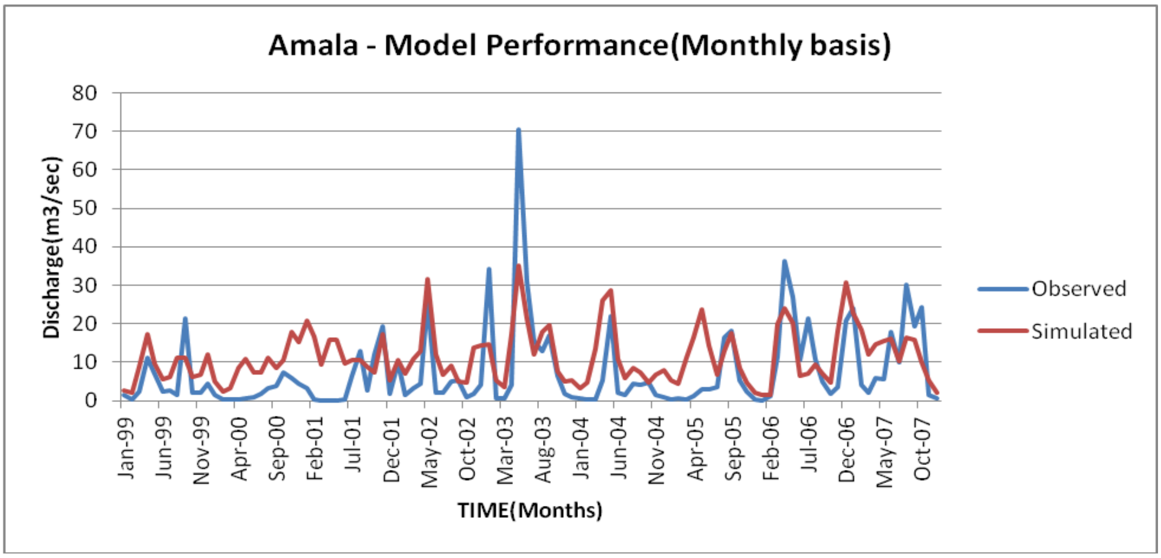

Figure 4.8 Amala model performance (monthly basis)

The less than satisfactory performance in simulating the Amala hydrograph can largely be attributed to the input data. While the main driver of hydrological processes is precipitation, it is noted that there is no precipitation data collected within the Amala catchment, the three rain gauge stations within the upper Mara catchment are all within the Nyangores catchment with an additional station Narok located outside the Upper Mara catchment. Amala is generally much drier than Nyangores from field observations and therefore applying precipitation data collected in Nyangores on Amala is likely to generate inaccuracies. This partly explains the simulated peak on the Amala hydrograph and which is not on the observed hydrograph, yet such a peak occurs on the observed Nyangores hydrograph. This could be a result of a highly localised, high intensity rainfall event that occurred in the Nyangores catchment and not within the Amala catchment.

4.5 Land use change scenarios

The land use scenarios described in chapter 3.5 were applied to both Amala and Nyangores catchments. These were applied in the model by changing the land use maps as the various scenarios were run. Three land use change scenarios were applied as described earlier and the results analysed for both the Amala and Nyangores rivers. Fig 4.9, 4.10 and 4.11 illustrate the results of the three scenarios for Amala river. In general it can be seen that for scenario1, there is minimal change in the hydrograph. Scenario 2 results in some change on the hydrograph, the flows increase whereas scenario 3 results in reduction of flows. This response pattern is similar for Nyangores River (See fig 4.12, 4.13 and 4.14.)

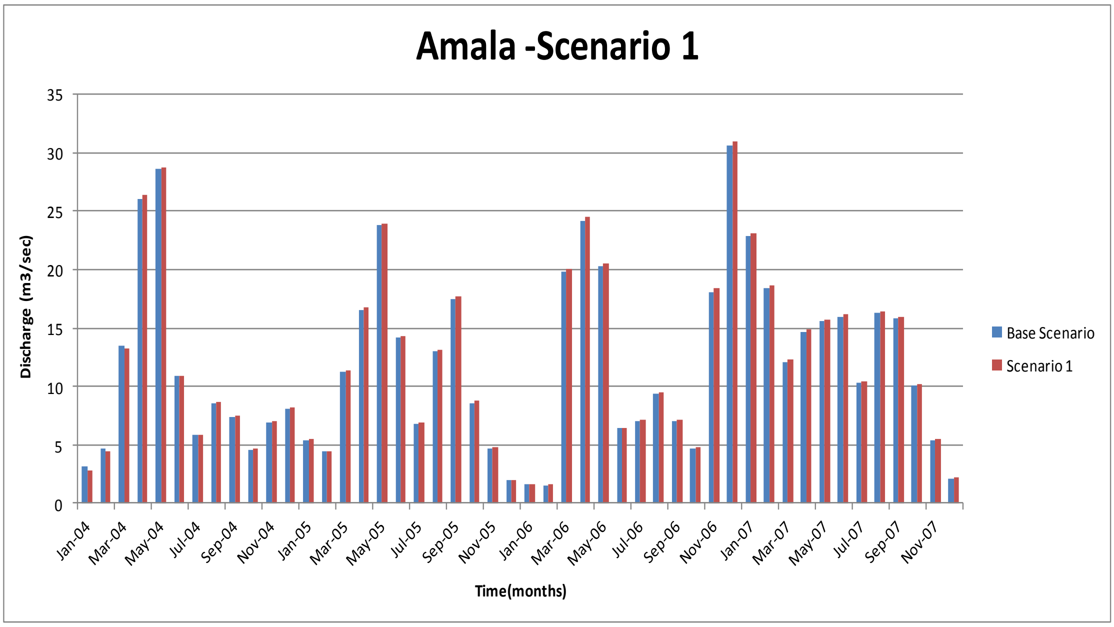

Figure 4.9 Amala Partial deforestation scenario

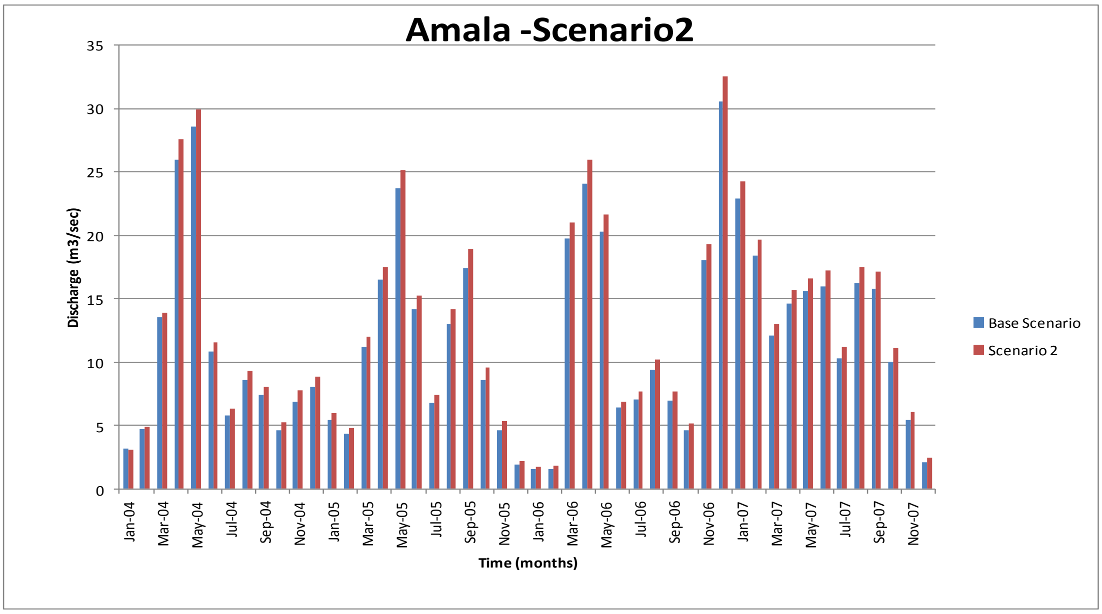

Figure 4.10 Amala conversion of forest to agriculture scenario

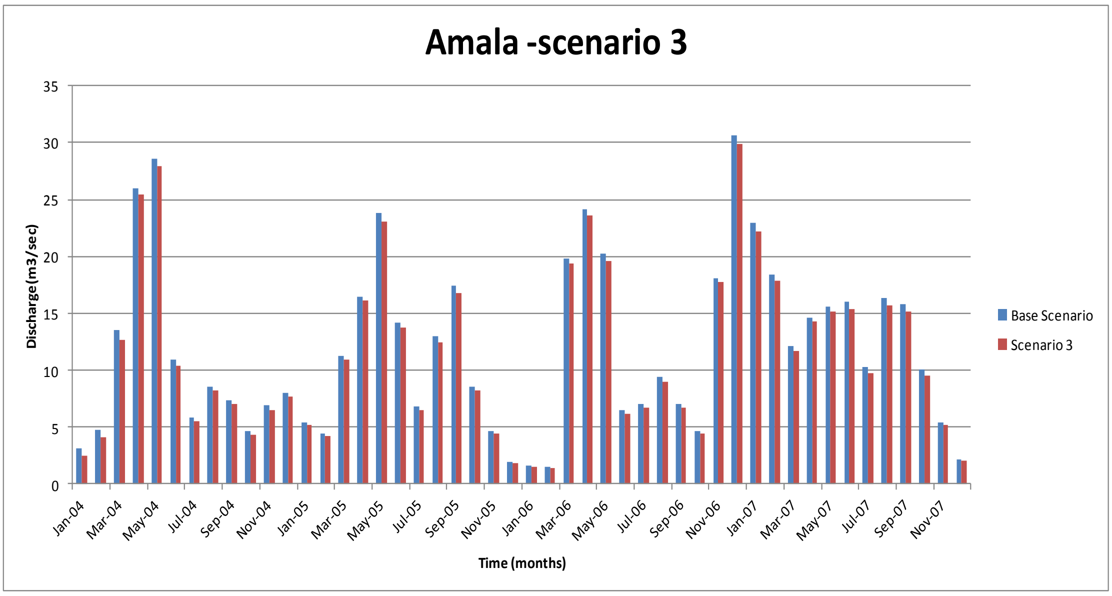

Figure 4.11 Amala re-afforestation scenario

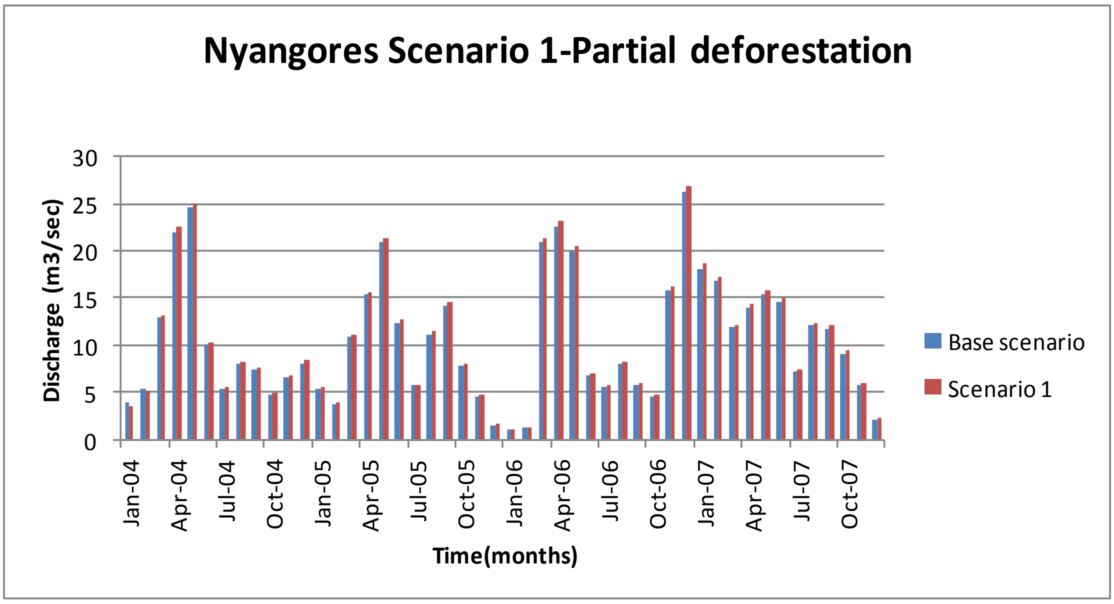

Figure 4.12 Nyangores partial deforestation scenario

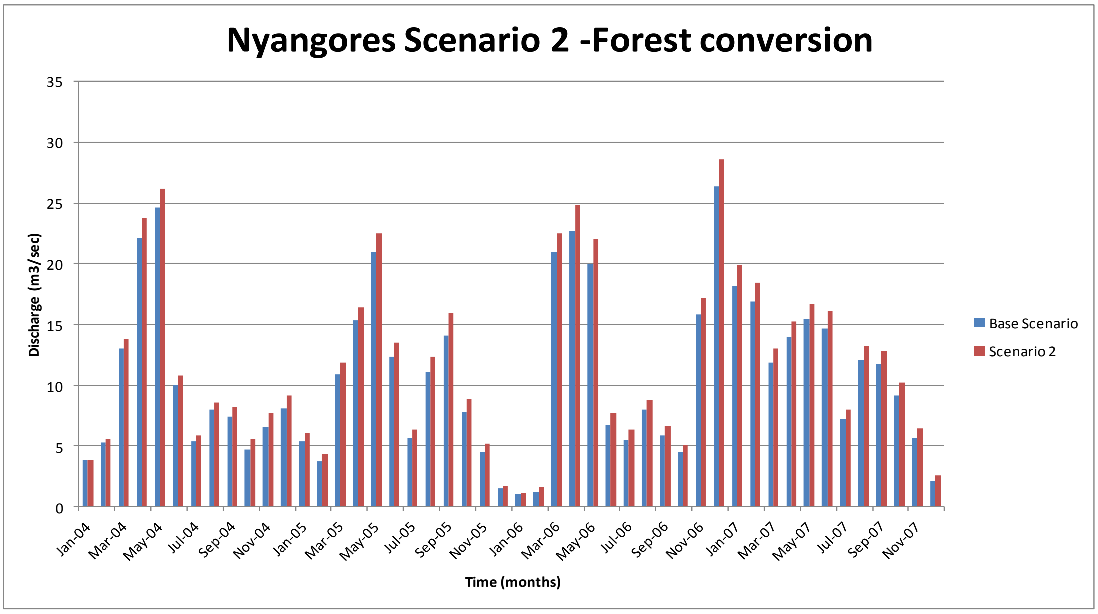

Figure 4.13 Nyangores conversion of forest to agriculture scenario

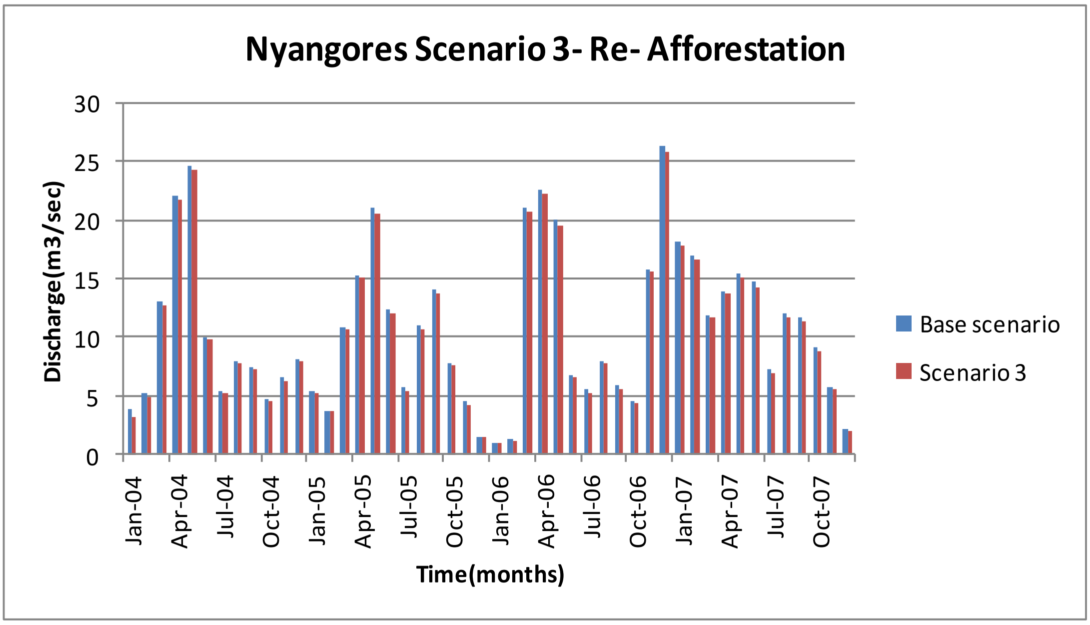

Figure 4.14 Nyangores re-afforestation scenario

4.5.1 Scenario hydrograph analysis

Table 4.5 below gives statistics of the simulated hydrographs. The data obtained from the simulation was analysed using specialised software (The Indicators of Hydraulic alteration software).

Table 4.5 Land use change scenario statistics for Nyangores River

No	Parameter	Base scenario	Scenerio1	%Change	Scenario2	%change	Scenario3	% change
			Deforestation	(+/-)	Conv to agric	(+/-)	Afforestation	(+/-)
1	7 day Minimum	0.8372	0.7641	-9	0.8894	6	0.6909	-17
2	30 day Minimum	2.059	2.009	-2	2.255	10	1.843	-10
3	90 day Minimum	5.443	5.647	4	6.226	14	5.225	-4
4	7 day Maximum	29.31	29.89	2	31.34	7	28.87	-2
5	30 day Maximum	23.87	24.37	2	25.69	8	23.45	-2
6	90 day Maximum	18.55	18.94	2	20.06	8	18.21	-2
7	days with zero flow	0	0	0	0	0	0	0
8	Annual CV	0.69	0.68	-1	0.67	-3	0.7	1
9	Mean Annual flow	10.64	10.9	2	11.67	10	10.37	-3
10	Base flow index	0.08	0.07	-12	0.08	-4	0.07	-16

N/b: The discharge figures are expressed in m^3/sec

Table 4.5 gives analysis statistics associated with the various land use change scenarios for the Nyangores River. Figures 4.15, 4.16 and 4.17 illustrate these statistics. It can be seen that the mean annual flow increases by 2% when there is deforestation, 10% when all forest is converted to agriculture and decreases by 3% when afforestation takes place.

The 7 day minimum flow decreases with deforestation by 9% and increases with conversion to agriculture by 6% and reduces by 17% when afforestation takes place. The peaks in the hydrograph show an increase with deforestation of 2%, and a marked increase of 7% with conversion to agriculture and a decrease of 2% with afforestation. Fig 4.18 and 4.19 illustrate the monthly mean flows and their associated changes with the various scenarios. It can be seen that the highest change in mean monthly flows occurs with conversion to agriculture. Deforestation results in average reduction in the monthly mean flow by 2.6%, while conversion to agriculture an increase of approximately 10% and reforestation a reduction of approximately 3%, this is illustrated in fig 4.19.

Base flow contribution to the river flow is 8% in the base scenario and decreases to 7% with deforestation, to 7.6% with conversion to agriculture and to 6.6% with afforestation. The associated percentage changes are 12%, 4% and 16% respectively. The highest impact of land use change is in the conversion to agriculture scenario and occurs in the months of September, October and November.

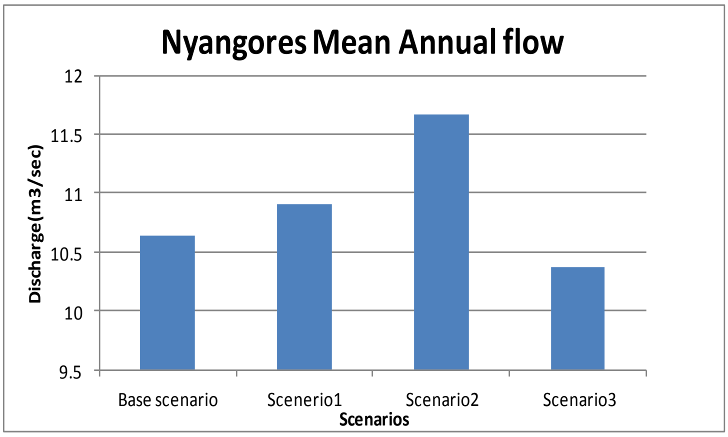

Figure 4.15 Nyangores scenario mean annual flows

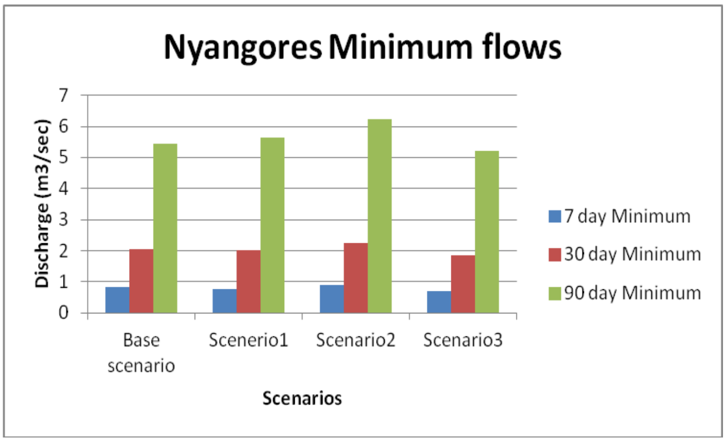

Figure 4.16 Nyangores scenario minimum flows

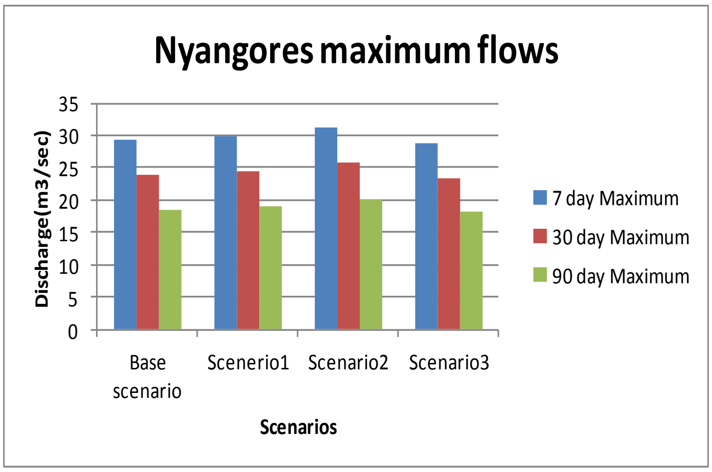

Figure 4.17 Nyangores scenario Maximum flows

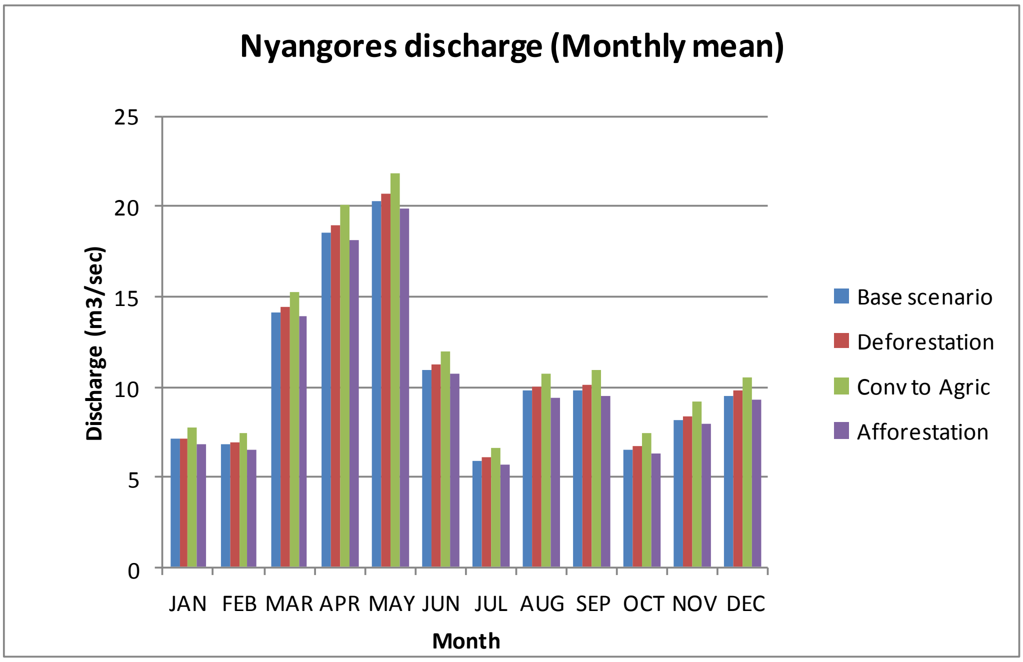

Figure 4.18 Nyangores scenarios mean monthly flows

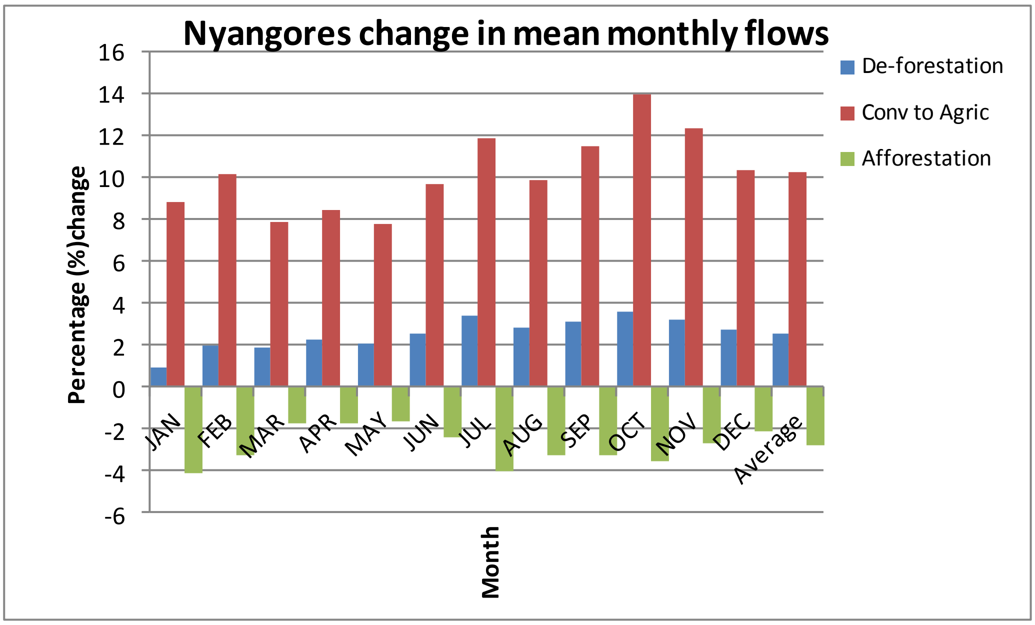

Figure 4.19 Nyangores scenario percentage change in mean monthly flows

Table 4.6 Land use change scenario statistics for Amala River

No	Parameter	Base scenario	Scenerio1	%Change	Scenario2	%change	Scenario3	% change
			Deforestation	(+/-)	Conv to agric	(+/-)	Afforestation	(+/-)
1	7 day Minimum	0.877	0.848	-3	0.941	7	0.781	-11
2	30 day Minimum	2.112	2.043	-3	2.284	8	1.864	-12
3	90 day Minimum	5.815	5.933	2	6.493	12	5.526	-5
4	7 day Maximum	34.9	35.38	1	36.85	6	34.29	-2
5	30 day Maximum	28.15	28.47	1	29.8	6	27.48	-2
6	90 day Maximum	20.41	20.62	1	21.66	6	19.83	-3
7	days with zero flow	0	0	0	0	0	0	0
8	Annual CV	0.7	0.71	1	0.69	-1	0.72	3
9	Mean Annual flow	11.82	11.95	1	12.71	8	11.38	-4
10	Base flow index	0.07	0.07	-5	0.07	-1	0.07	-8

N/b: The discharge figures are expressed in m^3/sec

Table 4.6 above gives the flow change statistics for the various land use change scenarios for Amala River. The mean annual flow increases by 1% with deforestation, 8% with conversion to agriculture and decreases by 4% for afforestation. The 7 day minimum flow decreases by 3% with deforestation, increases by 7% with conversion to agriculture and decreases by 11 % with re-afforestation. The contribution of base flow to river flow follows a similar trend reducing by 5%, 1% and 8% for deforestation, conversion and afforestation respectively. The peak flows (7 day maximum) increase by 1% and 6% for deforestation and conversion to agriculture respectively and decrease by 2% for afforestation (See fig 4.20,4.21 and 4.22).

The impact of land use change on the mean monthly flows is illustrated by figures 4.23 and 4.24. The average change in monthly means is an increase of 1%, 8% for deforestation and conversion to agriculture respectively and a decrease of 4% for afforestation. The highest impact is felt with conversion to agriculture, occurring mainly in the months of September, October and November.

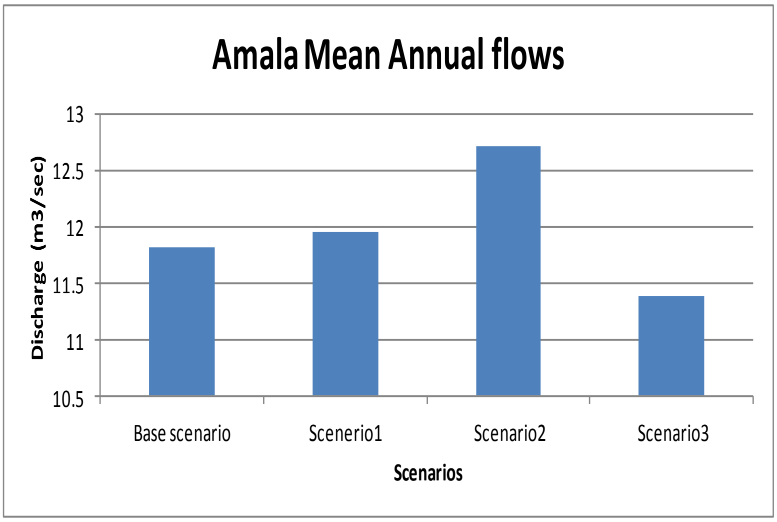

Figure 4.20 Amala scenario mean annual flows

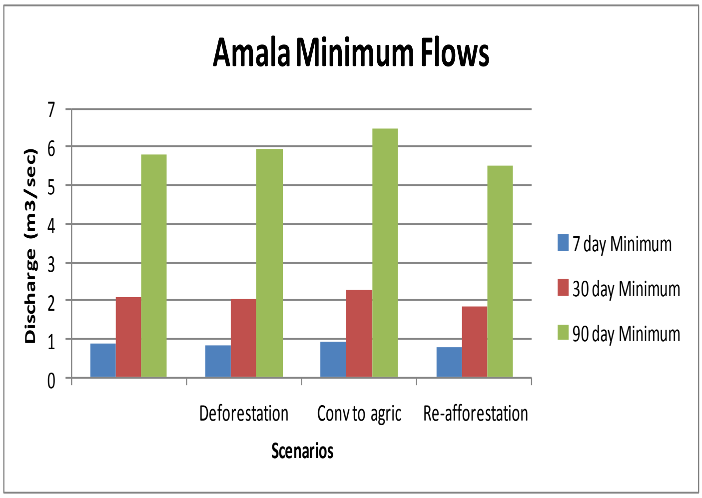

Figure 4.21 Amala minimum scenario flows

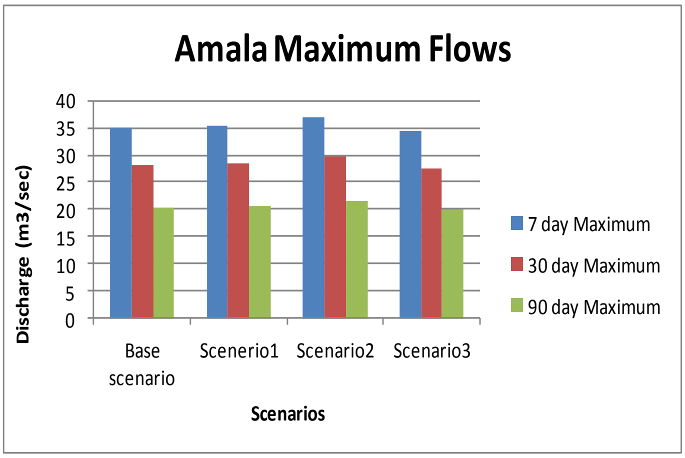

Figure 4.22 Amala scenario maximum flows

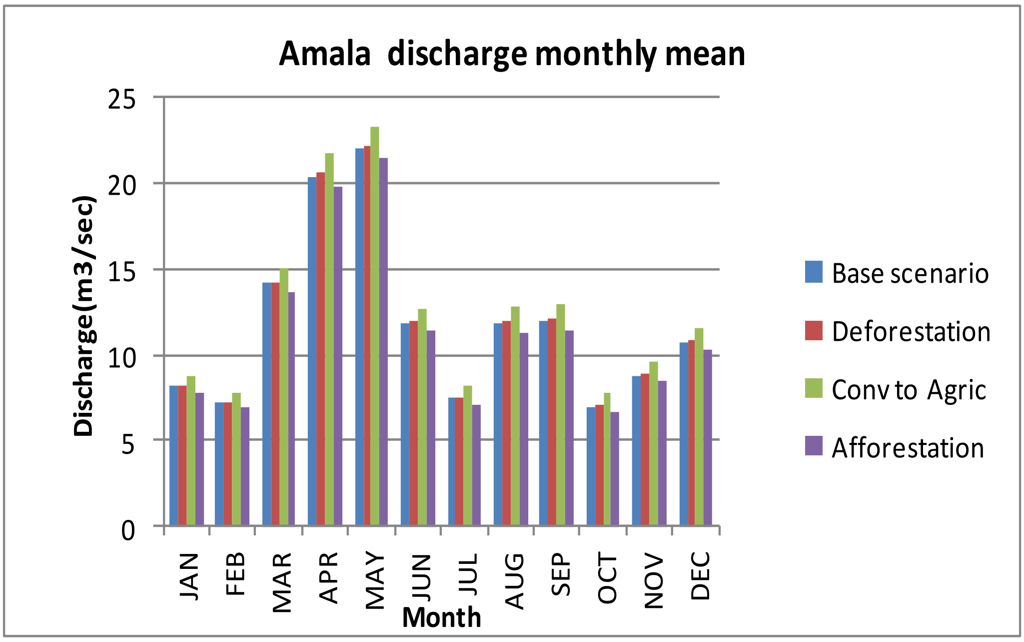

Figure 4.23 Amala scenario mean monthly flows

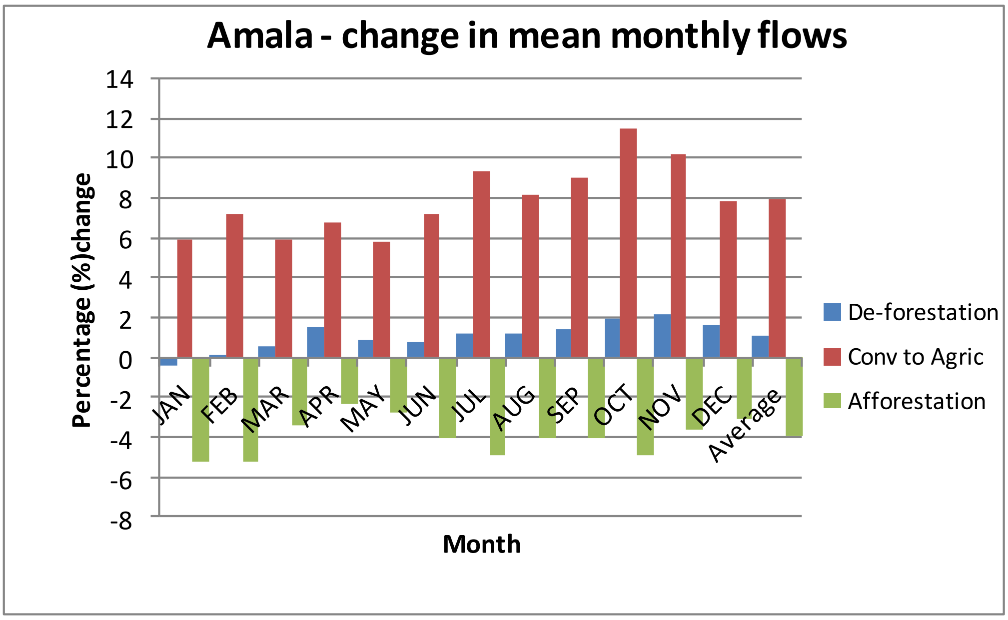

Figure 4.24 Amala percentage change in scenario mean monthly flows

4.5.2 Scenario hydrograph analysis discussion

The results presented above illustrate the simulated changes in the Mara hydrograph as a result of land use changes in the basin. Similar trends are simulated to both Nyangores and Amala but with slightly differing magnitudes.

The partial deforestation scenario results in increases in the mean annual flows, accompanied with increases in peak flows and decreases in the 7 day minimum flows. This results from decreases in evapotranspiration as forest cover is lost and most of the precipitation results in quick runoff and thus contributes less to recharge of the saturated zone hence a decrease in the minimum flows whereas the quick flows result in higher peaks. In the model by deforestation the size of the unsaturated zone is reduced, resulting in less actual transpiration, as this is dependent on the size of the unsaturated zone. In addition the interception value used in the model is lower hence less interception is calculated.

The conversion of forests to agricultural use has the highest impact on the Mara hydrograph for both Amala and Nyangores rivers. This scenario results in agriculture constituting 83% of the catchment area. Mean annual flows increase, with increases in both the 7 day minimum flows and peak flows. The increases are more marked in the Nyangores hydrograph than in Amala. However there is a reduction in both hydrographs in the contribution of base flow to the river flow, this could be as a result of less recharge of the saturated zone as more water from the unsaturated zone ends up as river flow. In timing of impacts the most pronounced impacts occur in the months of September, October and November, this coincides with the beginning of the hydrologic year and probably explains this occurrence.

STREAM models conversion of forest land to agricultural land as a reduction in the size of the unsaturated zone, hence less storage and less amount of water is calculated as actual transpiration. A high separation coefficient for agricultural lands means more water is available from the saturated zone as quick flow or surface overland flow.

The afforestation scenario results in reduction in the mean annual flows, this is accompanied by reductions in the 7 day minimum flows and 7 day maximum flows. The mean annual flows reduce by 3% in Nyangores and 4% in Amala. These findings are similar to those found by Wang et al. (2008), who found a 2.3% change in stream flow for a similar percentage increase in forested area. However the 7 day minimum flow reduces by a bigger magnitude than the peaks. The contribution of base flow reduces by up to 16% in Nyangores and 8% for Amala.
Afforestation is modelled as an increase in the size of the unsaturated zone, which results in calculation of more actual transpiration. This is further compounded by a higher interception value resulting in less water ending up in the saturated zone and hence lower flows.

It is well established that forested catchments have higher evapotranspiration than grassed catchments (Zhang et al., 2001) .Studies show that reductions in runoff can be expected following afforestation of grasslands and shrub lands and may be most severe in drier regions. It is found that annual runoff reduces on average by 44% (±3%) and 31% (±2%) when grasslands and shrub lands are afforested, respectively, depending on the tree species (Farley, 2005).

Findings of paired catchment studies indicate that reduction of forest cover of less than 20% apparently cannot be detected by measuring stream flow (Brown et al., 2005).This modelling effort has used vegetation changes of 25% for deforestation and 26% for afforestation. This explains the small magnitudes in flow change in percentage terms that were simulated in modelling the Mara hydrograph.

On the other hand, Zhang et al. (2001) finds that the main process responsible for changes in water yield as a result of alterations in vegetation at the mean annual scale is evapotranspiration. Similar conclusions can be drawn from the results discussed above mainly arising from simulation of higher actual transpiration as a result of changes in the size of the unsaturated zone and hence how much is available for transpiration.

In general the model results conform to what is generally known in terms of hydrological processes and the hydrological cycle, despite the marginal performance in simulating the hydrographs especially for the Amala River, which is largely attributed to data scarcity and problems associated with regionalisation of data.

4.6 Model inter-comparison

This modelling effort was largely motivated by previous modelling efforts in the Mara and more so by the inherent challenges associated with modelling in data scarce environments. To evaluate the performance of this modelling effort, it is compared to previous modelling efforts in the Mara. It is compared with the modelling efforts using the SWAT software (Mango et al., 2011) and USGS geo spatial stream flow model (Mutie et al., 2006).

4.6.1 Framework for model inter comparison

The comparison of the modelling efforts was carried out based on four (4) aspects, the structure of the model, the input data, the modelling efficiency indices and the scenario formulation.

In analysing the model structure the comparison outlines how the different models represent the various hydrological processes and whether the models are distributed or lumped. For comparison of model input data the comparison looks at the various dataset requirements and parametisation for the various models. Model performance efficiency is compared by use of the two most common efficiency indices the R^2 and Nash and Sutcliff efficiency index. Scenarios are compared in terms constituent composition of the various land uses in the various scenarios.

4.6.2 Model structure

In differentiating between models the first consideration would be the type of model followed by the structure. A hydrological model structure refers to how the construction of a model handles the various hydrological processes. However a detailed analysis of the various model structures is beyond the scope of this research. In analysing the type of model the description will be restricted to whether it is lumped or distributed and an outline of the modelling process.

The SWAT model is a semi distributed model that uses the concept of distinct homogeneous sub basins called hydrologic response units (HRU) based on soil and land use properties, SWAT models hydrological processes at HRU scale and subsequently routes the flow to user selected outlets.

Stream is a fully distributed model that models the hydrological processes on a scale determined by the input digital elevation model, often a 90m x 90 m DEM is used. The model then accumulates the flows from every cell or unit through a local drain direction network that is generated using the elevations on the digital elevation model to an outlet on the local drain direction network.

The Geospatial stream flow model GeoSFM is a physically based semi distributed geo spatial hydrological model. This model is an extension of the geographical information system software ARCVIEW .The model uses a digital elevation model for routing water to the outlet of the basin.

All the three models were operated on daily time steps.

4.6.3 Model Input data

The three models employ the basic data requirements for hydrological modelling such as precipitation, soils and meteorological data.

More specifically SWAT input data consists of a DEM, meteorological data, soils data, land cover data, and management data within the basin. This data is specified for each hydrological response unit. The GeoSFM model requires input data of a DEM, rainfall data, soils data, land cover data and evaporation data. Similarly the input data into the Stream model is a DEM, soils data, land use data and precipitation data.

Precipitation input data was obtained from Bomet, Nyangores, Kiptunga and Narok rain gauge stations for STREAM whereas SWAT modelling utilised data from Bomet and Kiptunga rain gauge stations. For temperature data, the stations Kericho, Kisii and Narok were used for STREAM modelling whereas for SWAT Kericho and Narok were used. Soils data was obtained from the SORTER data base and classified according to FAO classifications for both models. For the river discharge data, the two models have used gauging data of the same stations on Nyangores and Amala whereas Geo SFM makes use of the Mara Mines station. It is evident that for the SWAT and STREAM models similar data sets have been used. Table 4.7 below summarises the various datasets and their sources.

Table 4.7 Mara Datasets

DATA SET TYPE		MODEL TYPE		
		SWAT	**STREAM**	**GeoSFM**
Rainfall	Rain gauge	Bomet and Kiptunga	Bomet, Narok, Nyangoes and Kiptunga Rain gauge stations.	Kenya and Tanzania metreological departments
	Rainfall estimates	FEWS remotely sensed rainfall Data	--	--
Temperature (Max/Min)		Kericho, Narok Met stations	Narok, Kericho, Kisii Met stations	Kenya and Tanzania Metreological departments
Soils		KENSORTER database	KENSORTER database	KENSORTER database
Land use		Landsat thematic Mapper data(2008)	Landsat thematic Mapper(2000)	Landsat thematic Mapper (1973,1986,2000)
DEM		SRTM(90m x 90m)	SRTM(90m x 90m)	SRTM
Discharge		Amala and Nyangores discharge measurement Stations	Amala and Nyangores discharge measurement Stations	Mara mines gauging station

4.6.4 Land use change

In modelling the Mara using STREAM model three land use change scenarios were used a 25% decrease in forest cover with the lost forest cover being converted to shrub land, complete conversion of forest to agricultural use and a 26% increase in forest mainly taken from agricultural use.

Modelling the Mara using SWAT (Mango et al., 2011) uses similar scenarios for land use change. The scenarios used for SWAT are partial deforestation (26.59% Nyangores and 41.28% Amala), complete deforestation, and conversion to agriculture (See table 4.8). In the complete deforestation scenario the land is converted to grassland.

It is important to note that in SWAT model there was distinction between the Amala and Nyangores catchments whereas while modelling in STREAM there was no distinction. Distinction of the two catchments was not deemed necessary for STREAM as modelling is done at pixel level whereas it was necessary for SWAT as modelling takes place at the level of distinct hydrological response units.

Table 4.8 Comparison of various percentage changes in land use under various scenarios

SCENARIO	SWAT		STREAM	REMARKS
	Amala	Nyangores	Amala & Nyangores combined	
Partial deforestation	26.59%	41.28%	-25%	
Afforestation	--	--	26%	STREAM only
Conversion of forest to Agriculture	-100%	-100%	-100%	% loss of forest
Complete deforestation to grassland	-100%	-100%	--	% loss of forest SWAT only

Land use change in the Geo SFM model was analysed for a 30 year period (Mutie et al., 2006).This modelling effort did not develop scenarios but rather worked with historical land use data sets derived from historical satellite imagery. Analysis is done for land use of 1973, 1986 and the year 2000. Only the years 1999 and 2000 are shared between this model and the other two described above. This indeed does not lend a suitable platform for comparison of these results with those of SWAT and STREAM. In addition, this modelling effort uses the Mara mines gauging station to analyse the impacts on the hydrograph. However comparison of trends is still valid. Mutie et al. (2006) finds that with the changing land use there is an increase in flow peaks as well as earlier occurrences of these peaks. Using the year 2000 data set, the peaks were found to increase by 7%.

Similar results are found in using the STREAM model for both the deforestation and conversion to agriculture scenarios there are increases in peak flows. For Nyangores (7%) and Amala (6%) though it is noted these refer to totally different comparison time periods.

The SWAT model with similar periods to STREAM analyses the impacts of land use change in terms of percentage change in terms of total water yield. Table 4.9 summarises these findings. STREAM model indicates percentage increases in stream flows as opposed to SWAT model except for the complete deforestation scenario. However this comparison is viewed with caution as the scenarios are not the same though similar, a case in point is the fact that in STREAM model, where the exact location of where the land use change takes place has an impact on its effect on the hydrograph, whereas SWAT differentiates the two catchments and its unit of modelling is the HRU.

Table 4.9 Comparison of Impact of Land use change on Mara hydrograph (Annual mean)

LAND USE CHANGE SCENARIO	**SWAT**		**STREAM**	
	Amala	***Nyangores***	***Amala***	***Nyangores***
Conversion of Forest to Agriculture	- 4.28%	- 3%	8%	10%
Complete deforestation	0.25%	2.93%	-	-
Partial deforestation	- 3.88%	- 2 %	1%	2%
Afforestation	-	-	- 4%	-3%

There are marked differences in the simulated impacts of the various land use change scenarios, however these are only used for indicative comparison as the scenarios have differences. Specifically they differ in the areas of different percentage changes in land use, different land use maps and differences in location of the respective changes in land use.

4.6.5 Model performance efficiency

Model performance efficiency is often carried out using the established methods of NSE and R^2.In addition to these indices, visual evaluation of the hydrographs is carried out evaluating the simulation of peaks, low flows, recessions and timings. However for model comparison of these efficiencies it is challenging to do so if the modelling time periods are different as in the case of the GeoSFM model (1983 to 1992).But for purposes of how efficient the model was in simulating the observed flow the indices suffice, this is enumerated in Table 4.10.

Table 4.10 Comparison of model efficiencies

Statistic	**MODEL TYPE**						
	SWAT		**STREAM**			**GEOSFM**	
	Amala	***Nyangores***	***Amala***	***Nyangores***	***Amala***	***Nyangores***	***Mara Mines***
NSE(Calibration)	0.076	- 0.533	0.56	0.59	-	-	-
NSE(Validation)	0.407	- 0.057	0.35	0.52	-	-	-
R^2(Calibration)	0.303	0.085	0.81	0.80	0.76	0.74	0.83
R^2(Validation)	0.413	0.321	0.66	0.79	0.72	0.69	0.87
Modelling period	2002-2006	2002 -2008	1999 - 2007	1999-2007	1983 - 1991	1983 - 1991	1983 - 1991

From the above it is evident that the STREAM model performs well in comparison to SWAT or GEOSFM in simulating the hydrograph of Amala and Nyangores rivers, indicating higher efficiency except for modelling Amala with SWAT which has a higher NSE efficiency at validation.

5.0 CONCLUSIONS AND RECOMMENDATIONS

This research in modelling the Mara aimed at developing a model to simulate the Mara hydrograph and subsequently use it to simulate the impacts of land use change. A simple rainfall-runoff model was developed to accomplish this task. The model was used to predict impacts of land use change under various scenarios and subsequently a comparison was carried out against previous hydrological modelling efforts in the Mara.

On model performance evaluation the developed model achieved NSE indices of 0.59 and 0.56 for calibration and 0.52 and 0.35 during validation for Nyangores and Amala respectively. The R^2 values achieved were calibration 0.80 and 0.81 and validation 0.79 and 0.66 for Nyangores and Amala respectively.

There was a greater success in modelling Nyangores river than Amala River. The better performance in Nyangores can be attributed to the fact that the precipitation data used in modelling was collected within the Nyangores catchment with an additional station of Narok. Incidentally this works against the efficiency of modelling Amala as the same data is regionalised over the Amala catchment, whereas Amala is observed to be a much drier catchment than Nyangores. In effect this suggests that with better rainfall data the probability of successfully modelling Amala would be greatly increased.

The developed model was used to simulate the Mara hydrograph under three land use change scenarios of partial deforestation, conversion of forests to agriculture and reforestation. In the simulation, deforestation resulted in an increase in peak flows of 2% for Nyangores and 1% for Amala. The 7 day minimum flows reduced by 9% for Nyangores and 3% for Amala.

The afforestation scenario resulted in reduction of flows both peaks and 7 day minimum flows. 7 day minimum flows reduced by 17% for Nyangores and 11% for Amala. Peak flows reduced by 2% for both Nyangores and Amala. Reduction in peak flows could be a result of increased interception and evapotranspiration, indicating that with the current model parametisation most of the water is lost through interception and evapo-transpiration and little recharge of the aquifer takes place.

Conversion of land from forest to agriculture had a marked impact on the hydrograph. This is modelled as higher peaks (7% Nyangores and 6% Amala) and higher 7 day minimum flows (6% Nyangores and 7% Amala) but reduced contribution of base flow to the hydrograph (4% Nyangores and 1% Amala). Similar trends are indicated for both Amala and Nyangores rivers. This could be as a result of increased interflow and reduced recharging of ground water that yields base flow.

However it must be noted that these changes when viewed in terms of actual change in flow are very small and could easily fall within the margin of error.

The second objective of this research was to evaluate the performance of the developed model, in terms of simulating the Mara hydrograph in comparison to previous modelling efforts in the Mara basin. When compared to modelling the Mara using SWAT (Mango et al., 2011) the model records an NSE of 0.52 for Nyangores and 0.35 for Amala whereas SWAT records results of an NSE of 0.41 for Amala and - 0.06 for Nyangores. The SWAT model fails to simulate Nyangores, though has comparable results for Amala.

However when remotely sensed estimates are used in SWAT modelling there is an improvement in the results with an NSE of 0.586 and 0.622 for Nyangores and Amala respectively for calibration and validation values of 0.094 for Nyangores and 0.390 for Amala. However the improvements using RFE are marginal for the validation process, casting doubt on the effectiveness of RFE data.

As can be seen from the model inter-comparison results, the developed stream model performs fairly well in simulating the Mara hydrograph achieving higher modelling efficiencies as compared to the other models.

STREAM model is a simple model with limited complexity, but as is illustrated with modelling the Mara, its performance rivals that of complex, heavily parameterized models such as SWAT. The hypothesis that postulates that "the performance of a simple, conceptual distributed hydrological model in modelling the river discharge of the upper Mara is comparable to that of complex models" is supported by the findings of this research.

Modelling the Mara using PCRaster shows great potential for improvement. A fully distributed model is best suited for modelling spatially distributed changes like land use change, as these can be modelled exactly where they occur as opposed to lumped systems, but this is heavily dependent on the availability and quality of datasets. Unfortunately this remains a challenge in data scarce environments.

To improve modelling efforts in the Mara using PCRaster the following actions are recommended:

- Accurate determination of the model parameters through field investigations such as the size of the unsaturated zone (field capacity).
- Improvement of precipitation data input through use of remotely sensed data especially for the Amala catchment. This is informed by the fact that the previous research (Mango et al., 2011) has registered improved performance using remotely sensed data.
- Use of progressive land use change maps could improve accuracy in predicting changes on the Mara hydrograph.
- A sensitivity analysis should be carried out to identify important parameters, test the model conceptualisation and improve the model structure through better understanding of the underlying hydrological processes.

For management information purposes all the three models indicate that deforestation will result in higher peak flows but reduced minimum flows that are most needed during the dry season. Whereas afforestation efforts (as simulated in STREAM) will result in reduced peak flows.

References

Aerts, J. C. J. H., Kriek, M., & Schepel, M. (1999). STREAM (Spatial tools for river basins and environment and analysis of management options): 'set up and requirements'. [doi: 10.1016/S1464-1909(99)00049-0]. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 24*(6), 591-595.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., & Williams, J. R. (1998). Large Area hydrologic modeling and assessment PART I: MODEL DEVELOPMENT1. *JAWRA Journal of the American Water Resources Association, 34*(1), 73-89.

Barno, S. (2011). Water Allocation in Mara River basin, balancing Water for Agriculture and Environment. *Msc Thesis*.

Bergstrom, S. (1991). Principles and Confidence in Hydrological Modelling. *Nordic Hydrology, 22*, 123 - 136.

Beven, K. (1996). The limits of splitting: Hydrology. [doi: 10.1016/0048-9697(95)04964-9]. *Science of The Total Environment, 183*(1-2), 89-97.

Beven, K. (2005). Rainfall-runoff Modelling:Introduction. *Encyclopedia of Hydrological Sciences*.

Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H. G., Gräff, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., & Viney, N. R. (2009). Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use. [doi: 10.1016/j.advwatres.2008.10.003]. *Advances in Water Resources, 32*(2), 129-146.

Bronstert, A., Niehoff, D., & Bürger, G. (2002). Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities. *Hydrological Processes, 16*(2), 509-529.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., & Vertessy, R. A. (2005). A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. [doi: 10.1016/j.jhydrol.2004.12.010]. *Journal of Hydrology, 310*(1–4), 28-61.

Ewen, J. (2011). Hydrograph matching method for measuring model performance. [doi: 10.1016/j.jhydrol.2011.07.038]. *Journal of Hydrology, 408*(1-2), 178-187.

Farley, K. A., Jobbágy, E. G., & Jackson, R. B. (2005). Effects of afforestation on water yield: a global synthesis with implications for policy. *Global Change Biology, 11*(10), 1565-1576.

Gerrits, A. M. J. (2010). *The role of interception in the hydrological cycle.* VSSD, Delft.

Gerrits, M. (2005). Hydrological modelling of the Zambezi catchment for gravity measurements. *Msc thesis*.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. [doi: 10.1016/j.jhydrol.2009.08.003]. *Journal of Hydrology, 377*(1-2), 80-91.

Hargreaves, G. H., & Allen, R. G. (2003). History and Evaluation of Hargreaves Evapotranspiration Equation. *Journal of Irrigation and Drainage Engineering, 129*(1), 53-63.

International Law Association, C. (1967). *Helsinki rules on the uses of the waters of international rivers*. London: The International law association.

IPCC. (2007). An assessment of the Intergovernmental panel on Climate change,fourth assessment report. *Report*.

Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research, 29*(8), 2637-2649.

Karssenberg, D. (2002). The value of environmental modelling languages for building distributed hydrological models. *Hydrological Processes, 16*(14), 2751-2766.

Karssenberg, D., Burrough, P. A., Sluiter, R., & de Jong, K. (2001). The PCRaster Software and Course Materials for Teaching Numerical Modelling in the Environmental Sciences. [doi:10.1111/1467-9671.00070]. *Transactions in GIS, 5*(2), 99-110.

Legesse, D., Vallet-Coulomb, C., & Gasse, F. (2003). Hydrological response of a catchment to climate and land use changes in Tropical Africa: case study South Central Ethiopia. [doi: 10.1016/S0022-1694(03)00019-2]. *Journal of Hydrology, 275*(1-2), 67-85.

LVBC, W.-E. (2010). Assessing reserve flows for the Mara river *Report*.

Mango, L. M., Melesse, A. M., McClain, M. E., Gann, D., & Setegn, S. G. (2011). Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: Results of a modeling study to support better resource management. *Hydrology and Earth System Sciences, 15*(7), 2245-2258.

Mati, B. M., Mutie, S., Gadain, H., Home, P., & Mtalo, F. (2008). Impacts of land-use/cover changes on the hydrology of the transboundary Mara River, Kenya/Tanzania. *Lakes and Reservoirs: Research and Management, 13*(2), 169-177.

Mul, M. (2009). Understanding hydrological processes in an ungauged catchment in sub-saharan Africa. 2-10.

Mutie S.M, M. B., Home P,Gadain H, Gathenya J. (2006). Evaluating land use change effects on river flow using USGS Geospatial stream flow model in Mara River Basin, Kenya. *Centre for remote sensing and land surfaces,Bonn*, 141-148.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. [doi: 10.1016/0022-1694(70)90255-6]. *Journal of Hydrology, 10*(3), 282-290.

NBS. (2002). National Population census, Tanzania. National bureau of statistics, Ministry of planning ,economy and empowerment.

Savenije, H. H. G. (2001). Equifinality, a blessing in disguise? *Hydrological Processes, 15*(14), 2835-2838.

Schaefli, B., & Gupta, H. V. (2007). Do Nash values have value? *Hydrological Processes, 21*(15), 2075-2080.

Sieber, A., & Uhlenbrook, S. (2005). Sensitivity analyses of a distributed catchment model to verify the model structure. [doi: 10.1016/j.jhydrol.2005.01.004]. *Journal of Hydrology, 310*(1-4), 216-235.

Sivakumar, B. (2008a). Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stochastic Environmental Research and Risk Assessment, 22*(6), 737-748.

Sivakumar, B. (2008b). The more things change, the more they stay the same: the state of hydrologic modelling. *Hydrological Processes, 22*(21), 4333-4337.

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes, 17*(15), 3163-3170.

Uhlenbrook, S., Roser, S., & Tilch, N. (2004). Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. [doi: 10.1016/j.jhydrol.2003.12.038]. *Journal of Hydrology, 291*(3-4), 278-296.

Uhlenbrook, S., & Sieber, A. (2005). On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. [doi: 10.1016/j.envsoft.2003.12.006]. *Environmental Modelling & Software, 20*(1), 19-32.

Wang, S., Kang, S., Zhang, L., & Li, F. (2008). Modelling hydrological response to different land-use and climate change scenarios in the Zamu River basin of northwest China. *Hydrological Processes, 22*(14), 2502-2510.

Winsemius, H. C., Savenije, H. H. G., Gerrits, A. M. J., Zapreeva, E. A., & Klees, R. (2006). Comparison of two model approaches in the Zambezi river basin with regard to model reliability and identifiability. *Hydrol. Earth Syst. Sci., 10*(3), 339-352.

www.brc.tamus.edu.

Zhang, L. (2001). Response of mean annual evapotranspiration to vegetation changes at catchment scale. *Water Resour Res, 37*, 701-708.

Annex 1 Rating curves

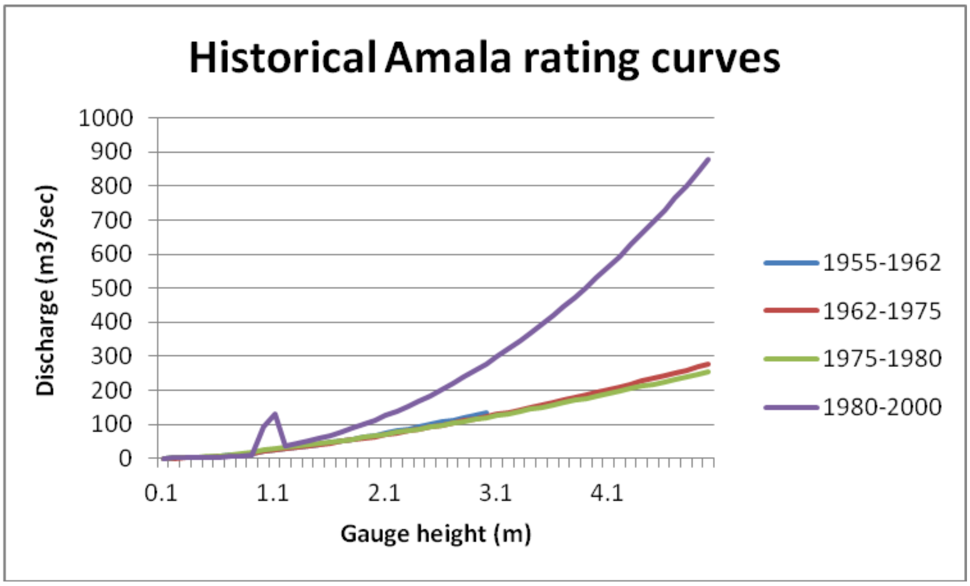

Fig A-1 Amala rating curve with validity periods. (Source: WRMA, Kenya)

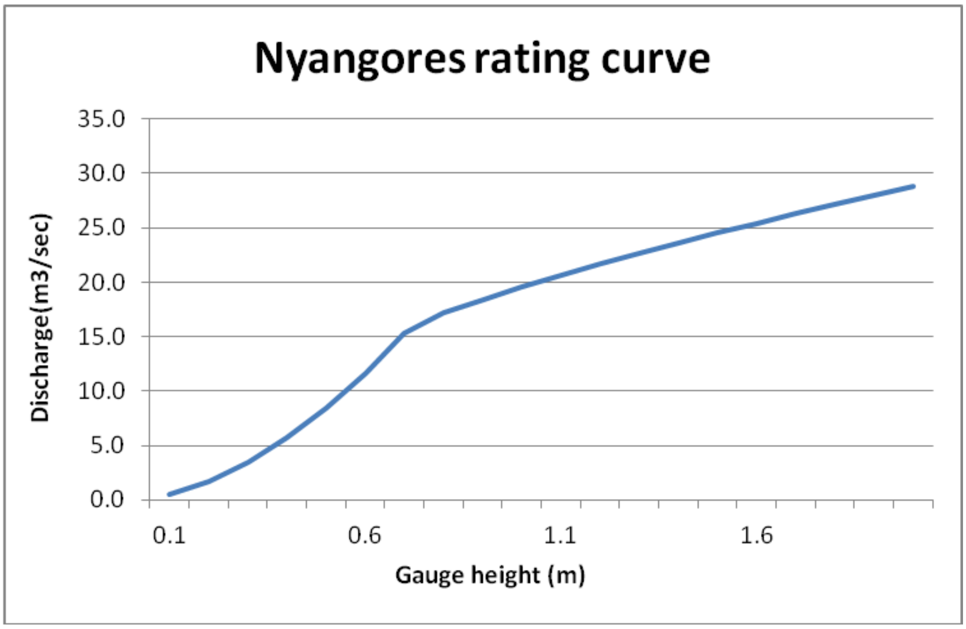

Fig A-2 Nyangores rating curve (Source : WRMA,Kenya (valid from 1963 to date))

Annex 2 Calibration Parameters

Interception

Land use	Interception(*D*)mm
Forest	4
Woodland	3
Bush land dense	2
Bush land sparse	2
Grassland	2
Agriculture dense	2
Agriculture Sparse	2
Plantation	2

Tree Top factor

Land use	F -factor
Forest	100
Woodland	70
Bush land dense	50
Bush land sparse	50
Grassland	15
Agriculture dense	25
Agriculture sparse	25
Plantation	25

Quickflow Coefficients

Soil type	Quick flow coefficients (q_c)
Clay	0.90
Sandy clay loam	0.72
Silt loam	0.50
Loam	0.80
Clay loam	0.30
Sandy loam	0.50

Separation Coefficients

Land use	Separation coefficient (C_r)
Forest	0.60
Woodland	0.60
Bush land	0.50
Grasslands	0.50
Agriculture dense	0.65
Agriculture sparse	0.60
Plantation	0.25

Annex 3 Model Script

```
# PCRaster rainfall Runoff Model
# Mara Catchment Kenya
# Adopted for the Mara basin from Zambezi STREAM PCR model by Winsemius

binding
# linking the disk file to the model variable
PrecipitationTimeseries = precipitation7.pcrtss;
EvaporationTimeseries = ET7.pcrtss;
MetStations = Metstns.map;
Evaporation = Evap;
RainStations = rainstns.map;
landuse = scenario4.map;
soil = soil30.map;
Moist = Smax.map;
Precipitation = prep;
runoff = run;
level = homdem8.map;
myldd = mask12ldd.map;
Qstations = flowstat8.map;
bound = maramask.map;
#-----------------------------------------------------------------------------------------
# Constants
#D   = 20;              # Interception threshold **************
k   = 38;               # Recession Su 2 GWS constant**********
surface = 0.008561241219;# surface of a gridcell (km2)

ConvConst = 0.0000990884; # From mm/month to m3/s: *1/1000 *
                        # 1/24 * 1/3600 * 92.5269^2
rtq = 18;               # Recession constant quick flow **********
rts = 800;              # Recession constant slow flow ***********
timestep = 2592000; # Timstep in seconds
#-----------------------------------------------------------------------------------------

timer
1 1461 1;

initial
# initialise these variables before starting the dynamic run
mask = defined(bound);
```

```
#-----------------------------------------------------------------------------------------------
Su = 50; # Initial soil moisture
GWS = 20;# Innitial storage
cap = 2;# capillary rise
TS = 1;
S1 = 0;
S2 = 0;
S3 = 0;
S4 = 0;
S5 = 0;
S6 = 0;
ROUT = 1;

#-----------------------------------------------------------------------------------------------


dynamic
# iterate for the number of timesteps over these statements
# Read the precipitation timeseries and assign the values to the rainstations map
 RainfallStat = timeinputscalar(PrecipitationTimeseries,RainStations);
# Interpolate between the rainstation to obtain local values for each rastercell
report Precipitation = inversedistance(mask,RainfallStat,2,0,0);
# Read the Evaporation timeseries and assign the values to the Metstations map
EvapStat = timeinputscalar(EvaporationTimeseries,MetStations);
# Interpolate between the Metstations to obtain local values for each rastercell
report Evaporation = inversedistance(mask,EvapStat,2,0,0);

#Calculate Net precipitation;
#----------------------------------------------------------------------------------------------
RainTss = Precipitation;       #timeinput(input\gauging\rainfall)
Ep      = Evaporation ;      #timeinput(input\gauging\potevapo)
#----------------------------------------------------------------------------------------------
#Calculation of interception
D = if(landuse == 1,4,2);   # Forest
D = if(landuse == 2,3,D);    # Woodland
D = if(landuse == 3,2.2,D);    # Bush land dense
D = if(landuse == 4,2.2,D);    # Bushland sparse
D = if(landuse == 5,2,D);    # Grassland
D = if(landuse == 11,2,D);    # Agriculture dense
D = if(landuse == 12,2,D);    # Agriculture sparse
D = if(landuse == 13,2,D);    # Plantation(tea)
D = 1.0 * D;                  # Calibration of total D
#----------------------------------------------------------------------------------------------
Int     = min(RainTss, D);
RainNet = RainTss - Int;
```

```
#------------------------------------------------------------------------------------------
Smax = Moist;
Smax = 0.60 * Smax;
#report semax = Smax;
#-----------------------------------------------------------------------------------------
#separation coefficient
cr = if(landuse == 1, 0.60,0.5);     # Forest
cr = if(landuse == 2, 0.60,0.5);     # Woodland
cr = if(landuse == 3, 0.50,0.5);    # Bush land dense
cr = if(landuse == 4, 0.50,0.5);    # Bushland sparse
cr = if(landuse == 5, 0.50,0.5);    # Grassland
cr = if(landuse == 11, 0.65,0.5);   # Agriculture dense
cr = if(landuse == 12, 0.80,0.5);   # Agriculture sparse
cr = if(landuse == 13, 0.50,0.5);   # Plantation(tea)

cr = 1.15 * cr;                     # Calibration of total cr
#----------------------------------------------------------------------------------------
#Quickflow coefficient
qc = if(soil == 1, 0.30,0.5); # clay
qc = if(soil == 2, 0.72,qc); # sandy clay loam
qc = if(soil == 3, 0.80,qc); # clay
qc = if(soil == 4, 0.9,qc); # clay
qc = if(soil == 5, 0.9,qc); # clay
qc = if(soil == 6, 0.5,qc); # Silt loam
qc = if(soil == 7, 0.30,qc); # clay
qc = if(soil == 8, 0.80,qc); # loam
qc = if(soil == 9, 0.25,qc); # clay
qc = if(soil == 10, 0.30,qc); # clay loam
qc = if(soil == 11, 0.30,qc); # clay loam
qc = if(soil == 12, 0.20,qc); # clay loam
qc = if(soil == 13, 0.30,qc); # clay loam
qc = if(soil == 14, 0.50,qc); # sandy loam
qc = if(soil == 15, 0.50,qc); # silt loam
qc = if(soil == 16, 0.50,qc); # sandy loam
qc = if(soil == 17, 0.50,qc); # sandy loam
qc = if(soil == 18, 0.70,qc); # clay
qc = if(soil == 19, 0.70,qc); # clay
qc = if(soil == 20, 0.50,qc); # sandy loam
qc = if(soil == 21, 0.50,qc); # sandy loam
qc = if(soil == 22, 0.50,qc); # sandy loam
qc = qc * 0.90;
#---------------------------------------------------------------------------------------
#treetop factor
```

```
f = if(landuse == 1, 100,25);  # Forest
f = if(landuse == 2, 70,f);    # Woodland
f = if(landuse == 3, 50,f);    # Bush land dense
f = if(landuse == 4, 50,f);    # Bushland sparse
f = if(landuse == 5, 15,f);    # Grassland
f = if(landuse == 11, 25,f);   # Agriculture dense
f = if(landuse == 12, 25,f);   # Agriculture sparse
f = if(landuse == 13, 25,f);   # Plantation(tea)

#report TOTtree = f;
#Smax = soil;
#-------------------------------------------------------------------------------------------
Su      = Su + (1 - cr) * RainNet;
Overtop = max(0, ((Su - Smax) / k));
#OLAND = (max(0, (Su - Smax))) - Overtop;
#report otop = Overtop;
#report oland = OLAND;
Su      = Su - Overtop;
#su = Su - OLAND;
Tp      = max(0,Ep - (1 - (f/100)) * Int);
#report trees = f;
#report trans = Tp+0;
Ta      = (Tp * min(1, ((2 / Smax) * Su)));
Ta      = min(Ta, Su);
#report transp = Ta;
#------------------------------------------------------------------------------------------

Su      = Su - Ta;
#-------------------------------------------------------------------------------------------
# Calculate satur. overland flow, quick flow, slow flow
#-------------------------------------------------------------------------------------------
# Sat. ov. flow
GWSdem = level;
gwsmax = 24 * ln(GWSdem);
GWS      = GWS + (cr * RainNet) + Overtop;
saof     = if(GWS>gwsmax, GWS - gwsmax, 0);
#report overland = saof;
saof = saof/15;
GWS      = GWS - saof;
#saof = saof ;

#-------------------------------------------------------------------------------------------
# Quick flow
```

```
#-----------------------------------------------------------------------------------------------
GWSquick = gwsmax * qc;    # Certain soil level which determines the maximum level, available for
quickflow
qflo     = max((GWS-GWSquick), 0) / rtq;
GWS      = GWS - qflo;
report quick = qflo;
#-------------------------------------------------------------------------------------------------
# Capillary rise
# PotentialCap = if(GWS < -25, 2, (2 * cap));
# CapRise      = PotentialCap;
# GWS          = GWS - CapRise;
# Su           = Su + CapRise;
#-------------------------------------------------------------------------------------------------
# Slow flow
sflo = max(GWS, 0) / rts;
GWS = GWS - sflo;
#report slow = sflo;
#-------------------------------------------------------------------------------------------------
Runoff = saof + qflo + sflo;
#-------------------------------------------------------------------------------------------------
B1 = 0.10;
B2 = 0.90;
B3 = 0.00;
Q4 = Runoff * B1;
#--------------------------------------------------------------------------------------------------
S1 = S2 + Q4 ;
S2 = S3 + (B2 * Runoff);
S3 = B3 * Runoff;

#---------------------------------------------------------------------------------------------------
Runoff = S1;
#TS = TS + 1;
#TS = if(TS == 4,1,TS);
#---------------------------------------------------------------------------------------------------
#report riverflo = Runoff;
  RunoffNeg = if(Runoff < 0, abs(Runoff), 0);
  RunoffPos = if(Runoff >= 0, Runoff, 0);

#RunoffTot     = abs(RunoffPos) - abs(RunoffNeg);
#report TOT = RunoffTot;
#RunoffTotal = RunoffTot * ConvConst;
#report TOTrun = RunoffTotal;
#RunoffTotConv  = accuflux(myldd, RunoffTotal);
```

```
#report runcon = RunoffTotConv;

  RainNeg   = if(RainNet < 0, abs(RainNet), 0);
  RainPos   = if(RainNet >= 0, RainNet, 0);
  RunoffTotNeg  = accuflux(myldd, RunoffNeg);
  RunoffTotPos  = accuflux(myldd, RunoffPos);
  saofTot       = accuflux(myldd, saof)*ConvConst;
  sfloT         = accuflux(myldd, sflo)*ConvConst;
  qfloTot       = accuflux(myldd, qflo)*ConvConst;
  #OLAND2       = accuflux(myldd, OLAND)*ConvConst;
  RunoffTot     = RunoffTotPos - RunoffTotNeg;
  RunoffTotConv = RunoffTot * ConvConst;

#report allrun = RunoffTotConv;
  RainNetNeg    = accuflux(myldd, RainNeg);
  RainNetPos    = accuflux(myldd, RainPos);
  RainNetTot    = (RainNetPos - RainNetNeg)*ConvConst;
  basinarea     = accuflux(myldd, surface);
  Runoffmm      = RunoffTot/basinarea * 0.008561241219;
#  RainTot       = accuflux(myldd, RainTss)*ConvConst;
  RainTot       = accuflux(myldd, RainTss)/basinarea * 0.008561241219;
  GWSav         = catchmenttotal(GWS,myldd)/basinarea * 0.008561241219;
  Suav          = accuflux(myldd, Su)/basinarea * 0.008561241219;
#====================================================================
#REPORTING
#====================================================================

#report timeout\transpir   = Ta;
#report timeout\biomass    = Biomass;
#report timeout\GWSmap     = GWS;
#report timeout\Sumap      = Su;
  report qneg.tss    = timeoutput(flowstat8.map,RunoffTotNeg);
  report qpos.tss    = timeoutput(flowstat8.map,RunoffTotPos);
  report qoutmm.tss   = timeoutput(flowstat8.map,Runoffmm);
  report qout.pcrtss    = timeoutput(flowstat8.map,RunoffTotConv);

  report saof.tss    = timeoutput(flowstat8.map,saofTot);
  report qflo.tss    = timeoutput(flowstat8.map,qfloTot);

  report Runoff.tss  = timeoutput(flowstat8.map,Runoff);
  #report Oland.tss  = timeoutput(flowstat8.map,OLAND2);#-------------
  report gws.tss     = timeoutput(flowstat8.map, GWS);
  report su.tss      = timeoutput(flowstat8.map, Su);
  report raineff.tss = timeoutput(flowstat8.map, RainNetTot);
```

```
report rainbrut.tss = timeoutput(flowstat8.map, RainTot);
report basinpointarea.out = timeoutput(flowstat8.map, basinarea);
report basinpointarea.tss = timeoutput(flowstat8.map, basinarea);
report gwsav.tss   = timeoutput(flowstat8.map, GWSav);
report Suav.tss    = timeoutput(flowstat8.map, Suav);
report sflo.tss    = timeoutput(flowstat8.map, sfloT);
```