EFFECT OF SURFACE WATER ABSTRACTIONS ON THE INFLOW INTO LAKE NAIVASHA (KENYA) USING CREST MODEL FORCED BY EARTH OBSERVATION BASED DATA PRODUCTS

OCHIENG KENNEDY OTIENO
March, 2017

SUPERVISORS:
Dr. Ir. R. van der Velde
Drs. R. Becht
Ir. Henry Munyaka (Advisor)

EFFECT OF SURFACE WATER ABSTRACTIONS ON THE INFLOW INTO LAKE NAIVASHA (KENYA) USING CREST MODEL FORCED BY EARTH OBSERVATION BASED DATA PRODUCTS
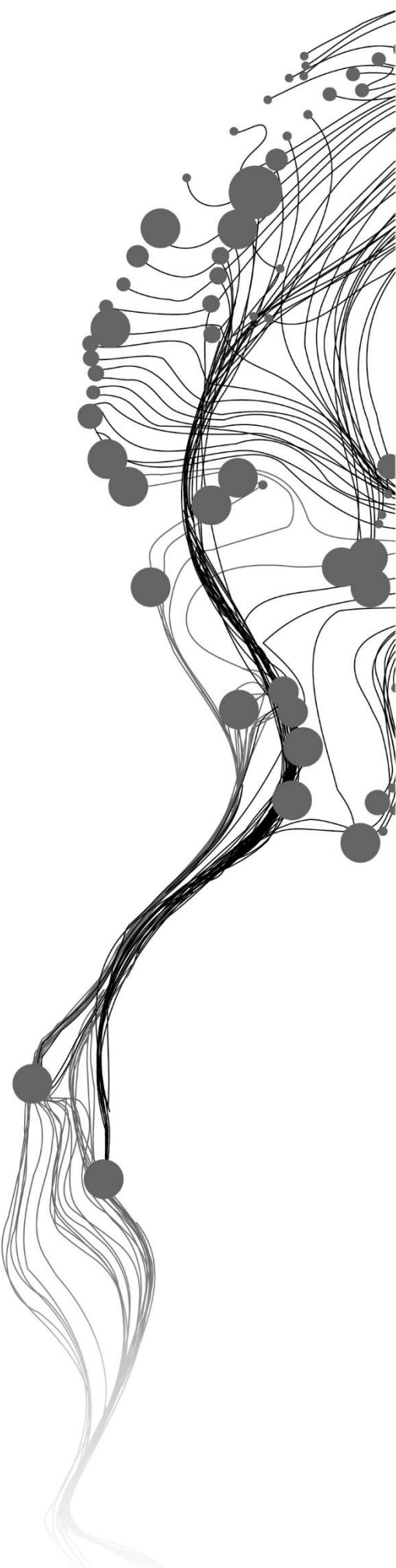
OCHIENG KENNEDY OTIENO
Enschede, The Netherlands,
[March, 2017]

SUPERVISORS:
Dr. Ir. R. van der Velde
Drs. R. Becht
Ir. Henry Munyaka (Advisor)

THESIS ASSESSMENT BOARD:
Dr. M. W. Lubczynski (Chair)
Dr. W. van Verseveld, (External Examiner, Deltares, The Netherlands)

ABSTRACT

Drying up of Lake Naivasha would be detrimental to the existence of both lives and ecosystems in Kenya. The lake is also the backbone of robust economic activities since it supports horticulture, floriculture, tourism, among others. Meteorological forcing are major driving factors of the water balance of the basin. However, anthropogenic influences such as abstraction of water from both the Lake and rivers can cause a drawdown of lake levels and consequently drying up if it is not regulated. Lack of enough water can cause conflicts between upstream and downstream users. To ensure there is environmental flow on the rivers, there has to be a withdrawal limit. Stream flows below the Q_{80} are considered as normal flows and should not be allocated while flood flows are those above this threshold and can thus be allocated for abstractions. This study aims to quantify the effects of surface water abstractions on the inflow into the lake. A lake balance model was used to assess the approximate inflows required to sustain the high lake levels. The quantification of the amount of stream flow was done using a distributed hydrological model which was developed by the University of Oklahoma and NASA SERVIR. Important data include *In-situ* stream flows, rainfall and evaporation. In this study, hydrological data collected from Water Resources Management Authority in Naivasha was assessed for credibility and reliability before application. There were gaps in the both stream flow and *in-situ* rainfall data. Flow duration curves for different stations were also prepared and it was seen different sections of the rivers experience different flow regimes. Reliability of the stream flows was also assessed using double mass curves. These curves showed that different stations had varying levels of discharge data reliability. Since the rainfall data also had gaps, satellite retrieved rainfall products provided an alternative data source for calibrating the CREST model. TRMM3B42v7 and CHIRPS rainfall products were used for this purpose and their reliability was assessed using *in-situ* data as a benchmark. Cumulative values showed both products underestimated rainfall amounts. Both satellite products showed improvement in performance in terms of RMSE when monthly values are compared against the daily values. The RMSE values for TRMM3B42v7=7.616 mm/day, CHIRPS=6.922 mm/day and TRMM3B42v7=2.454 mm/month and CHIRPS=2.564mm/month. Only one evaporation station is available in the basin i.e. the Lake station and this could not be used since it was not representative and it had gaps. FEWSNET PET was used as an alternative input to CREST model. In application of the satellite data for calibrating the CREST model over the period 2001 to 2010, the model gave lower values of calibration objective functions. PEST calibration technique was used for this purpose. The *NS* performance for Malewa (TRMM3B42v7=0.283, CHIRPS=0.036), Gilgil (TRMM3B42v7=0.111, CHIRPS=0.102) and for Karati (TRMM3B42v7=-0.49, CHIRPS=0.175). Both products failed to adequately simulate the peak flows but they did well in simulating the base flows. Comparison of rainfall hydrographs also showed discrepancies between the observed flows and the *in-situ* rainfall and this could have been a source of uncertainties evident with the model simulations. This performance was improved when CREST was ran on shorter windows of 4, 3 and 2 years. Reduced model ran windows improved *NS* efficiency values of each basin. However, the period 2005-2008 showed low performance and this could be attributed to the low data availability during this period. Validation results for Malewa (TRMM3B42v7=-0.87, CHIRPS=-0.096), Gilgil (TRMM3B42v7=-0.067, CHIRPS=-0.026) and for Karati (TRMM3B42v7=-1.702, CHIRPS=0.405). The poor validation results were attributed to scarce stream flow data available during the validation period of 2011 to 2013. The resulting monthly stream flows of model calibration were compared to the *in-situ* obtained stream flows and those used to prepare the lake balance model. CHIRPS had the best correlation to these stream flows and its monthly simulated values were used for the resilience test of the lake volume and level in response to river abstractions. The resilience test was also conducted during a wet and a dry year. Permit limits for legal permits were used as the basis for variation techniques. It was concluded that the lake is sensitive to reduction of stream flows with the enhanced sensitivity during the dry years. The level of sensitivity, moreover, was also found to be related to the lake levels. Low lake levels were more sensitive to reduced stream flows as a result of abstractions than when the lake levels are high. The research shows that a stream flow hydrological model can be used to simulate discharge in a catchment. However, the performance of the model in such instances largely depends on the credibility of the data used as input. This resilience test can be used by water authorities for allocation limits purposes.

Keywords: CREST model, Lake Naivasha, resilience, spatio-temporal, water allocation, TRMM3B42v7, CHIRPS, FEWSNET, scenario analysis.

ACKNOWLEDGEMENTS

DEDICATION

Dedicated to my dear parents, Mr. and Mrs. Ochieng.
Your prayers and words of encouragement kept me going always. To my siblings, family and in-laws, thank you for the moral support while I was away in The Netherlands.

To my fiancé, Florence Mpho Ogollah,
I do not have enough words to express how grateful I am for your support and prayers.
THANK YOU. It's finally over, for now!!

~I LOVE YOU ALL~

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AOI	-	Area of Interest
ASCII	-	American Standard Code for Information Interchange
CHIRPS	-	Climate Hazards Group Infrared Precipitation Data
CREST	-	Coupled Routing and Excess Storage
CSV	-	Comma Separated Value
DEM	-	Digital Elevation Model
FAC	-	Flow Accumulation
FAO	-	Food and Agriculture Organisation
FAR	-	Frequency Alarm Ration
FDR	-	Flow Direction
FEWSNET	-	Famine Early Warning System Network
FORTRAN	-	Formulae Translation
IDL	-	Interactive Data Language
ILWIS	-	Integrated Land and Water Information System
ISOD	-	*In-situ* and Online Data
ITCZ	-	Intertropical Convergence Zone
LNB	-	Lake Naivasha Basin
LULC	-	Land use and Land Cover
NASA	-	National Aeronautical and Space Administration
NetCDF	-	Network Common Data Form
PDB	-	Permit Data Base
PEST	-	Parameter Estimation
PM	-	Passive Microwave
POD	-	Probability of Detection
RGS	-	Regular Gauging Station
RMSE	-	Root Mean Square Error
SRE	-	Satellite retrieved Estimate
SRTM	-	Shuttle Radar Topographic Mission
TIFF	-	Tag Image File Format
TIR	-	Thermal Infrared
TRMM	-	Tropical Rainfall Measuring Mission
WAP	-	Water Allocation Plan
WAS	-	Water Abstraction Survey
WRMA	-	Water Resources Management Authority

1. INTRODUCTION

1.1. Research background

Fresh water is a vital component of the Earth system that enables life and sustains ecosystem health. In their book, Bengtsson et al. (2014) points out that despite the importance water holds on earth, it is still a relatively undervalued resource compared to other natural resources occurring in the world. Water has an effect on development of the economy, shaping the world's weather and climatic conditions and food security. In their book on the World's water and lands resources, (FAO, 2011) notes that water scarcity or low quality water has negative effects on agricultural, industrial and also domestic use of water since it is a key resources in carrying out economic development activities. The Stockholm International Water Institute (2005) details different ways of investing in water resources. To enhance economic growth in a country, water and its related services need to be integrated with other businesses since it is an economic good. This means understanding the water cycle and its components by decision makers and system managers.

Hydrology can be described as a science that studies the occurrence, dispersal, conveyance and assessment of water quantity and quality on earth (United States Geological Society, 2016). Studies have been conducted to ascertain how processes such as precipitation, evaporation and condensation affect the occurrence of water in different parts of the world. Shaw (1994) describes evapotranspiration process of the water cycle as a source of all fresh water in the world with the most contribution coming from the oceans. Freshwater resources have so far been maintained by natural cycles globally and regionally (Zhang & Luo, 2015). The water cycle is described as a continuous process where water is heated with energy from the sun, it evaporates, gets cooled and condensed in the atmosphere after which it returns to the earth in the form of precipitation. The quality of water determines its usability and fresh water is in high demand all over the world.

The world's fresh water reserves occur unevenly in terms of distribution. It is estimated that a number of countries in different parts of the world suffer from extremely high water stress levels in terms of withdrawal and availability ratios (Zhang & Luo, 2015). Pressure on water resources has increased tremendously due to a surge of the population, which is estimated to have increased more than fourfold over the past century and infrastructure developments (Bengtsson et al., 2014). In a bid to quantify fresh water reserves in the world, hydrologists have embarked on research to quantify different aspects of the hydrological cycle and this has been done at both local and meso scales. These studies have been done by the use of hydrological models. A hydrological model is a miniature characterization of the real world hydrological situation (Sorooshian et al., 2008). Models have produced in-depth information of the water situation in most countries around the world. FAO (2003) presented a report of the water situation in most parts of the world and it is seen that most countries in Africa and the Middle East are in the category of water-scarce regions. A report by the International Water Management Institute also shows that a quarter of the world's population are in localities characterised by water scarcity physically or economically. This can be seen in *Figure 1-1*.
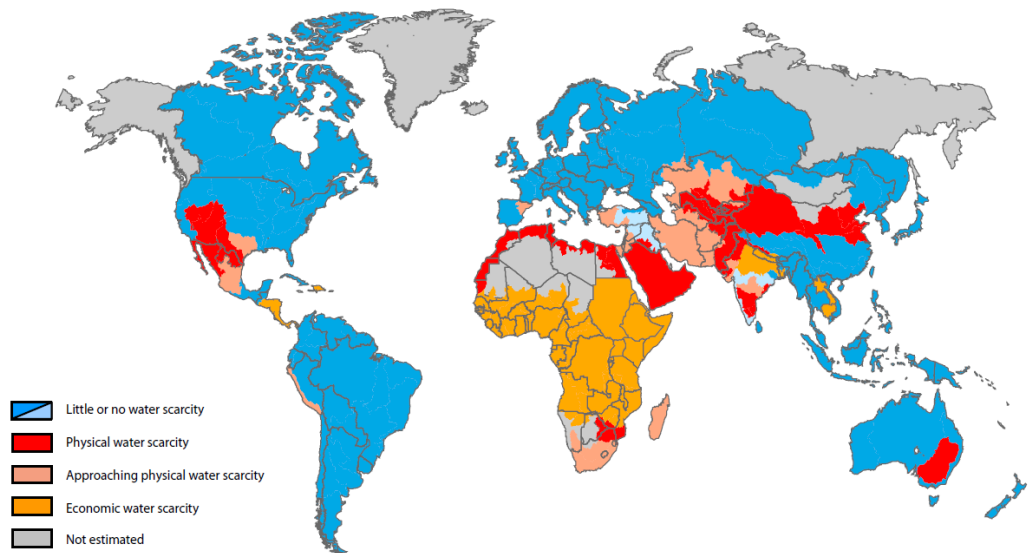
Figure 1-1: World map of economic or physical water scarcity, adapted from International Water Management Institute, (2006)*)*

Water scarcity leads to loss of lives and can result to conflicts. On smaller scales, water scarcity has escalated to demi-wars which have led to the loss of lives and properties in different regions around the world since different groups fight for control of the available water. One of the most conspicuous conflicts regarding water allocation between different stakeholders is the Nile river water allocation in Africa. The Nile river waters are shared amongst 10 countries. However, the most pronounced conflict is between Ethiopia and Egypt which serve as the source and outlet of the Blue Nile and the Nile respectively. 86% of the Nile waters is from highlands in Ethiopia, which makes a 95% contribution of water that flows into Egypt (Mason, 2004). In his publication, Mason (2004) notes that different users in the upstream and downstream of rivers have socio-political and economic impacts towards each other. The trend has been witnessed at meso and micro scales of water management in different catchments.

Lack of proper water allocation practices has also made water resources management challenging. Dourojeannni (2001) emphasizes that water allocation procedures have to meet the expectation of all stakeholders at the same time ensuring sustainability of the resources. Water resources situation of most basins are a concern to a variety of stakeholders since water serves various functions that each have a different set of requirements with respect to the quantity and quality of water. For instance, there are users who depend on the water for tourism income, fishing as a source of livelihood, transport, recreation and it also has an effect on sustainability of nature. Environmental flow of water ensures that living organisms are able to survive and ecosystems are maintained. Meeting these demands is a challenge in any basin around the world.

In solving these challenges faced in efficient water resource management, studies need to be done on the availability of water resources. However, these studies are hindered by a lack of sufficient *in-situ* spatial and temporal data that can be used to provide sound information. This is because weather stations are scarce and unevenly distributed across catchments. In hydrological studies, use of satellite data forms a platform for augmenting the spatially scarce distribution of observations (Khan et al., 2011). It is however noted that satellite products need to be assessed for applicability and for certain areas and spatial scales (Dinku et al., 2007).

In this research, an evaluation of the performance of a hydrological model in Lake Naivasha basin is evaluated based on *in-situ* data and uses of satellite-retrieved estimates (SRE) of rainfall and evapotranspiration as model inputs. The model chosen for this purpose is the Coupled Routing and Excess Storage (CREST), which is a grid based streamflow model developed by University of Oklahoma and NASA SERVIR (Wang et al., 2011). Water allocation information was retrieved from Water Resources Management Authority (WRMA) Lakes Nakuru-Naivasha Sub-region permit database (PDB).

1.2.   Problem statement and importance of this research

Classified as a Ramsar site, Lake Naivasha has a delicate ecosystem that needs special attention regarding water resources management. It is an important ecological and economic hub for Kenya and being a freshwater lake, it is suitable for irrigation agriculture (Van Oel et al., 2013). The basin hosts a robust horticultural, subsistence agriculture and geothermal power production industries. However, as an important source of livelihood for thousands of people and animals, competition for accessing clean water by different users has resulted in water related conflicts between different users in the basin. Low water quality and quantity has different effect to different users. For farmers, low quality water increases production cost since they have to treat the water before use and this increases business operating costs while for domestic users, it forces them to search for alternative sources which may be expensive or unavailable. Low quantity of water has led to conflicts between the farmers, the Maasai herders in the basin and the wildlife since they compete for the scarce resources.

Fayos (2002) noted that most water conflicts in the LNB result from deteriorating quantity and quality of both the rivers and lake waters. This kind of conflicts is characterised by downstream users being concerned about the decrease in water quantity and quality due to water resources development in the upstream. Upstream users, on the other hand, are concerned about downstream users hindering their water resources development initiatives (Mason, 2004). Providing water of acceptable quality and adequate quantity remains a challenge because of diverse uses of water. Seasonal variability of water quality and quantity has led to challenges in providing water for domestic use, agriculture, horticulture, recreation and nature in the Lake Naivasha basin. A possible solution to these challenges is the use of informed scientific strategies of water allocation such as the use of hydrological models that assist in developing scenarios identifying under which hydrological conditions the availability of water will be insufficient for sustaining all functions.

Several studies have been done to quantify the water balance of different areas in the world. Most of these studies have been done using hydrological models such as CREST, TOPMODEL, Hydrologiska Byråns Vattenbalansavdelning (HBV), MIKE-SHE and Soil and Water Assessment Tool (SWAT). In recent years, CREST and SWAT models have been used for studies in the LNB. Gathecha (2015) used the CREST model to reconstruct the streamflow into Lake Naivasha while Meins (2013) used SWAT to try to understand how spatial scales influence a model's performance. Other studies have been on groundwater and have used groundwater models such as MODFLOW. Several studies have proved that carrying out hydrological analyses are important for providing platforms for technical decision making to enable efficient use of available water in the LNB (Yihdego, Reta, & Becht, 2016). In Kenya, water allocation is placed under the authority of Water Resources Management Authority (WRMA).

WRMA was established by an act of Parliament in 2002. The authority uses the Water Act 2002 as a guiding tool for its operations. It is tasked with issuing permits based on the feasibility of the abstraction to the prevailing water quality and quantity conditions (Government of Kenya, 2002). WRMA has a permit database (PDB) where all water allocations in different basins are processed and stored. Permits are classified under four categories of A, B, C and D. The limits defining each class are dictated by the hydrological state of the river which varies from one river basin to another. For the Lake Naivasha basin, the surface water thresholds per permit class is as indicated in *Table 1-1*.

Table 1-1: Surface water thresholds per category of permit

Thresholds in m³/day			
Permit category A	Permit category B	Permit category C	Permit category D
Up to 20	>20 to 500	>500 to 1000	>1000

The Mean Annual Flow (MAF) in cubic meters per day can be used to estimate the amount of water in a river system on a daily basis. The naturalised flow duration curve can be used to set the Q_{80} threshold of a stream. The 80 represents the amount of flow occurring 80% of the time. A flow duration curve shows the probability of a certain amount of discharge to be equal or surpass a set threshold. The Q_{80} threshold

is used to distinguish between the normal and flood flows with normal flows being below this threshold. (Government of Kenya, 2006).

The main challenge with the water allocation process is defining limits for "normal" or "flood" water. These limits dictates the amount of water prioritized for each use during the wet and dry seasons (Government of Kenya, 2009). This is a challenge also experienced in the LNB since the amount of precipitation received in this basin varies spatially and throughout the year. Sufficient amount of data is needed to enable effective water allocation with regards to the flow in the rivers. In this research, discharge records accumulated over 50 years is used to create the flow duration curves of the different rivers in the basin in order to establish the Q_{80} threshold for each sub-catchment. It is noted that some uses such as ecological and domestic use require a consistent supply of water all year round while others like irrigation can be varied depending on the crop growth stage. *Table 1-2* shows the defining categories of water use.

Table 1-2: Category of water use activities, adapted from Government of Kenya, (2006)

Category	Description
A	Water use activity deemed by virtue of its scale to have a low risk of impacting the water resource. Applications in this category will be determined by Regional Offices
B	Water use activity deemed by virtue of its scale to have the potential to make a significant impact on the water resource. Permit applications in this category will be determined by Regional Offices.
C	Water use activity deemed by virtue of its scale to have a significant impact on the water resource. Permit applications in this category will be determined by Regional Offices in consultation with the Catchment Area Advisory Committees.
D	Water use activity which involves either two different catchment areas, or is of a large scale or complexity and which is deemed by virtue of its scale to have a measurable impact on the water resource. Permit applications in this category will be determined by Regional Offices in consultation with the Catchment Area Advisory Committees and approval by Authority Headquarters

Ecological demand has the highest priority for allocation. Domestic water demand has the highest demand when it comes to water allocation for use (Government of Kenya, 2009; WRMA, 2010). In severe drought, rationing of domestic water supplies may take place.

1.3   Research objective

In Kenya, and more so in the Lake Naivasha basin, water allocation is based on regular discharge measurements conducted on rivers. Even though this is an acceptable way of allocating water according to the water allocation guidelines (Government of Kenya, 2009), it is not sustainable for rapid decision making since neither seasonal nor inter-annual variations of discharge are taken into account in the permit details. Discharge from the rivers varies depending on whether the season is dry or wet. One way of improving the water allocation is by the use of hydrological models.

Hydrological models can be used to simulate the seasonal and inter-annual variation of streamflow or recharge. Models are able to simulate the amount of water available as streamflow during wet and dry seasons and can also be used to simulate the spatio-temporal variability of streamflow production across catchments. The variation of water availability across the two seasons can be used for scenario analysis and an acceptable means of varying the permit limits to maintain the minimum ecological flow of the Lake Naivasha Basin. The main objective of this research will be to quantify the resilience of the water system under various scenarios of water scarcity by using water abstraction records from streams and rivers in the basin. A hydrological model will be used to quantify the recharge into the basin and this will be used to propose ways of varying permits issued by WRMA in the basin.

The specific objectives of the research are to:

- Prepare spatially distributed rainfall and evaporation fields from available EO-based data products as primary inputs for the hydrological model and their validation/justification using available *in-situ* measurements.
- Determine a method of quantifying surface water abstractions from rivers in the basin using CREST model.
- Calibrate and validate the CREST model that includes the surface water abstraction module using stream flow measurements on a 10-day basis.
- Asses the resilience of the water system from various hydro-meteorological scenarios and levels of surface water abstraction.

1.4 Research questions

Research questions to be answered include:

1. How do earth observation products perform in characterizing the spatial-temporal distribution of rainfall and evapotranspiration in the basin when used as input for a hydrological model?
2. What is an effective method to account for surface water abstractions in the CREST model streamflow calculations?
3. Is the CREST model capable of adequately simulating the hydrological regime of the Lake Naivasha Basin with surface water abstractions included?
4. How resilient is the current water system with all the surface water abstractions and how can the resilience be further improved/optimised?

1.6 Thesis outline

Chapter 1 of the thesis gives an introduction to the research. The aim of the research and the objectives are expressed here. In chapter 2, a literature review is presented on applied methods and previous studies in the field. The uncertainties involved in the research are also analysed in this chapter. Chapter 3 describes the study area, the datasets and the choice of the hydrological models used in the study. The fourth chapter entails the methodologies employed to attain the research objectives such as those used to process the dataset and the results after the analysis process. The performance of the satellite products to the *in-situ* data is also analysed in this chapter. Chapter 5 covers the results of satellite products analysis, CREST model results and results of the resilience tests. While Chapter 6 summarises the conclusion and recommendations of the study.

2.   LITERATURE REVIEW

2.1.   Hydrology, water resources management and use of hydrological models

In the LNB, the water cycle plays an important part in ensuring availability of freshwater for different stakeholders of the water resources. However, there has been water scarcity in the basin during different seasons of the year. This is more pronounced in the recent years due to the ever increasing socio-economic pressure in the LNB (Van Oel et al., 2013). The situation has led to over-exploitation of the available water resources.

The basin experienced fluctuations in the amount of water flow to the lake and this almost led to the lake dying out between the years 1945 and 1955. (Becht & Harper, 2002). In their paper, Becht and Harper (2002) have identified the need for enhanced studies to better manage the waters of the basin so as not to experience what happened in the dry years since it could be worse in the future. Another study by Awange et al. (2013) estimated lake levels to have fallen at a rate of 10.2 cm/year which led to a shrinkage of the area at a rate of 1.02 km²/year in the period 2000-2010. In their analysis, the authors attributed these changes to human activities such as increased abstraction for floriculture and horticulture, evaporation, groundwater level fluctuations and climatic changes.

Management of anthropogenic and natural effects on the water resources is a vital part of ensuring existence of sustainable fresh water reserves in the world (Treut et al., 2007). Water resources management plays a vital role in ensuring water is available at an acceptable quality and quantity. Management practices have an effect at basin or river levels and are crucial in alleviating conflicts between users (Dourojeannni, 2001). Harper, Nic, Caroline, Ed, and Richard (2013) in their paper have analysed how different stakeholders are working together in the basin to improve the water situation. An integrated model for management of the water resources for the basin has been proposed by Odongo et al. (2014). The proposed model is based on coupling of socio-economic aspects and eco-hydrological processes of the basin. Socio-economic developments are known to have direct influences on hydrological processes on a basin (Loucks, Van Beek, Stedinger, Dijkman, & Villars, 2005).

One of the main goals of mankind is to understand the water cycle and control the processes involved in it by ensuring availability of freshwater in areas where it is needed (Rientjes, 2015). Freshwater is not as distributed as humans may wish even though it is plentiful (Loucks et al., 2005). It is estimated that eight out of every ten users of freshwater are located on the downstream of river basin (Vorosmarty et al., 2005).More studies need to be done to ascertain the effect of different flow regimes in a basin. The use of models provides a means of understanding the real conditions of catchments. In the LNB, some of the recent uses of hydrological models have been to reconstruct streamflow (Gathecha, 2015), assess effects of spatial scales when using hydrological models (Meins, 2013) and application of integrated model for demand and supply (Alfarra, 2004).

Different models have different capabilities based on the algorithms applied to achieve the main goal of water balance closure. These models are classified into two different categories of lumped and distributed models. Lumped models are those which average spatial characteristics over a model domain and represent these values using a single value while distributed models discretize the catchment characteristics within a model domain and consider their spatial variability. Distributed models employ use of data from diverse sources to enhance the simulation of hydrological components at different scales. *In-situ* datasets are the main inputs for most models but lack of or low accuracy of the available data have necessitated use of other data sources such as satellite retrieved estimates (SREs). Correct input data improves the

performance of models significantly as they work on the principle of "garbage in more garbage out" which stipulated that when a system is fed with bad data as input, it will definitely output undesired results.

2.2.   Representation of topography

Topographic representations are key inputs to hydrological models. They are represented using digital elevation models (DEM) which are used to retrieve geometry of river terrains and the catchment boundaries (Tarekegn, Haile, Rientjes, Reggiani, & Alkema, 2010). In hydrological models where the catchment is discretised into cells, the DEM is used to define flow patterns within a cell and the connectivity between different cells (Rientjes, 2015; Wang et al., 2011b). In this kind of representation, the runoff assumes the gradient within a cell and the same direction is also followed by subsurface flow of the cell.

The precision and accuracy of the DEMs differ depending on the source. Low cost DEM such as Advanced Space-borne Thermal Emission And Reflection radiometer (ASTER) and Radar Topographic Mission (SRTM) give coarse resolution of around 30 to 90 meters while costly DEMs from Light Detection and Ranging (LiDAR) technique has high precision and accuracy of up to 1 meter (Md Ali, Solomatine, & Di Baldassarre, 2015). Use of DEMs reduces the cost of carrying out ground surveys when topographical data is required. LNB falls under the 1^0S 36^0E tile and this was used for delineation basin.

2.3.   Use of *in-situ* and satellite data in research

Measuring of weather elements requires an elaborate network of measuring instruments. World Meteorological Organisation (WMO) gives guidelines of the network density of these instruments in different areas of the world (Plummer, Allsopp, & Lopez, 2003). *Table 2-1* show the minimum density of precipitation stations as proposed by WMO. It is also suggested that ten percent of these stations should be self-recording to record rainfall intensities.

Table 2-1: Minimum density of precipitation stations based on WMO provisions

Regions	Minimum density range (km²/gauge)
Temperate, Mediterranean and tropical zones	
Flat areas	600-900
Mountainous areas	100-250
Small mountainous islands (<20,000 km²)	25
Arid and polar zones	1,500-10,000

Lack of *in-situ* data for use in hydrological studies is one of the main hindrance to efficient water resources management and carrying out hydrological studies. This problem can be solved by combining globally available satellite products. Bhatti, Rientjes, Haile, Habib, and Verhoef (2016), advocate for use of satellite data in hydrological studies even though these datasets require correction. One of the main advantage of satellite based products is their area coverage. Their spatial and temporal resolutions are continuously increasing as more products are launched (Tufa Dinku, Asefa, Hilemariam, Grimes, & Connor, 2011).

Even though uncertainties are encountered while using satellite data, recent studies show that satellite images are increasingly offering reliable data for hydrological modelling and climatological studies (T. Dinku, Chidzambwa, Ceccato, Connor, & Ropelewski, 2008; Chen et al., 2014; Bhatti, Rientjes, Haile, Habib, & Verhoef, 2016; Khan et al., 2011; Miralles, Gash, Holmes, De Jeu, & Dolman, 2010; Knoche,

Fischer, Pohl, Krause, & Merz, 2014). These studies have shown different abilities of SRE's and a need for their calibration depending on the location where they are used.

There are a number of satellite retrieved products which include the Tropical Rainfall Measuring Mission (TRMM) data which uses thermal infrared and passive microwave sources to produce the final product (NASA, n.d.). Climate Hazard Group Infrared Precipitation with station data (CHIRPS) is also another high resolution product that uses the thermal infrared Cold Cloud Duration (CCD) and rain gauge data to produce a reliable product for use in estimating rainfall data (Funk et al., 2015). These products can be used to provide continuous rainfall values which can be used for running hydrological models.

2.4.   Uncertainties in hydrological modelling

Conceptual hydrological models help in understanding in-depth processes of the water cycle. It is known that use of hydrological models can be very complex. However, a more complex model does not necessarily mean better results. This is more so when the model is over-parameterised (Bergström, 2006). An over-parameterised model can be said to be one of high data demand which cannot be easily met with gauged values from standard hydrological and climatological data found in established stations. It also has to keep the calibration parameters to a minimum.

In modelling, uncertainties may arise from atmospheric forcing data used as input, incorrect model algorithms, wrong formulation of the model physics, inclusion of unknown parameters in a model structure, an oversimplified model, or uncertainties introduced when initialising the model. These uncertainties may eventually lead to variations between the observed and simulated values. The systematic differences observed in model simulations is referred to as bias. The bias can have a multiplying, summative or non-linear effect on the simulation results of a model. They cannot be eliminated by averaging a series of values.

Uncertainties are functions of scale and a hydrologist's knowledge of a particular spatial scale influences his understanding of the model. Beven. (2001), discusses challenges modelers' face when developing and applying models such as (1) Non-linearity of hydrological systems, (2) influences of spatial scales in modelling, (3) differences between watersheds, (4) Equifinality of model parameter sets and (5) uncertainty of parameters and the model selected. The definition of input parameters influencing the performance have to be carefully assessed before running a hydrological model. It is important that uncertainties are kept at a minimum and to reduce residual errors and eventually improve model performance.

The effect of these bias cannot be eliminated by averaging a series of values. (Smith, Arkin, Bates, & Huffman, 2006). It is important to analyse the effects of these bias in the long run so as to quantify their effect to the final performance of the application of the retrieved data.

2.5.   Model selection

In choosing a model for use in a hydrological study, the performance of the model based on previous studies should be considered. A model which better simulates the observed characteristics with little variations to field values should be considered (Rientjes, 2015). Hydrological models are either distributed or lumped. Distributed models take into account the spatial disparity of the model domain by taking into account differences in land use land cover, differences in meteorological forcing and topography while the lumped models average values in model domain and use a single value to represent different values of the same phenomena.

In recent years, a water balance of the Malewa sub-catchment of the LNB was studied by Meins (2013). The model was used to evaluate the effect of scales in application of hydrological models. The model used was the Soil Water Assessment Tool (SWAT). According to the author, one shortcoming of the model is the data requirements. SWAT model requires a lot of parameters which are not available for the study domain and the calibration exercises was also noted to be time intensive. This had an influence on the performance of the model.

Gathecha (2015) used the CREST model to reconstruct stream flow in the LNB. CREST model was also used in Nzoia basin which drains into Lake Victoria to simulate the runoff in the area and it performed exceptionally well in this East African basin (Wang et al., 2011). CREST model has also been used in Usangu catchment in Tanzania to assess its performance when satellite retrieved weather data is used (Mbaga, 2015). In the CREST model, runoff generation and routing are coupled hence giving a convincing interaction of the lower layers of the atmospheric boundaries, the terrestrial surfaces and water in the subsurface and this can be applied at catchment, regional or global scales. These are the distinguishing characteristics of the CREST model compared to other distributed models.

A lake water balance model proposed by Van Oel et al. (2013) was used to determine the amount of stream flow required to maintain the desired lake levels. Lake levels are influenced by the precipitation received, amount of water that flows in as runoff from rivers, the rate of evaporation of lake water, the amount of abstractions and the amount of outflow experienced in the lake.

3. STUDY AREA, DATASETS, AND MODEL

3.1. Study area and background

3.1.1. Location

Lake Naivasha is a fresh water body situated in the Kenyan Rift Valley between longitudes 0^045' S, 36^020' E at an altitude of 1890 m.a.s.l. The lake covers an area of approximately 140 square kilometres while the basin covers an estimated area of 2,500 square kilometres. In Kenya, it is second to Lake Victoria in terms of size of a freshwater body. Its catchment is shared between Nyandarua County to the North-West, Nakuru County to the North and North-East and Narok County to the south-West. It is approximately 80km from Nairobi to the North-West. The basin has three main sub-catchments namely; Malewa, Gilgil and Karati. The Malewa and Gilgil are permanent rivers while Karati sub-catchment only contribute water during the high rainfall seasons (Becht & Harper, 2002). Marmanet River sub-catchment from the Mau Escarpment which ends before it enters the lake as it recharges the Ndabibi plains. *Figure 3-1* shows the location of LNB in Kenya (*Source: United Nations Cartographic section*), elevation and Regular gauging stations at the outlet of the three main sub-basins.
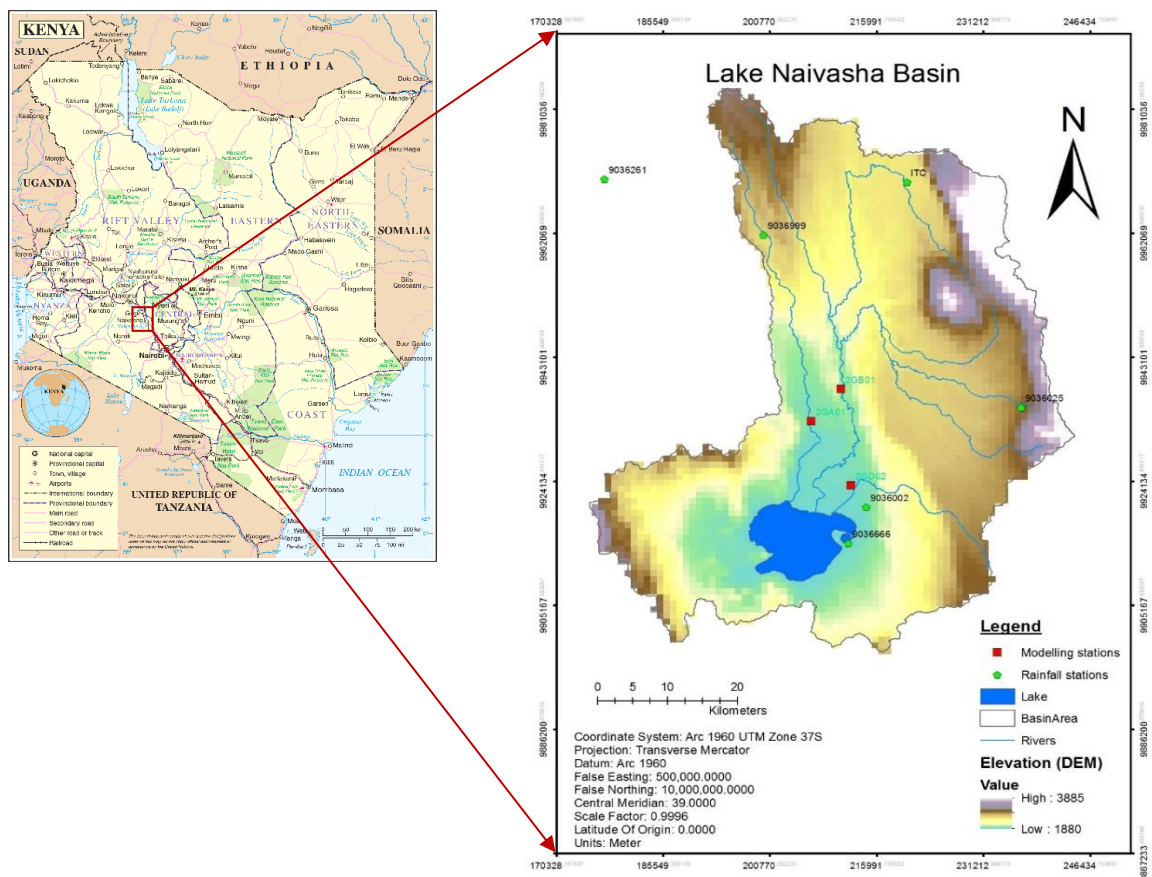
Figure 3-1: Lake Naivasha Basin map showing its location in Kenya, rainfall station location and sub-basin outlet discharge stations.

It is noted that apart from Lake Naivasha, the Kenyan Rift Valley has six other lakes from the north to the south. These include Lakes Turkana, Bogoria, Baringo, Nakuru, Magadi, and Elementaita. Lake Naivasha elevation makes it the highest in the Rift Valley lakes in East Africa (Armstrong, 2002).

3.1.2. Importance of study area and water use

Lake Naivasha was designated as a Ramsar site (No. 724) on 10th April, 1995. Within its locality there is the crater lake (Sonachi), a river delta and Lake Oloiden which is blue-green algae dominated and has soda tolerant plants (The Ramsar Convention on Wetlands, 2014). The basin supports complex terrestrial vegetation, littoral and riparian plants such as papyrus. These plants offer breeding grounds to both migrant and resident species of birds. The riparian of the lake is also a home to wildlife including buffaloes, waterbucks and hippopotamus.

Socio-economically, the lake is one of the most important sources of income to the country and more so to the people who depend on it. Commercially, the lake hosts some of the biggest horticultural and floricultural greenhouses in Kenya. Exports from the flower industry in Kenya amounted to Kenyan shillings 62.92 billion (673 million USD) at 0.96% of the Country's Gross domestic product (Global Finance, 2017) for the year 2015 with 75% of these exports coming from the LNB (Kenya Flower Council, 2016; The Ramsar Convention on Wetlands, 2014). The lake also supports fishing, pastoralism and is also a tourist site with a number of tourist hotels around it.

3.1.3. Topography of the basin

The maximum altitude of the basin is 3990 meter above sea level (m.a.s.l.). This is the North-Eastern side of the Aberdare range. The lowest point of the basin is 1980 m.a.s.l. This point is at the floor of the Rift Valley (Gathecha, 2015). The Aberdare ranges border the basin on the eastern side while eastern border of the basin is capped by the Mau escarpment. The Mount Longonot and Eburu Hills border the basin on the Southern and Northern side respectively (Becht & Harper, 2002).

In representing the topography of the study area, three elevation zones were established. The zones include the highlands which range from 2827 to 3558 m.a.s.l., the plateau which has an elevation range of 2243 to 2827 m.a.s.l. and the low land areas whose elevation range from 1880 to 2243 m.a.s.l. These zones show the source of the rivers (highlands) and the areas where the rivers flow over a relatively flat area (plateau) and the lake where the river drains. *Figure 3-3* shows a cross-section of LNB and the major features while *Figure 3-4* shows the LNB map with the 3 elevation zones and where the cross-section was derived.

Figure 3-2: Cross section of LNB
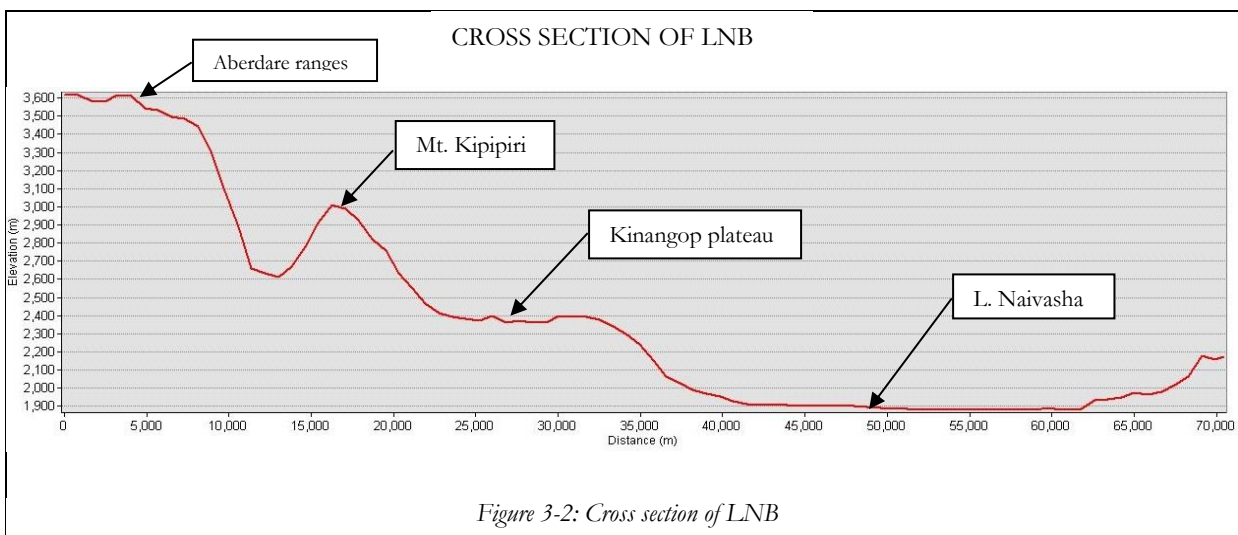
Figure 3-3: LNB elevation zones and cross section line
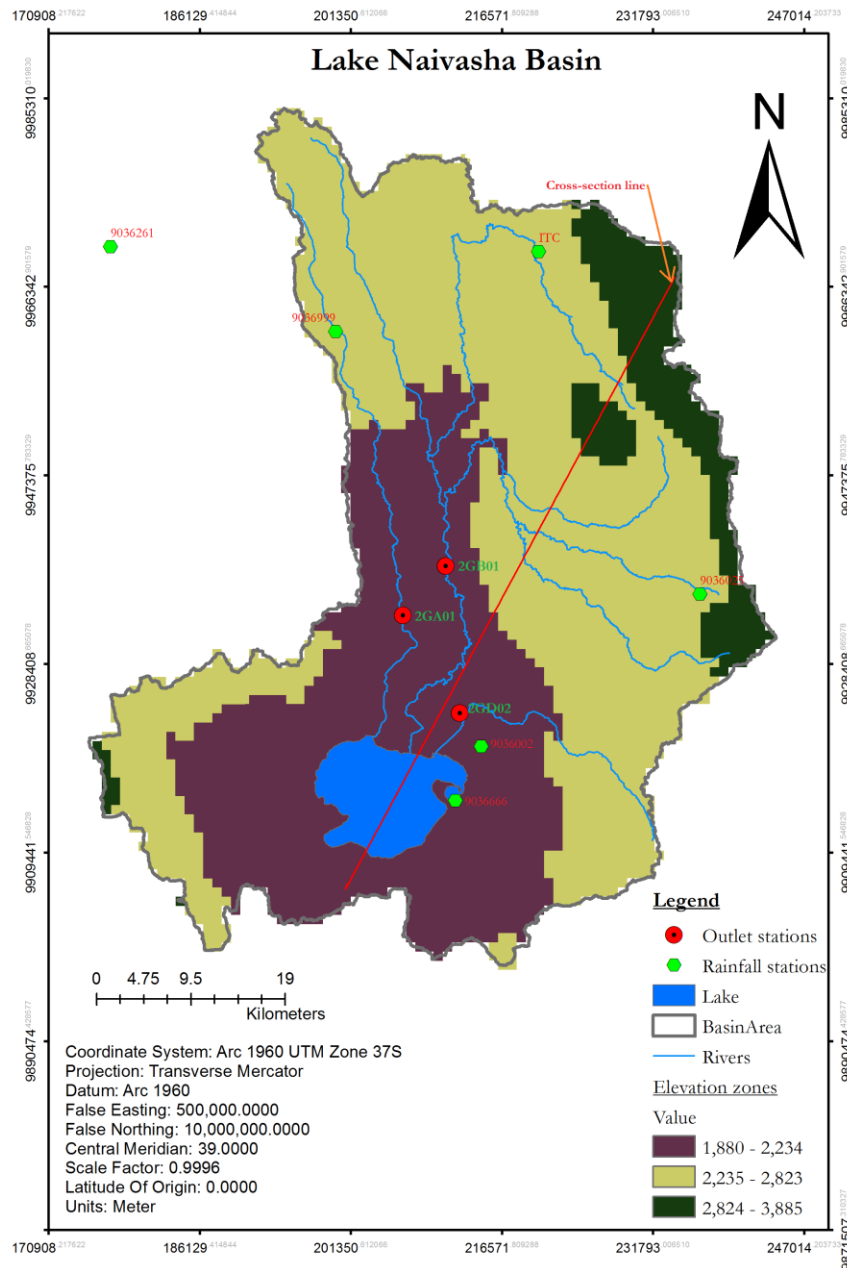
Figure 3-4 LNB map with elevation zones and the line where cross section was derived:

3.1.4. Climate of the basin

The region is within the Intertropical Convergence Zone (ITCZ) and this has a strong influence on wind and rainfall patterns of an area. Mt Kenya and the Aberdare ranges capture most of the moist monsoon winds which make LNB to fall on the rain shadow side. The area has a bi-modal rainfall regime with long rains being experienced between April and June while the short rains are experienced during October and November months (Becht, Odada, & Higgins, 2005). The long rains refers to periods of low intensity rainfall that falls over a long duration of time after the hot months of February and March while the short rains are of heavy intensity and short duration and normally occurs between the months of November and December. *Figure 3-5* shows the average monthly rainfall received in the basin.

The large difference in elevation of the LNB (1880-3990 m.a.s.l.) results to high variations in the amount of rainfall experienced. Naivasha town receives a rainfall amount of approximately 600 mm/year while the slopes of the Aberdare ranges receives precipitation in the amounts of 1,700 mm/year. The Kinangop plateau receives annual rainfall in the range of 1,000-1,300 mm. The evaporation of Lake Naivasha has been estimated at 5.95 mm/day (Ahmed, 1999). The Malewa sub-catchment is the biggest contributor of water to the lake at 79% compared to the Gilgil (19%). Karati River and other areas to the south of the lake do not contribute significantly as they are ephemeral and flow roughly two months in a year with most of them only contributing in terms of flash floods. Mean annual temperatures vary between 16^0c to 25^0c in the north western part of the basin and on the lake respectively. The daily temperatures however vary from 5^0c to 25^0c.
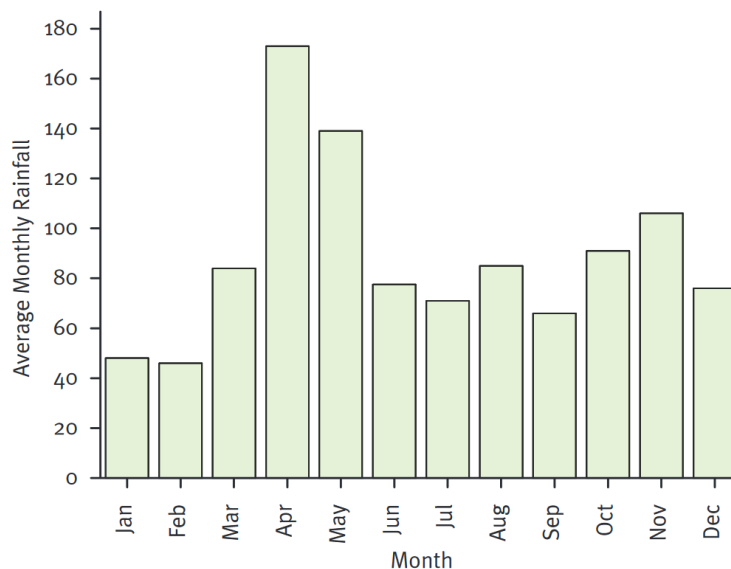
Figure 3-5 Average monthly rainfall distribution for Lake Naivasha region over a 60 year period. (Adapted from Becht et al., 2005)

3.1.5. Soil types found in the basin

The lake basin lies in a volcanic area with Mt. Longonot being a dormant Volcano on the southern border of the lake. This has influenced on the development of the soils found on the floor of the basin. The major soil type found around the lake are volcanic soils which are composed of pumice layers and pumice grains.(Becht et al., 2005). The interaction of the water with these volcanic soils has led to occurrence of zeolite. Pumice is highly permeable and this reduces the water holding capacity of the basin.

Volcanic rocks are highly saline and this influences the quality of the irrigation water in the basin. This quality also influences the use of groundwater for irrigation. Water from the catchments that flow through rivers Malewa and Gilgil Rivers are of good quality but due anthropogenic activities and interaction with other rock formations the quality changes as it flows into the lake. The soil type influences the infiltration rates in the basin.

3.1.6. Land use and land cover of the Basin

Land use and land cover (LULC) varies in the basin and this is more so because of the type of climate in the area and the distribution of population. Odongo et al. (2014) notes that human activities give influence the LULC in an area and this is mostly attributed to an increase in population. These activities have an influence on the quality and quantity of the water resources in the basin. It is estimated that LULC transformations in the LNB between the years 1961-1985 led to increase of runoff by 32% compared to the runoff for the period 1986-2010. The basin land cover has transformed over the years from forested

areas to human settlements and clearing of land for agricultural purposes. This led to changes in the lake volumes and the sediment yields observed in the basin.

3.2. Datasets

3.2.1. *In-situ* data

The ground data used for this research was retrieved from different stations in the basin. The ground stations were used as reference to the data retrieved from the satellite values. There are a total of 29 regular gauging stations (RGS) for quantification of discharge but only 15 have reliable data. Reliable stations refers to those that had data between the years 1960–2010. The discharge measurements of these stations was based on rating curves prepared from the staff gauge recordings of the height of water (Meins, 2013). The water level measurements are taken on a daily basis while the stream flows used to derive the rating curves were calculated using the velocity-area technique.

The discharge measurements used for calibration of the CREST model was obtained from the outlet stations of each sub-basins. Rainfall and evaporation data was obtained from weather stations located at both the upper and lower zones of the basin. The rainfall data is collected using both manual gauges and the tipping bucket rain gauges. Consistent evaporation data was only available for the Lake station. This station uses the open pan evaporation estimation method to collect evaporation rates of the lake.

The lake levels were obtained from a lake station located at the shores of Lake Naivasha. The lake level measurements cover the period from the year 1965 to 2010. The same data was used in preparing the lake model used in this study as a reference to ascertain the effect of surface water abstractions on the inflow into Lake Naivasha. Most of the RGS stations have staff gauges with a few having automatic recorders (divers). In the time series data used from some stations, some stations have recordings of both manual and automatic recordings. It was also noted there is difference of time-step recording of the observation. These datasets were aggregated to give daily records for all the stations used as CREST model inputs. A schematic drawing of all the stations is as shown in *Figure 3-6*.
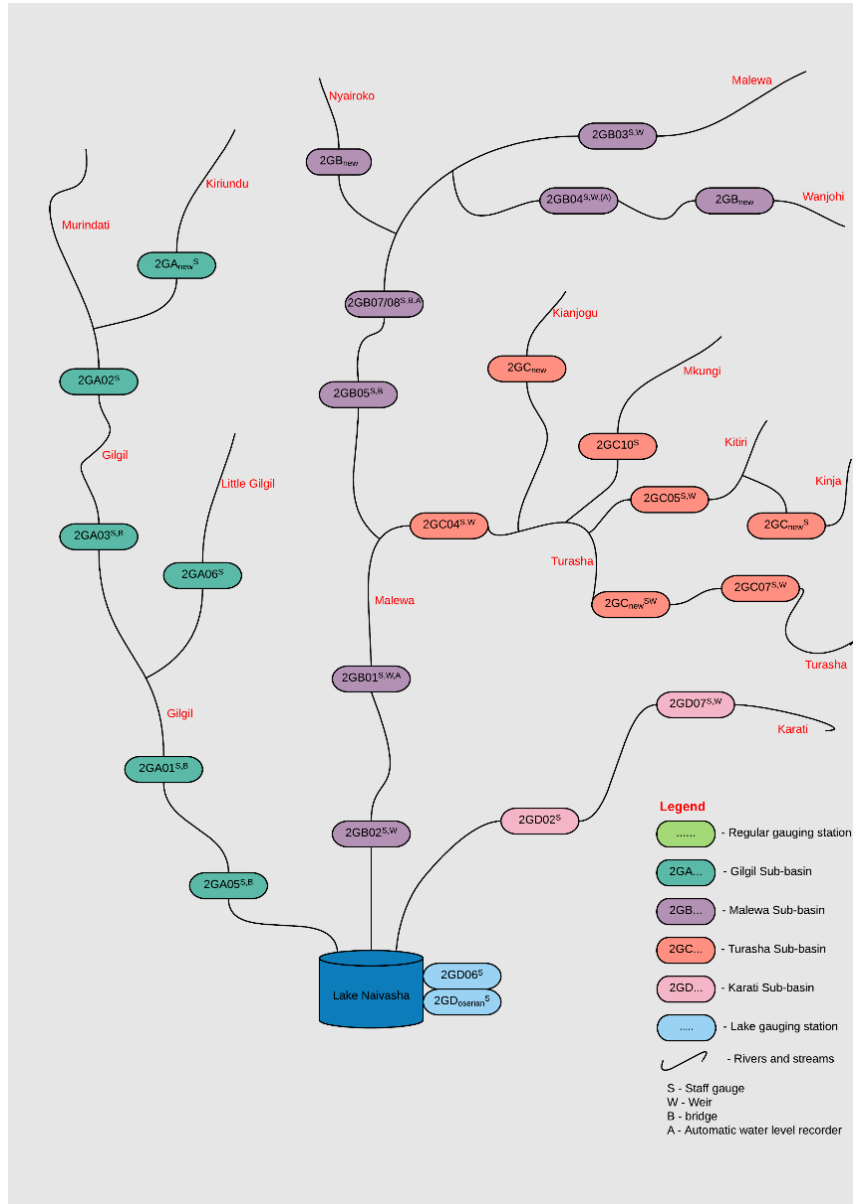
Figure 3-6: Schematic drawing of gauging stations and the devices used for measurements in Lake Naivasha basin

In naming of the RGS stations, the prefix "2G" represents the catchments in this case the Lake Naivasha catchment. The Prefix "A, B, C or D" represents the sub-catchments and the last two numbers refer to the station's number. The station's number indicates the sequence of the establishment of the station and not necessarily the position of the station towards the outlet of the river. Initial stations were established from the outlet towards the headwaters on the upper side of the catchment.

The first station from the outlet was given the prefix "01" and this changed as the stations were established upstream. With time, the water authorities realised that more stations needed to be established in between the already established stations and this caused the naming sequence to change since older stations maintained the same name. *Figure 3-7* shows a map of the location of all the regular gauging stations in the LNB. It can be seen from *Figure 3-7* there are a number of streams without gauges. This is because most of the ungauged streams are either ephemeral or they have very low discharge.
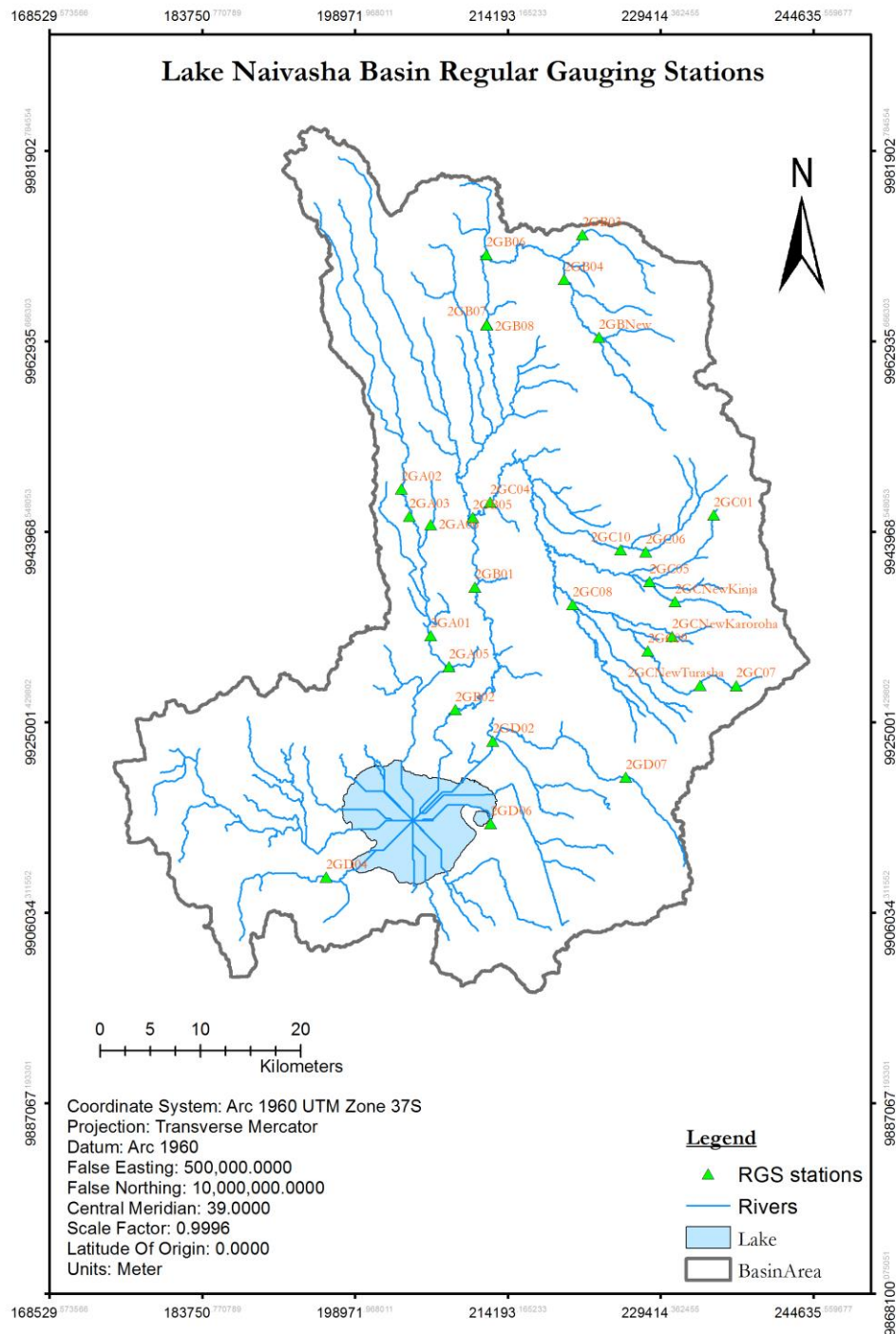
Figure 3-7: Map showing location of RGS stations in LNB.

For this research, a total of six rainfall stations were available for assessing the performance of the satellite retrieved rainfall products. These stations are located at different parts of the basin. In choosing these stations, data availability and representative weight of each station to the total area of the basin was considered. The spatial support provided by the rainfall station was determined via Thiessen polygon. Thiessen polygons can be used to analyse the spatial weight of a rainfall station in the basin. When using this method of rainfall interpolation, a single rainfall value is assigned to any point within the boundary of a polygon. *Figure 3-8* shows the location of rainfall stations and the resulting Thiessen polygons

Stations used to assess performance of SREs
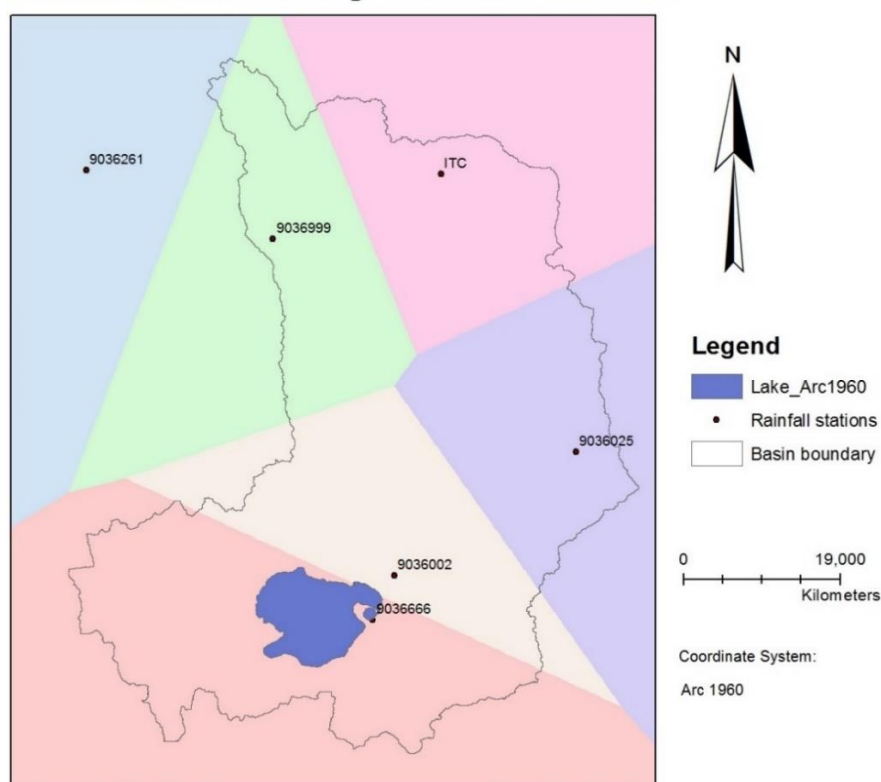
Figure 3-8: Rainfall station locations and resulting thiessen polygons.

The coordinates of the stations and their elevations is as shown in *Table 3-1*. The selection of these stations was based on how representative they are to the whole catchment and the amount of data available from each station.

Table 3-1: Used rainfall station locations and their elevations

Station name	Station ID	X-Coord	Y-Coord	Elevation (m.a.s.l)
Naivasha D.O.	9036002	214500	9920200	1923
N. Kinangop Forest Station	9036025	236545.6	9935511.5	2617
Gilgil Kwetu Farm	9036999	199826	9961909	2391
Kijabe Farm	9036666	211924.8	9914724.7	1907
New Holland Flowers	ITC	220260	9969946	2400
Nakuru Meteorological Station	9036261	177161.6	9970453.1	1910

In previous studies conducted in the basin, it was noted that stream flow data gaps were more than 75% during the period 2001 to 2010 (Meins, 2013), while Gathecha (2015) estimated the gaps for the three main sub-catchments amounting to average of 79.3%. The gaps were mostly filled using interpolation by use of rating curves produced with the available data but this introduced uncertainties since interpolation introduce overestimations or underestimations of the real scenario. The gaps in the sub-catchments stream flows can be summarised as shown in table for the period 2001 to 2010 (total number of days being 3651).

Table 3-2: Gap analysis (2001-2010)

	Gaps	Total gaps in % of 3651 days
Malewa	2935	0.80
Gilgil	3012	0.82
Karati	2768	0.76

Discharge measurements are used calibration and validation of the CREST model. The stations used in this research were selected based on their location in the basin and the quantity of data available. The routing of the sub-basin formed the basis for selection of the stations used in the study. Basin outlet discharge stations were used for CREST model input in cases where the outlet stations had gaps, the data from upstream stations were used to fill the data gap. From the digital elevation model (DEM) of the basin, three sub-basins can be distinguished and the accompanying stations at the outlet are as shown in *Table 3-3*.

Table 3-3: Location of discharge stations and their elevations as used in the study

Station	Sub-basin	River	Location (WGS84)		Elevation
2GA01	Gilgil	Gilgil	36.362856	-0.602006	1920
2GB01	Malewa	Malewa	36.403195	-0.558904	1951
2GB05	Malewa	Malewa	36.401537	-0.495355	1987
2GC04	Malewa	Turasha	36.41701	-0.480988	2000
2GD02	Karati	Karati	36.419773	-0.697605	1896

3.2.2. Earth Observation based meteorological forcing.

In hydrological modelling, use of high spatio-temporal precipitation data is highly desired. Existence of a reliable rain gauge network is not available in most areas and practically non-existent in the oceans (Joyce, Janowiak, Arkin, & Xie, 2004). In their paper, Joyce et al. (2004) notes that in many areas around the world, rain gauge data is available in time durations of 6 hours or daily. Remotely sensed rainfall estimates are available on a 3 hour or less intervals and this makes them most desirable for hydrological modelling. Another advantage of these data is their spatial resolution which is higher compared to any rain gauge station. The performance of the data needs to be verified before application. In this study, TRMM3B42v7 and CHIRPS rainfall products were used.

3.2.2.1. TRMM data

The Tropical Rainfall Measuring Mission (TRMM) data is the product of a joint program between the Japanese Space agency and NASA. The satellite was launched in 1997 and continued to give reliable data until April, 2015 ("TRMM Mission Overview|Precipitation Measurement Missions," n.d.). The TRMM3B42-v7 data is available at a resolution of approximately 25km (0.25^0). It is a rainfall product that combines both thermal Infrared data collected from different geo-stationary satellites (TIR) and passive microwave (PM) data from a number of sensors. Sources of the PM data used in the TRMM3B42 data include the Special Sensor Microwave Image (SSM/I), TRMM Microwave Imager (TMI), the Advanced Microwave Sounding Radiometer-Earth Observing System (AMSR-E) and the Advanced Microwave Sounding Unit (AMSU). It blends algorithms from this sensors to give a final rainfall product.

TRMM rainfall has several products. These products depend on the level of processing of input data and the outputs have different resolutions. The 3B42 product has the highest resolution of all the TRMM products at 0.25^0 x 0.25^0.(National Aeronotics and Space Administration, 2016). In their paper, Prakash, Mahesh, and Gairola (2013) compared the performance of v6 and v7 of 3B42 products over oceanic rain gauges and conclude that v7 performs better in terms of Root Mean Square Error (RMSE) and bias performance in high precipitation events. In the LNB, Gathecha (2015), compared the performance of TRMM 3B42_v7 rainfall product to CMORPH data where the former performed better.

The TRMM data was downloaded via: http://mirador.gsfc.nasa.gov/cgi-bin/mirador/homepageAlt.pl?keyword=3B42. A subset for the study Area of Interest (AOI) was made and converted to ASCII format for input into the CREST model. A code for this task was already developed for Gathecha (2015).

3.2.2.2. CHIRPS data

Climate Hazard Group Infrared Precipitation with station data (CHIRPS) spans the 50^0S-50^0N longitudes and has 0.05^0 resolution images. In creating CHIRPS, data from satellites are blended with data from rain gauges to create rainfall time series (Climate Hazard Group, 2015). Most derived rainfall SREs have long latency and short period of availability of data or vice versa. CHIRPS data was created to reduce the latency and increase the data range availability that is not met by most SREs. CHIRPS is available for periods of almost more than 30 years (Funk et al., 2015). In their study, Funk et al. (2015) notes that land surface models can be used in conjunction with CHIRPS to produce precipitation forecasts in data scarce areas.

In Africa, a comparison of the performance of CHIRPS, TAMSAT and FEWSNET RFE rainfall was done (Toté et al., 2015). In this study, it was noted that CHIRPS SRE is more effective under intense rainfall experiences compared to periods of light rain. The author explains the performance test conducted in Mozambique show that CHIRPS outperformed TAMSAT and FEWSNET RFE. In the LNB, Njuki (2016) used CHIRPS rainfall study in his and concluded that the product performance increases when used over shorter time periods compare to longer periods and the difference was attributed to the accumulation of the bias over a longer time period.

The CHIRPS data was downloaded in batch from an ftp site. The ftp link used is ftp://chg-ftpout.geog.ucsb.edu/pub/org/products/CHIRPS-latest/africa_daily/tifs/p05/yyyy. The CHIRPS data downloaded from this site is in tiff format and subsets for our AOI were produced on a daily basis. This data was then converted to ASCII format for CREST model input.

3.2.2.3. FEWSNET Potential Evapotranspiration data

Evapotranspiration data from the basin was obtained from the Famine Early Warning Systems Network (FEWSNET). FEWSNET was developed to aid in prediction of famine in West and East Africa by USAID (FEWS-NET, 2006). FEWSNET potential evapotranspiration is available on a daily basis. Potential evapotranspiration occurs when an area has sufficient water. The FEWSNET PET is calculated at a resolution of 1.0 degree from 6-hour meteorological model output. The calculation uses the Penman-Monteith equation. The Penmann-Monteith equation is shown in equation (Eq 3.1).

$$\lambda ET_0 = \frac{\Delta(R_n - G) + [86400\frac{\rho_a C_p (e_s^0 - e_a)}{r_{av}}}{\Delta + \gamma \left(1 + \frac{r_s}{r_{av}}\right)} \qquad \text{(Eq 3.1)}$$

where ρ_a is the air density (kgm-3), C_p is the specified heat of dry air, e_s^0 refers to the mean saturated vapor pressure (kPa) computed as the mean, e^o at the daily maximum and minimum air temperature (^{0}C), r_{av} is the bulk surface aerodynamic resistance for water vapor (s m-1), e_a refers to mean daily ambient vapor pressure (kPa), r_s refers to the Canopy surface resistance (s m-1), R_n is the Net radiation (W m2), G is the Ground heat flux (W m2) and γ refers to Psychrometric constant (kPa °C−1)

The FEWSNET data was download using the *in-situ* and Online Data (ISOD) toolbox in ILWIS. This is platform developed to aid in downloading data near real-time and real time data from different databases around the world. It can be used when the downloading machine is connected to the internet (Maathuis et al., 2014). The data is in daily estimates and was prepared for input into CREST model as ASCII files.

3.3. Abstraction data

Water resources allocation in the LNB is done by WRMA. The authority maintains a permit database (PDB) which has records of all the legal abstractions in the basin. A client has to apply for a permit when he/she intends to draw water from a natural resource. The authority issues permits to control the abstraction of water from both surface and groundwater resources to ensure sustainability and equity in utilisation based on demand and availability.

Water resources management authority (WRMA) has conducted abstraction surveys in the LNB. These surveys were done to quantify the number of abstraction points in the basin and the quantity of water being abstracted from both surface and ground water resources. The number of compliant abstractors is also quantified during such surveys since there are also illegal abstractors and those who abstract beyond their permit's daily limits. De Jong (2011) produced a report of an abstraction survey for the LNB that was conducted in May, 2010.

The main goal of the abstraction survey was to use the abstraction data for preparing a LNB water allocation plan (WAP). From the abstraction survey, it was estimated that the surface water resources are over abstracted by 480% and this may be one of the main sources of conflicts in the basin since people are competing for the scarce resource. This percentage included both the legal and illegal abstractions that were recorded during the exercise. In this study, abstraction data from the WRMA database will be used to test the resilience of the Lake Naivasha volume to the effects of abstracting water along the three main sub-catchments of the lake. This will be done by deducting the surface water abstractions from the model simulated flows of each sub-basin. Only legal abstractions will be considered for the study since their limits can be established.

3.4. CREST model design and purpose

In this research, the CREST version 2.1 coded in FORTRAN was used. It is a distributed model where the user can define the resolution of the grid cells (Wang et al., 2011). This multi-scale use of the model enables it to simulate the occurrence of water at different areas and time spans (Spatio-temporal). The model is also able to simulate storages and energy fluxes at different scales. The CREST model uses a variable infiltration curve for distinguishing precipitation that is let to flow as runoff and the amount left for infiltration. The infiltration curve is affected by LULC characteristics of the cell i.e. Soil type and vegetation cover. For example when the soil is porous, there is more precipitation converted to infiltration than when the soil is not porous. There is also more infiltration in densely vegetated areas since the water is held by the plants and flows steadily into the soil layer. CREST uses two-linear reservoirs for runoff generation. Routing is achieved by the use of the fully distributed linear reservoir routing (FDLRR).

In the model design, a cell has four excess storage reservoirs which represent vegetation storage and three sub-surface layers. Two more separate linear storage reservoirs representing sub-grid overland cell routing and sub-surface runoff are also included. *Figure 3-9* shows a schematic presentation of the structure of the CREST model and its core components.

(a) *Vertical profile of a cell including rainfall-runoff generation, evpotraspiration, sub-grid cell routing and feedbacks from routing*

(b) *Variable infiltration curve of a cell*

(c) *Plan view of cell and flow directions*

(d) *Vertical profile along several cell including sub-grid cell routing, downstream routing and subsurface runoff redistribution from a celll to its downstream cells.*

Figure 3-9. An overview of CREST model core components (Adapted from (Wang et al., 2011))

In the **canopy interception**, the capacity of water the canopy can hold is dictated by the land cover. The Canopy Interception Capacity (CIC) equation can be summarised as follows:

$$CIC = k_c \cdot d \cdot LAI \qquad \text{(Eq 3.2)}$$

where K_c is the coefficient of land cover CIC, d is vegetation coverage (m²) and LAI is the Leaf area index. Information about d and LAI can be retrieved form Land cover maps.

Runoff is only generated when the canopy is full and consequently the excess precipitation reaches the soil. This is quantified using (Eq 3.3.

$$P_{soil} = P - (CIC - CI) \qquad \text{(Eq 3.3)}$$

where P_{soil} is the precipitation (mm) reaching the soil surface, CI is depth of water intercepted by the vegetation canopy in millimetres.

The **variable infiltration curve** (tension water curve) of each cell defines the amount of precipitation reaching the soil is converted into either runoff or infiltration. This infiltration is governed by soil characteristics and land cover of the area. (Eq 3.4 shows the curve algorithm.

$$i = i_m \left[1 - (1 - A)^{1/b_i}\right] \qquad \text{(Eq 3.4)}$$

where i is the point infiltration capacity (mm/h), i_m is maximum infiltration capacity of a cell, A is fractional area (m²) of the cell corresponding to i and b_i refers to the exponential of the curve.

The three soil layers maximum water capacity (W_m) influences the maximum infiltration capacity (i_m) of a cell and are defined by the following (Eq 3.5) and (Eq 3.6)

$$i_m = W_m(1 + b_i)$$ (Eq 3.5)

$$W_m = W_{m1} + W_{m2} + W_{m3}$$ (Eq 3.6)

where W_1, W_2, and W_3 refers to mean soil water depth of each cell layer. Water infiltrates the layers sequentially with the upper layer first followed by the middle and lower layers. The available infiltration water can be quantified as shown in (Eq 3.7):

$$\text{For } i + P_{soil} \geq i_m \quad I = (W_m - W)$$

$$\text{For } i + P_{soil} < i_m$$

$$I = (W_m - W) + \left[1 - \frac{i + P_{soil}}{i_m}\right]^{1+b_i}$$ (Eq 3.7)

In **runoff generation**, excess rain can be obtained using (Eq 3.8) and the rate of infiltration at the top layer of the soil determines whether the rain is converted to overland excess rain (R_o) or interflow (R_I). These situations can be described with equation (Eq 3.9).

$$R = P_{soil} - I$$ (Eq 3.8)

$$\text{For } P_{soil} > K \quad R_I = K\frac{R}{P_{soil}} \quad R_o = R - R_I$$

$$\text{For } P_{soil} \leq K \quad R_I = R \quad R_o = 0$$ (Eq 3.9)

where K is the saturated hydraulic conductivity of the soil, R_o and R_I are discussed in the routing section.

In calculating **evapotranspiration**, the availability of data in the study area necessitates different ways of estimating evapotranspiration. Satellite retrieved potential evapotranspiration can be used in areas with no data. For CREST model, the FEWS NET PET is used.

In the CREST model, evapotranspiration firsts removes water from the top soil reservoir and the other reservoirs are followed after this layer runs dry. These situation can be summarised in (Eq 3.10).

$$\text{For } CI > E_p \quad E_c = E_p$$ (Eq 3.10)

$$\text{For } CI \leq E_p \quad E_c = CI$$

where E_c refers to water lost from canopy via evapotranspiration. The potential evapotranspiration applied in the upper layer of the soil (E_{P1}) is as shown in (Eq 3.11).

$$E_{p1} = E_p - E_c$$ (Eq 3.11)

E_p is initiated when the as storage tank is depleted and is achieved when the canopy is totally devoid of water. Soil water loss rate from the first layer (Layer 1) can be defined as:

$$\text{For } W_1 > E_{p1} \quad E_{s1} = E_{p1} \tag{Eq 3.12}$$

$$\text{For } W_1 \leq E_{p1} \quad E_{s1} = W_1$$

where W_1 refers to water depth in reservoir of the upper layer of the soil. Second soil layer evaporation (E_{P2}) can be determined as seen in (Eq 3.13) and occurs when layer 1 of the soil is completely depleted. Water is depleted from top layer first as it continues to the lower layers.

$$E_{p2} = \left(E_{p1} - E_{s1}\right)\sqrt{\frac{W_2}{W_{m2}}} \tag{Eq 3.13}$$

The right hand side of the (Eq 3.13) represents a reduction term and it shows how difficult it is to evaporate lower layers of the soil. (Eq 3.14) shows how CREST model calculates the evaporation of the second layer.

$$\text{For } W_2 > E_{p2} \quad E_{s2} = E_{p2} \tag{Eq 3.14}$$

$$\text{For } W_2 \leq E_{p2} \quad E_{s2} = W_2$$

The potential evaporation of layer 3 (E_{P3}) is calculated as Layer 2 gets depleted as shown in (Eq 3.15).

$$E_{p3} = \left(E_{p2} - E_{s2}\right)\frac{W_3}{W_{m3}} \tag{Eq 3.15}$$

The reduction term on the right is greater in (Eq 3.15) than in (Eq 3.13) and this results to faster depletion of E_{P3} than E_{P2}. The rates of depletion in Layer 3 is calculated as shown in (Eq 3.16). ES3 refers to the evaporation rate.

$$\text{For } W_3 > E_{p3} \quad E_{s3} = E_{p3} \tag{Eq 3.16}$$

$$\text{For } W_3 \leq E_{p3} \quad E_{s3} = W_3$$

An aggregation of the evaporation in the vertical profile of a soil is as shown in (Eq 3.17) and it combines the evaporation in all the 3 layers.

$$E_a = E_c + E_{s1} + E_{s2} + E_{s3} \tag{Eq 3.17}$$

The evaporation process begins from drying of the canopy to the lowest layer of the soil and water lost through evaporation is not used in calculating the water budget in the CREST model.

In **runoff routing**, the excess overland flow (R_O) and the interflow (R_I) enter into two linear reservoirs as shown in the *Figure 3-9*. The depths of these reservoirs are calculated as with subscript t indicating time steps, S_O and S_I represents overland and Interflow reservoir depths. Subsurface (Interflow runoff ($R_{I,out}$)) and overland (Overland runoff ($R_{O,out}$)) reservoir discharges are calculated as shown in (Eq 3.20) and (Eq 3.21).

$$S_{o,t+1} = S_{o,t} + R_{o,t} \tag{Eq 3.18}$$

$$S_{I,t+1} = S_{I,t} + R_{I,t} \tag{Eq 3.19}$$

$$R_{O,out} = k_o \cdot S_0 \qquad \text{(Eq 3.20)}$$

$$R_{I,out} = k_I \cdot S_I \qquad \text{(Eq 3.21)}$$

where K_O and K_I are constants determined during calibration with their values being functions of basin area, slope and the cell size used for the model. The two discharges from the reservoirs ($R_{I,out}$, $R_{O,out}$) are then used in the modules used for routing downstream.

In **downstream routing**, the CREST model uses the cell-to-cell routing design describing both interflow and overland flows. (Eq 3.22) shows calculation of concentration time for output of the overland and interflow runoff.

$$T^j = \frac{i^j}{K_X \sqrt{S^j}} \qquad \text{(Eq 3.22)}$$

where spatial index of a cell is represented by superscript *j* as shown in *Figure 3-9* (c) and (d), concentration time is given by T^j which is denotes the time taken from cell *j* to the downstream cell (*j+1*). The euclidean distance between the cells is given by l^j while slope is to the (*j+1*)th from *Jth* cell is given by S^j. K_x denotes the coefficient of runoff velocity and can be estimated from roughness of the land surface.

The CREST model uses the water balance equation (Eq 3.23) for the water balance of each cell. *Sto* refers to total water stored in a cell (see *Figure 3-9* (a)).

$$\frac{dSto}{dt} = P - E_a + \sum R_{O,in} - R_{O,out} + \sum R_{I,in} - R_{I,out} \qquad \text{(Eq 3.23)}$$

In **coupling rainfall-runoff generation and routing**, the model incorporates run-off from one cell to the other to generate rainfall-runoff and this is catered for using three processes in the model. Firstly, runoff from cells in the upper section of the model is assumed to be direct precipitation to upper layers of the lower cells in the model and this calls for modification of the P_{soil} equation used in (Eq 3.3). The resulting equation is represented in (Eq 3.24).

$$\hat{P}_{soil} = P_{soil} + \sum R_{O,in} \qquad \text{(Eq 3.24)}$$

where $\hat{P}_{soil}$ refers to the adjustment made on P_{soil} and the additional water is can infiltrate the soil layers through the upper layer. The second process caters for lateral movement of soil moisture which increases through interflow from cells in the upper zones. This calls for adjustment of infiltration amount which is adjusted as indicated in (Eq 3.25).

$$\hat{I} = I + \sum R_{I,in} \qquad \text{(Eq 3.25)}$$

In (Eq 3.25), $\hat{I}$ refers to adjusted interflow while I is established by summing upstream interflows. In the third process, a modification of the depth of the reservoir of overland flow from the channels is done during coupling of routing and rainfall runoff generation. This runoff contributes to the depth of the overland reservoir. The modified S_O (from (Eq 3.18)) is as indicated in (Eq 3.26).

$$\hat{S}_O^{t+1} = S_O^t + R_O^t + \sum R_{O,in} \qquad \text{(Eq 3.26)}$$

where $\hat{S}_O^{t+1}$ refers to the adjusted S_O^{t+1} which represents the increase by channel runoff contribution.

3.4.1.   Model calibration and validation

In hydrological modelling, there are parameters which cannot be estimated by calculation or measurements and can only be inferred on a trial and error basis. The trial and error method of determining the optimum values of these parameters can be termed as a calibration process. Calibration can also be done using automatic methods which tend to optimise the model parameters to match the observed values as much as possible by setting the calibration limits of each parameter and letting a defined algorithm match the observed values as much as possible. Calibration of a model is done to determine the optimum values of parameters of a model which will yield the best simulation of the observed values.

In this model, the calibration was done with reference to the river discharges recorded at outlet stations of each sub-basin. The Parameter Estimation (PEST) method of calibration was used in this research. PEST utilises the Gauss-Marquardt-Levenberg algorithm. This algorithm works best when the simulated results complement the model parameters and these parameters should be able to change as the model is continuously calibrated (Doherty, Brebber, & Whyte, 2005). PEST works around the CREST model as a shell by optimizing the parameters of the model independently (Van der Velde, 2010). The sequence of the PEST tool operations can be summarised in *Figure 3-10*.
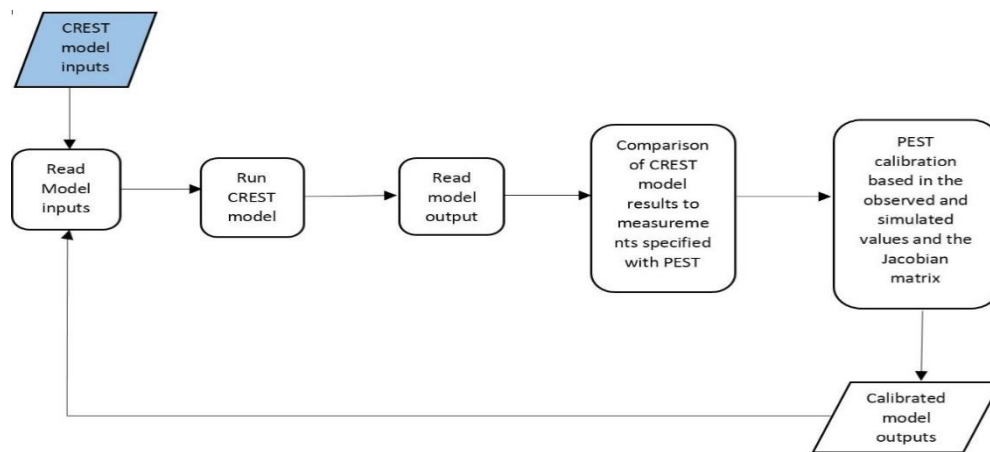

Figure 3-10: PEST operating sequence

The sequence for adjusting the parameters is repeated until the optimum values are attained. The number of reiteration of the adjustments can be specified in PEST setup. The optimization procedure can be summarised as shown in *Figure 3-11*.
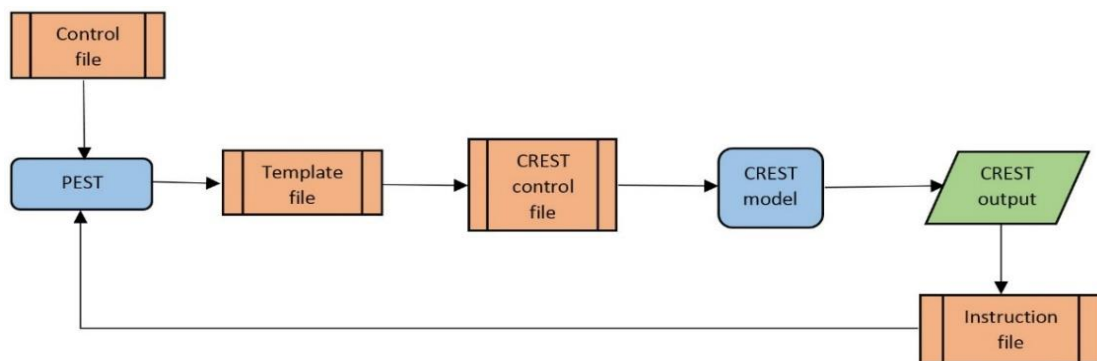

Figure 3-11: PEST optimization procedure

In the optimization procedure, PEST integrates a number of files to the CREST model. PEST is integrated into the CREST model using three files: the template file (denoted as *.tpl*), this file is used to instruct PEST where the optimised model is written in the CREST model input file, the instruction file (denoted as *.ins*), this file is used to instruct PEST the position in the CREST model output should it use to recalibrate the model based in the observed values. The third file is the PEST control file (denoted as *.pst*) which is used to define the whole PEST setup. The control file is used for configuring the method used for optimization, has the calibration parameters and their limits, contains sections where the observed data is defined and has a model interface needed for communication with CREST model.

In his thesis, Gathecha, 2015 defines the PEST objective function as *phi* (Φ)which is given by summing the values of squares of the deviations seen between the observed and the simulated values. (Eq 3.27) denotes the PEST's objective function.

$$\Phi \overrightarrow{(b)} = \sum_{i=1}^{n} [W_i (R_{obs,i} - R_{sim,i})]^2 \qquad \text{(Eq 3.27)}$$

Where Φ refers to the PEST objective function, $\overrightarrow{(b)}$ denotes the vector containing fitting parameter while W_i is associated weight of a measurement at a specific point. $R_{obs,i}$, and $R_{sim,i}$ refers to the observation at time interval i and the corresponding prediction given by the model respectively.

The results of calibration of the CREST model are given in the results folder of the model setup. They include the Nash Sutcliff Efficiency (NS), the Correlation Coefficient (R) and the bias. (Wang et al., 2011). *NS* objective function has a maximum value of 1 and the closer one approaches this value the better the performance of the modelled data. (Eq 3.28) shows how the *NS* is derived.

$$NS = 1 - \frac{\sum_{n-1}^{N} (Q_{obs} - Q_{sim})^2}{\sum_{n-1}^{N} (Q_{obs} - \bar{Q})^2} \qquad \text{(Eq 3.28)}$$

where N refers to total number of events (data elements), Q_{sim} refers to the simulated runoffs at a certain time interval while Q_{obs} refers to the observed runoffs nth time intervals. $\bar{Q}$ refers to the mean of the runoffs recorded over time. The correlation coefficient (R) analyses how well the simulated and the observed discharges relate to each other. R value of 1 gives the best performance of the model. (Eq 3.29) shows the equation used to derive R

$$R = \frac{\sum (Q_{obs} - \bar{Q}_{obs})(Q_{sim} - \bar{Q}_c)}{\sqrt{\sum (Q_{obs} - \bar{Q}_{obs})^2 (Q_{sim} - \bar{Q}_c)^2}} \qquad \text{(Eq 3.29)}$$

where $\bar{Q}_c$ refers to the average of all simulated daily discharge values. The other result produced by the CREST model is the bias ratio value. This is used to give an indication of systematic bias of the discharge simulated. A 0% bias value is considered to be the best performance of a model. (Eq 3.30) shows how to derive the relative bias.

$$Relative\ Bias = \frac{\sum Q_{sim} - \sum Q_{obs}}{\sum Q_{obs}} \times 100\% \qquad \text{(Eq 3.30)}$$

To determine whether a model is performing well, it should be subjected to different datasets but with the same conditions as that used during the calibration process. This can be described as the validation process. Validation of the model results is done after the calibration process. This process is used to determine the performance of the calibrated model. The calibration of the model was done for the period

2001 to 2010 with 2001 being the warming up period of the model and the successive years being the simulation period. Validation was done for the years 2010 to 2013.

3.4.2. Calibration procedure

Sensitivity analysis is done after the calibration process. It is done by applying stress to the model parameters to determine the effect of these stresses to the achieved objective function values of model calibration. Sensitivity analysis quantifies how different model parameters affect the overall performance of the model and it can be done to reduce the uncertainties associated with the model structure when subjected to different conditions.

In running the CREST model for Lake Naivasha basin, Gathecha, (2015) conducted a sensitivity analysis of the model parameters as applied to the basin. The *NS* objective function was used as the benchmark to the effects of varying these model parameters. The results were categorized into 3 based on the degree each parameter had on the overall performance of the model. The findings can be summarized in *Table 3-4*.

Table 3-4: Sensitivity analysis results of CREST model (Adapted from Gathecha, (2015))

Category	Parameters	Sensitivity level
1	Ksat	Low sensitivity
	WM	
	IM	
2	B	High sensitivity
	RainFact	
	CoeS	
	KE	
3	KI	Very sensitive
	KS	
	CoeR	
	expM	
	coeM	

However, the calibration period of the CREST model in this study was increased to ten years from the previous period of four years in 2014. The calibration was done based on the values that were previously established for calibration of TRMM3B42v7 rainfall since it performed better than CMORPH in the basin. The validation period of the model was from the year January 2011 to December 2013. The calibration will be done for the Malewa, Gilgil and Karati sub-basins.

3.5. Lake balance model

The lake balance equation can be expressed as shown in (Eq 3.31). The units of the change is in cubic meters per month ($m^3 month^{-1}$).

$$\frac{\Delta V}{\Delta t} = P - E + Q_{inflow} - Q_{aq} - Q_{abs\_lake}$$

(Eq 3.31)

Where the rainfall received on the lake is denoted by *P*, evaporation of the lake surface by *E*, Q_{inflow} represents the inflow by surface water into the lake. Loss of water from the lake through the ground water aquifer is represented by Q_{aq} while the abstractions from the lake are represented by Q_{abs_lake}. Inflow into the lake is formulated as shown in (Eq 3.32) (Units: $m^3 month^{-1}$),

$$Q_{inflow}^t = \sum_{i=1}^{n} inflow - \sum Q_{abs\_rivers}$$

(Eq 3.32)

where the *inflow* is the runoff from a total of *n* sub-basins and Q_{abs_rivers} refers to abstractions from the rivers. The outflow Q_{aq} through the aquifer is derived using (Eq 3.33) (units: $m^3\,month^{-1}$).

$$Q_{aq}^{t} = C * (H_{lake}^{t-1} - H_{aq}^{t-1}) \qquad \text{(Eq 3.33)}$$

where *C* refers to hydraulic conductance of the lake aquifer in $m^2 month^{-1}$, the H_{lake} is the level of water of the lake (m) while H_{aq} refers to the water level of the aquifer. The aquifer water level is updated using the calculated inflow and outflows using (Eq 3.34).

$$H_{aq}^{t} = H_{aq}^{t-1} + \left(\frac{Q_{aq} - Q_{outflow} - Q_{abs\_aq}}{A * S_y} \right) [m] \qquad \text{(Eq 3.34)}$$

In (Eq 3.34), the $Q_{outflow}$ refers to an outflow to an external area at a constant rate while *A* is the aquifer surface area. The specific yield of the aquifer is indicated by S_y while the abstractions from the aquifer are denoted by Q_{abs_aq}.

A relationship between the lake area and its volume of water was used to establish the lake water balance. In their paper, van Oel et al., 2013, used the relationship between the lake-area to its volume based on a bathymetric survey conducted in 1983. Sub-basin outlet stations were used for discharge measurements while the 9036002 (Naivasha DO) station was used as a source of rainfall data. Evaporations rate records used were those obtained from Kenya Ministry of Water Development (MoWD, 1982) records. The discharges used for the lake balance model were obtained from the same stations used in running the CREST model.

According to Van Oel et al. (2013), the alarm stage of the lake levels is 1885.3 m.a.s.l, while the stage level for water scarcity is 1884.5 m.a.s.l. The lake level reached the water scarcity level in late 1999. During this year, the abstraction of both the ground and surface water abstractions contributed to this drastic decrease in the Lake volume. The model quantifies how lake and ground water abstractions affect the lake levels. A time series depiction of the Lake Naivasha water levels over the period January 1965 to June 2015 is as seen in *Figure 3-12*.
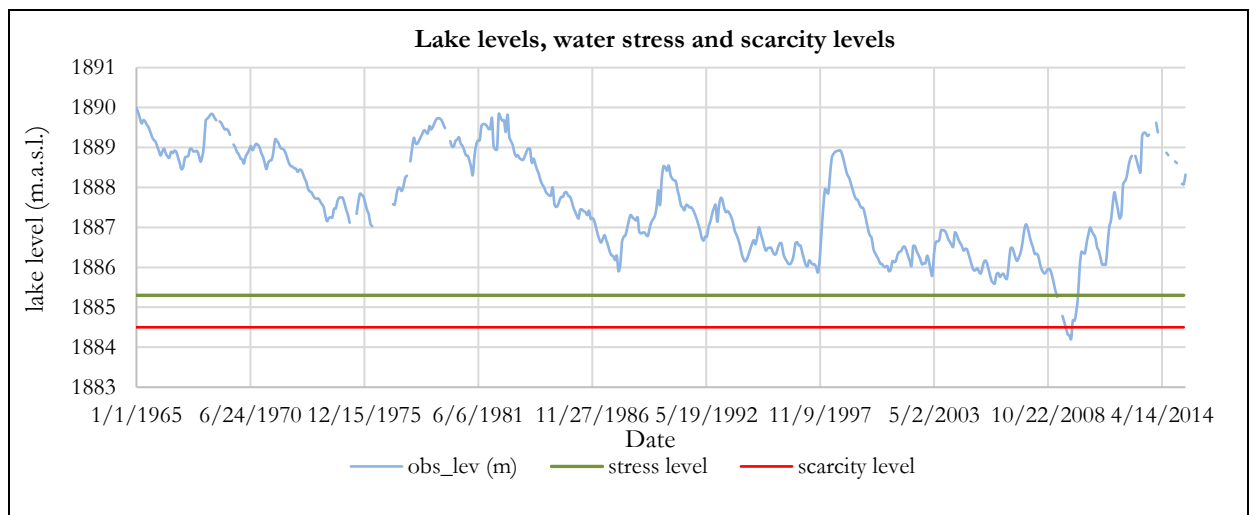

Figure 3-12: Graph of Lake Naivasha levels and the water stress and scarcity levels.

According to the Water allocation Plan (WAP) of LNB (WRMA, 2010), there are 3 thresholds of abstraction restrictions with four indicators based on water availability that can be applied to both the lake

levels and the discharges experienced in the rivers. The four indicators include the green zone, the amber zone, the red zone and the black zone.

The black zone of the Lake level was set at 1882.5 m.a.s.l. which is approximately one meter of the lake water depth. At this level, most irrigation abstractions that draw water from Lake Naivasha are not feasible since the shoreline recede far away from the abstraction points. *Table 3-5* show the abstraction restrictions of the lake and the rivers.

Table 3-5: Water abstraction restriction rules

Color coding zones on abstractions restrictions & reserve water	Water Resource Status i.e. Periods when		
	River flows	Lake Naivasha levels meters above mean sea level	
		Lake water	Ground water within conservation area
Green i.e. satisfactory. (*Abstraction allowed up to permit limits*)	Higher than Q_{80}	Higher than 1885.3	Higher than 1885.3
Amber i.e. stress. (*Slight abstraction restrictions imposed*)	Between Q_{80} and Q_{94}. Abstractions allowed: 1. Domestic and public water supplies 2. Others cease to pump from water bodies and revert to their 90 day storages	Between 1885.3 and 1884.6 Abstractions allowed: 1. Domestic and public water supplies 2. Others draw 75% of their water use	Between 1885.3 and 1884.6 Abstractions allowed: 1. Domestic and public water supplies 2. Others draw 75% of their water use.
Red i.e. Scarcity (*Severe abstraction restrictions imposed*)	Between Q_{95} to Q_{96}. Abstractions allowed: 1. Domestic and public water supplies draw 50% of their water use. 2. Others continue to draw from their 90 day storages	Between 18884.5 to 1882.5 Abstractions allowed: 1. Domestic and public water supplies draw 75% of water used 2. Others draw 50% of their water use.	Between 1884.5 to 1882.5 Abstractions allowed: 1. Domestic and public water supplies draw 75% of water use. 2. Others draw 50% of their water use.
Black i.e. reserve (*Full protection: Surface water used restricted to basic human/livestock need and nature (ecosystem) only*)	Lower than Q_{96}. Abstraction allowed: 1. Domestic and public water supplies draw amounts for basic body needs i.e. 25 litres per person/livestock unit per day.	Lower than 1882.5 Abstraction allowed: 1. Domestic and public water supplies draw amounts for basic body needs i.e. 25 litres per person/livestock unit per day.	Lower than 1882.5 Abstraction allowed: 1. Domestic and public water supplies draw 75% of water use. 2. Others draw 50% of their water use.

In recent years, 2006-2009, the lake levels receded below the scarcity level. This was attributed to the abstractions that were taking place in the upper areas of the catchment and in the lake (Odongo et al., 2014). The shore receded to levels that necessitated the farm owners to dig canals of up to 15 meters deep to access the lake water. An analysis of time series Landsat images of Lake Naivasha captured during this period can be used to show the extent of reduction of the lake level. The images are shown in *Figure 3-13* and *Figure 3-14*.

Figure 3-13: Landsat 4 TM image (January, 2010)

Figure 3-14: Landsat 8 image (February, 2016)

Figure 3-13 show the lowest lake levels and while *Figure 3-14* shows a recovered lake after conservation measures were made to mitigate the drying up experienced in the year 2009. The Landsat images were processed using ArcMap. The images show the difference in area of the lake between the two periods, there is also a difference in the surrounding waters of the island (Crescent Island) on the north-east of the lake. The proximity of the green houses to the lake is also seen to be different between the two periods with the furthest distance being seen on the January, 2010 image. There is also some difference in distance between the main Lake and Lake Oloiden on the south west of the images.

As at June, 2015, the lake levels were at 1888.54 m.a.s.l. This is above the lake's stress level of 1885.3 m.a.s.l. and the scarcity level of 1884.5 m.a.s.l. The amount of streamflow to the lake plays a key role in maintaining the lake levels and volumes. This is in relation to the amount of precipitation received, the amount of evapotranspiration and the abstractions from the lake and its related aquifers.

The thresholds for lake levels that were determined after the WAS 2010 report (De Jong, 2011) and the WAP (WRMA, 2010) will be used as a basis for lake system resilience analysis that will be conducted in this study. Allocation data from the WRMA database will also be used.

4. METHODOLOGY, DATA PROCESSING AND ANALYSIS

4.1. Methodology flowchart

A flowchart of the processes that were followed to achieve the research objectives are summarised as indicated in *Figure 4-1*.
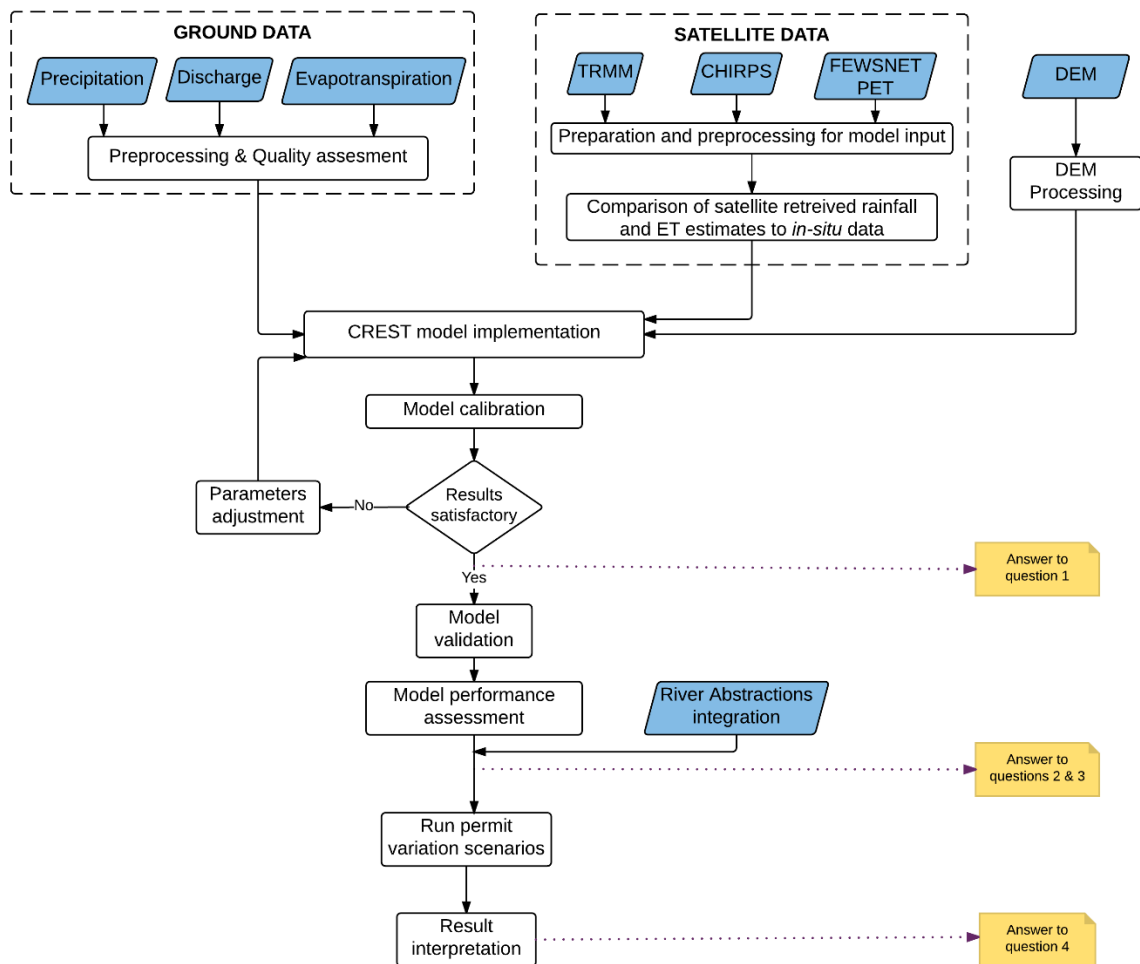
Figure 4-1: Methodology flow chart of processes used to achieve the research objectives.

4.2. Data analysis

4.2.1. Flow duration curves

Time series projections of flows can be used to visualise the variability of discharge on a stream. Discharge data used in this research was collected from RGS stations located on strategic locations at the outlet of the sub-basins. These stations were assessed for data availability and consistency. One method used to assess the discharges is by developing and analysing the flow duration curves of all the stations. Flow duration curves are used to assess the probability of exceedance of a certain amount of flow in a channel. The flow duration curves can be used to determine the environmental flow based on a set threshold.

The Q_{80} is used to distinguish the normal and flood flows. Flows above this threshold can be allocated for anthropogenic uses while those below this threshold are considered as environmental flow and should not be allocated. In creating the flow duration curves, the discharge data is sorted from the largest to the smallest and these values are assigned a probability of exceedance. *Figure 4-2* shows the flow duration curves of different stations located at each sub-basin which were developed from discharge data collected between 1960 and 2015.
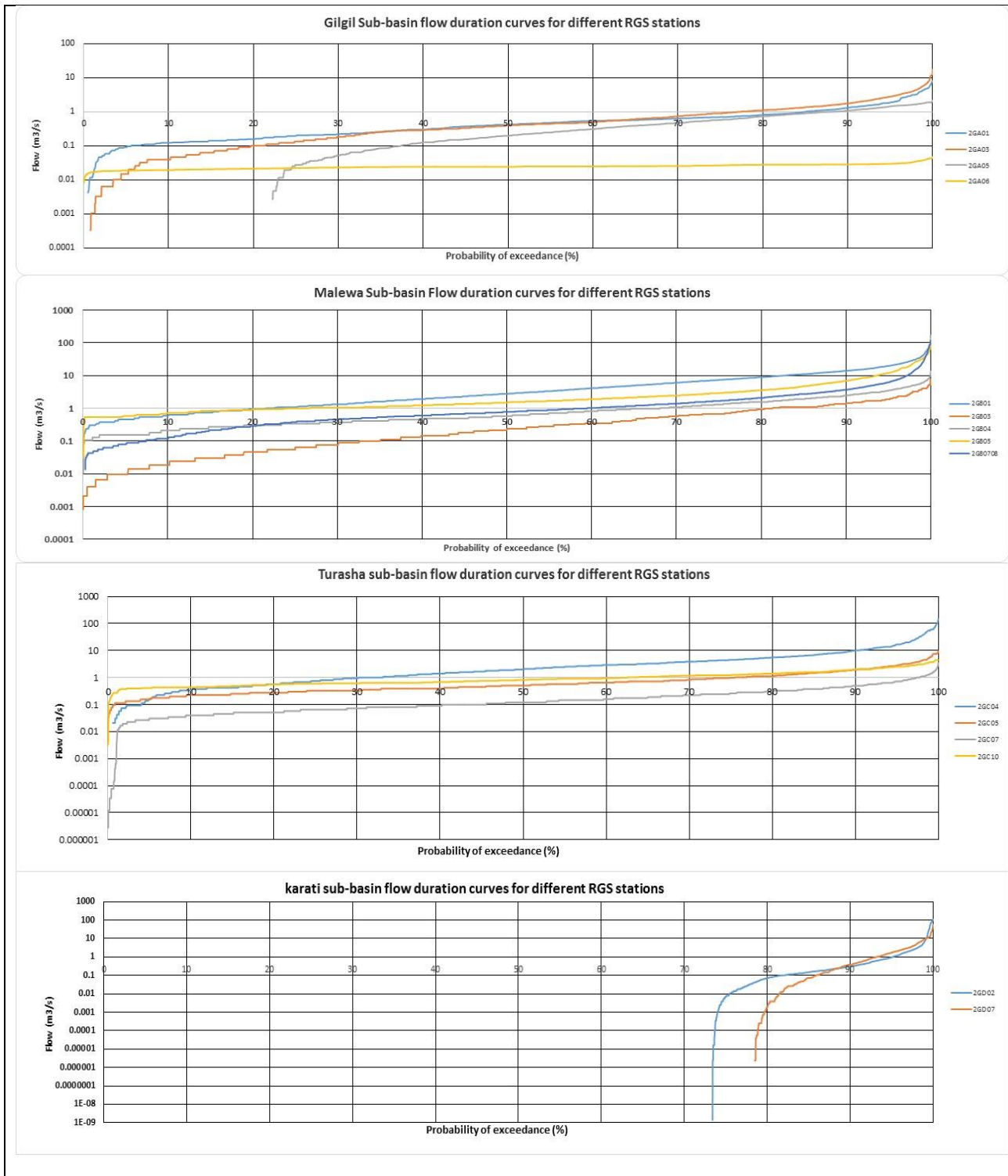
Figure 4-2: Flow duration curves (Exceedance probability) for different RGS stations in Lake Naivasha Basin

From *Figure 4-2*, it can be seen that different sections of the basin show highly variable discharges. 2GA03 is the second last station before the outlet of the sub-basin and has a Q_{80} discharge of 1.11 m³/s. 2GA01 station which is the last station on the downstream of the Gilgil sub-basin has a lower Q_{80} compared to 2GA03. This may be attributed to a combination of changes in the gauging stations and abstractions from the river between these two stations.

It can be seen that 2GA05 dries up at certain times. This means that the allocation of water at the section represented by this station should be considered so as to prevent the drying up scenario. 2GA06 experiences the lowest discharges in the Gilgil sub-basin. It has a Q_{80} discharge of 0.03 m³/s. Even though this station experiences low discharges, it does not dry up like 2GA05.

The source of Malewa river sub-basin are in the highest parts of the Lake Naivasha basin. This is the Aberdares ranges on the North-eastern part of the basin. This sub-basin can be partitioned into two sections, the upper Malewa and the Turasha sub-basins. These areas receive high rainfall due to orographic effects. Consequently, there are no seasonal streams on this sub-basin. In this study, the headwater stations i.e. 2GB03 and 2GB04 located at short distances from the sources show lower Q_{80} values. These stations experience 0.91 m³/s and 1.56 m³/s Q_{80} discharges respectively. This is higher than the discharge of all downstream stations of the Gilgil sub-basin.

Turasha sub-basin shows higher discharge values. 2GC04 was selected for the analysis as it is located after the confluence of all headstreams in this sub-basin. 2GC04 has a Q_{80} discharge of 5.14 m³/s. This is more than all the discharges contributed from the upper section of the Malewa sub-catchment which is calculated at 2GB05 which has a Q_{80} discharge of 3.47 m³/s. This shows that the Turasha sub-catchment contributes the most amount of runoff in the Lake Naivasha basin. 2GB01 is located after the confluence of 2GB05 and 2GC04 stations and it combines the flows of the two sub-catchments i.e. Upper Malewa and Turasha. A combined Q_{80} of 2GB05 and 2GC04 gives a discharge of 8.89 m³/s. This shows that the Malewa sub-basin can withstand more abstractions than the other sub-catchments.

The Karati sub-catchment experiences seasonal flows. This is shown by the flow duration curves retrieved from 2GD07 and 2GD02. The Q_{80} of these two stations show values of 0.00 m³/s but there are flows at some point but these flows are still lower than 0.10 m³/s. This is a low amount of streamflow and cannot be reliably allocated. The source of the Karati stream is from lowlands and this part of the basin receives less rainfall.

A noticeable trend in these flow duration curves is the shape of the curves which are relatively flat in all the sub-catchments except Karati. This shape shows that the flows have some contribution from ground water. The steep slopes of the Karati sub-catchment stations shows that these stations have less influence from ground water recharge.

The Malewa sub-basin is the largest contributor of runoff into the LNB. The Turasha sub-basin drains into the Malewa sub-basin. Even though the basin has a large volume of runoff, there is variation in the amount of runoff received at different stations in the basin. This influences the amount of abstractions that should be authorised for abstraction in each of the rivers. Stations located at the lower sections of the sub-basin experience high runoff volumes. In these sections, the amount of abstraction allowed can be higher than in the upper stations where the runoff is very low. A graphical analysis of runoff generated from stations in the Malewa sub-basin during the period 2000-2010 is shown in *Figure 4-3*.
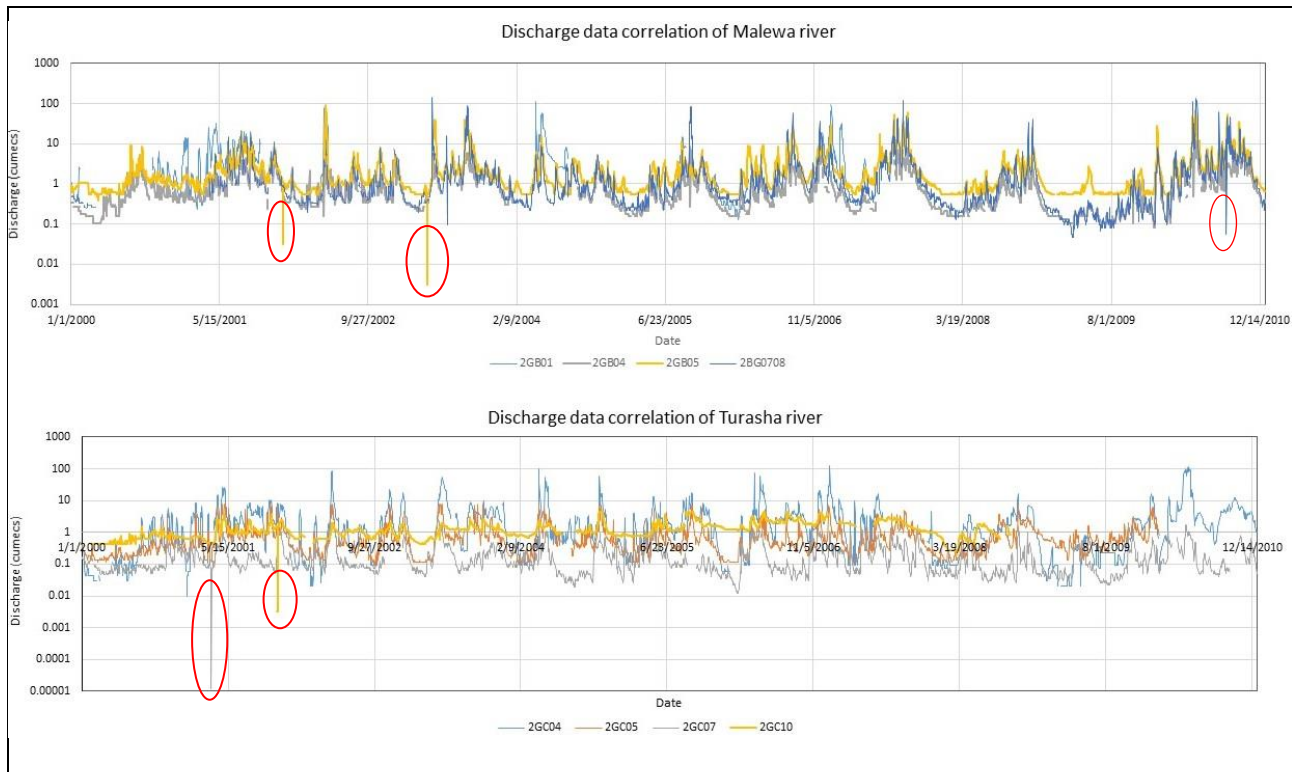
Figure 4-3: Discharge experienced at different stations in Malewa and Turasha Sub-basins for the period 2000 - 2010

A visual trend analysis of *Figure 4-3* shows that there is a close relation of the peaks and troughs of discharges trends exhibited at the different stations in each basin. On the Malewa basin, it is seen that the outlet stations such as 2GB01 and 2GB05 have higher volumes of runoff compared to 2GB04 and 2GB07/08. Sections of the basin with high runoff can be used to allocate more water for uses than those with lower runoff values. The matching peaks and troughs give an indication of the routing in the sub-basin. It shows that the stations are well connected and an effect on the amount of flow experienced on one station will be experienced on a downstream station. This may also give an indication on the spatial extent of rainfall experienced with the catchment.

However, there is evidence of few anomalies in the flow data analysis. There are a few discontinuities in the data series. The discontinuous lines arise when there were no discharge measurements taken from at the stations. The red oblongs in *Figure 4-3*, are used to show uncertainties in the data which may have been caused by wrong data recordings. These instances are however not so pronounced as they are very few within the data series and do not affect the reliability of the data for use in our research.

2GB01 and 2GC04 can be seen to have the highest runoff peaks compared to the other stations in the two sub-basins. Both stations are located at the outlets of the two sub-basin with 2GC04 draining into Malewa from the Turasha sub-basin. It is also noted that during the period 2007 to 2009, the amount of discharge received at station 2GB05, which is the outlet station of upper section of upper Malewa continued to produce high amounts of runoff compared to other stations located at the outlet of the other sub-basins.

4.2.2. Double mass curves

Double mass curves can be used in assessing the consistency of hydrological datasets that has been collected over a number of years. Inconsistencies can be caused by changes in observation methods employed during data collection, shifting of stations or changes in the ecosystem. These inconsistencies

can also be caused by degradation of the works at the stations such as degradation of a weir or meandering of a river over the years which may drift the river away from the gauge.

It is important to have consistent data so as to make informed decision on the changes experienced in river flows. Flow durations curves were prepared by calculating the cumulative of a station in a sub-basin. The discharge at the outlet of the sub-basins were also calculated. These two cumulates were then plotted against each other with the cumulative of a station on the y-axis and the cumulative of all the station on the x-axis. The data series used runs from 1960 to 2015.

Double mass curves of four stations were assessed in the Gilgil sub-basin. *Figure 4-4* shows the resulting double mass curves. 2GA01 has an inconsistent double mass curve as the graph shows an erratic trend from the onset. This could be due to erosion or change in discharge estimation methods. 2GA03, 2GA05 and 2GA06 have consistent double mass curves with fairly small deviations experienced at 2GA06. 2GA05 has fewer discharge records compared to the other four stations.

Figure 4-4. Graphs showing discharge double mass curves of stations in the Gilgil Sub-basin
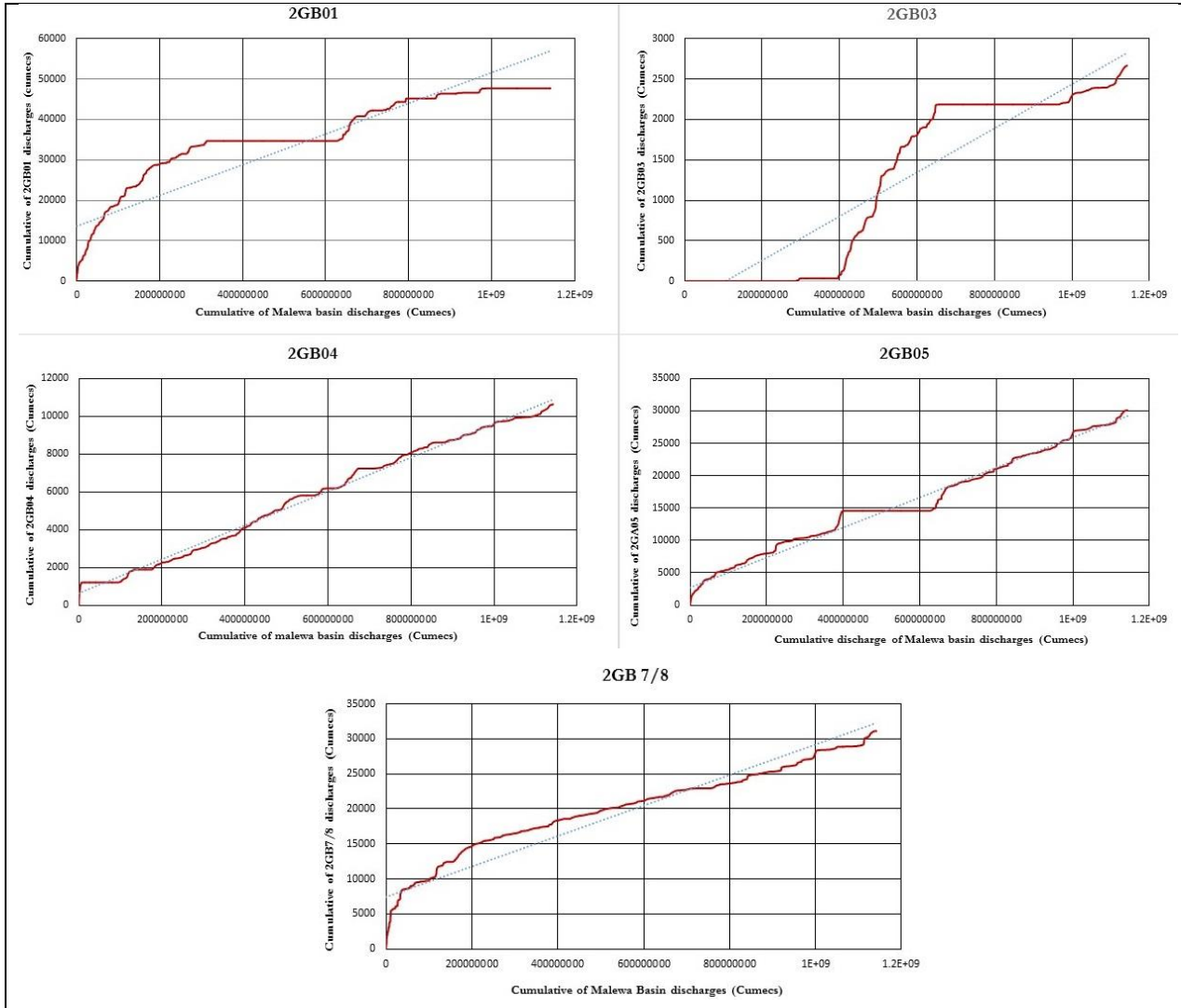
Figure 4-5. Graphs showing discharge double mass curves of stations in the Malewa Sub-basin

Malewa sub-basin has five stations that were used for construction of the double mass curves. The trends of stations in this sub-basin are generally good compared to what was exhibited at 2GA01 as seen in *Figure 4-5*. However, 2GB01 and 2GB03 exhibited some significant deviations from the main trend line. This brings out some discrepancies in the data that need to be investigated. The deviations are on both sides of the trend line and this suggests that the changes experienced have caused both increase and reduction in the amount of streamflow recorded. The trend of the double mass curves also exhibits zero gradients at some points.
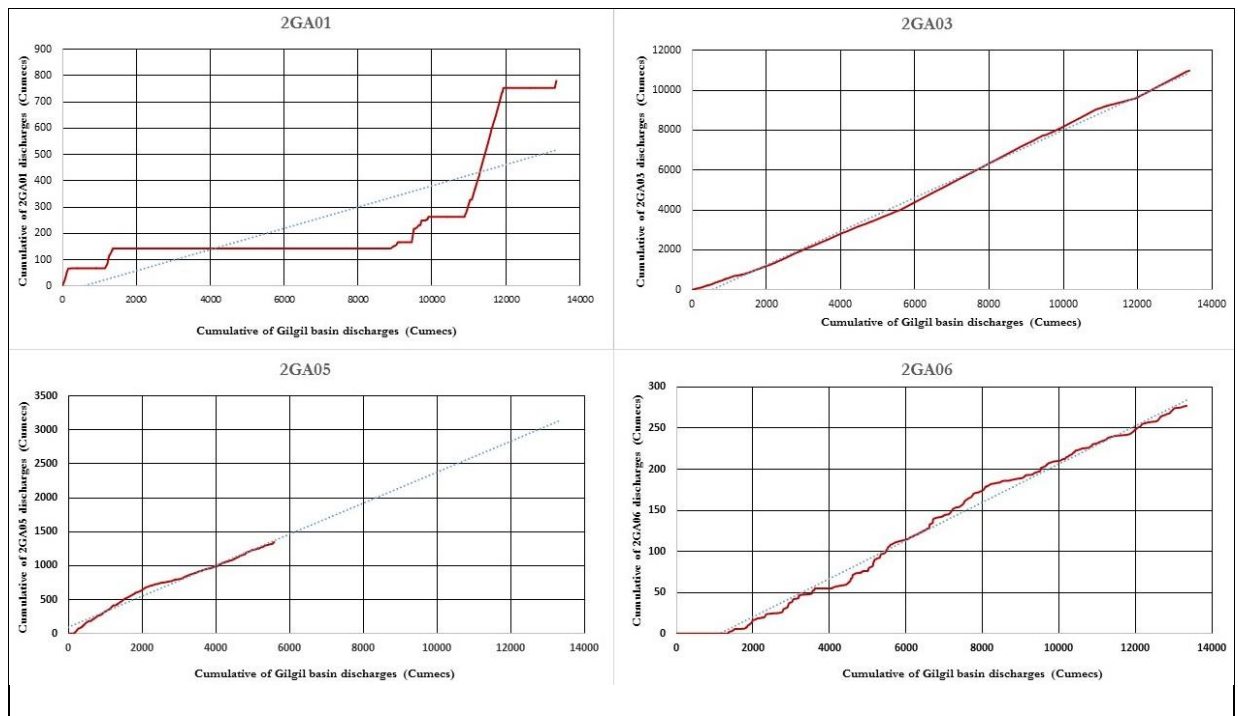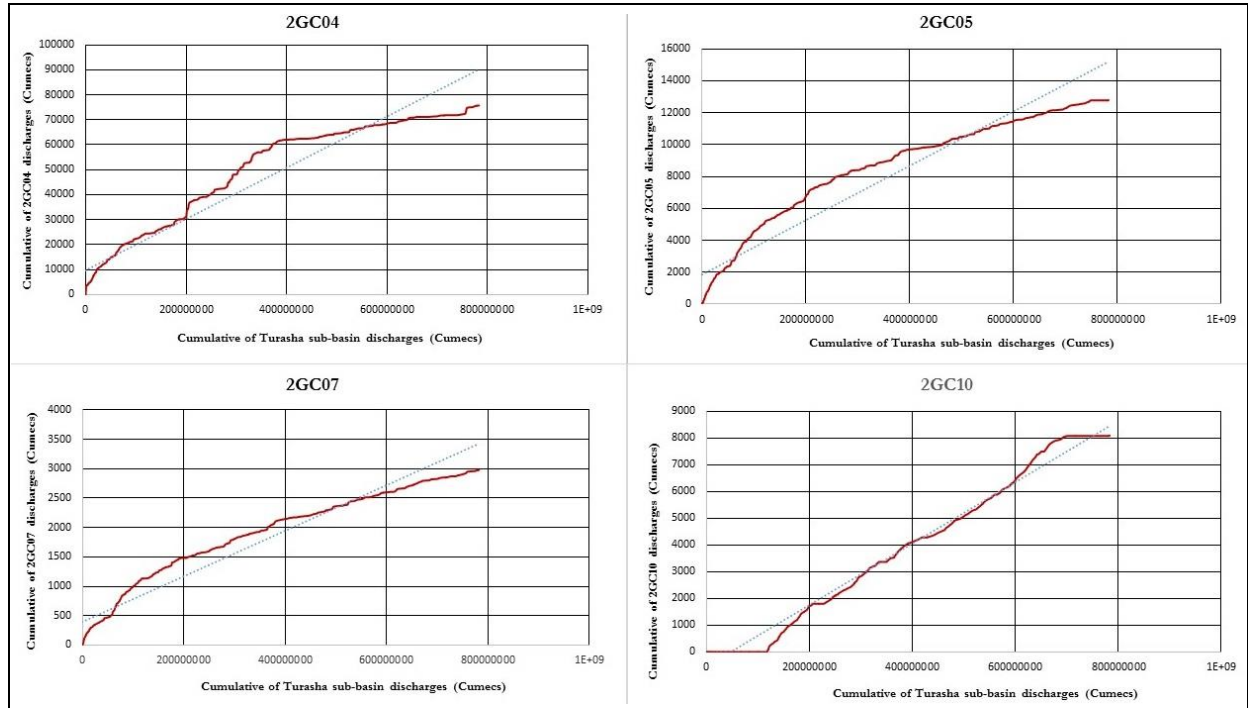
Figure 4-6. Graphs showing discharge double mass curves of stations in the Turasha Sub-basin

Figure 4-6 shows the double mass curve of stations in the Turasha sub-basin. The curves generally show consistent trends amongst all the stations in the basin with small deviations from the trend lines. The values of discharge in these station is also relatively high. The double mass curve of these station show fewer deviations from the trend line.
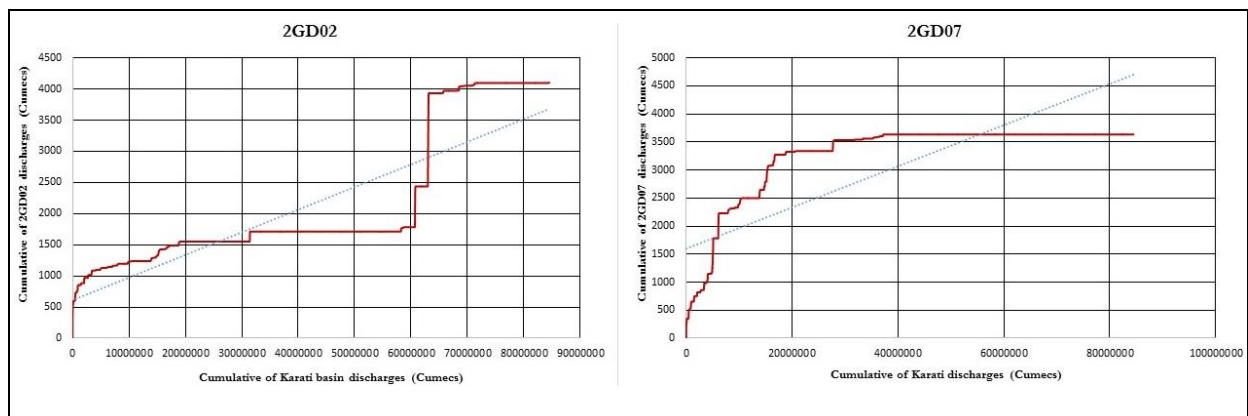
Figure 4-7.Graphs showing discharge double mass curves of stations in the Karati Sub-basin

The two reliable regular gauging stations of Karati sub-basin are 2GD02 and 2GD07. *Figure 4-7* show the double mass curves from the two stations. The curves show enhance deviations of the discharges from the anticipated trends. The basin has ephemeral rivers that only flow during the rainy seasons of the year. This means that discharge measurements can only be scheduled only during the wet seasons.

The deviations exhibited in these double mass curves can be attributed to changes experienced in the basin such as deforestation in the upper catchment of the basin, population increase which has caused more abstractions from the rivers and deviation of river from the main cause due to sedimentation or erosion. A change in frequency of discharge estimation (how often the flow measurements are taken) is also a cause of the deviations exhibited as some measurements were done on a daily basis while some

discharge measurements were done months apart. Degradation of structures such as weirs at the gauging stations is also a cause of deviation from the steam flow trends of different stations over time. A sample picture of changes in some stations in the basin is shown in Appendix 1: .

The use of manual gauging methods were also substituted with automatic methods. Some manual stations were substituted with divers which employ algorithms to convert air pressure into discharges. Divers were installed in stations such as 2GB01. In other stations divers were used for a short period of time and were later uninstalled or vandalised. The algorithms used by these divers may also introduce uncertainties to the final stream flows obtained from a stations. Some divers use algorithms to convert air pressure to river stage height and they need to calibrated before installation on rivers in different parts of the world.

4.3. Satellite rainfall estimates preparation

Daily global satellite products were downloaded and a subset of the Lake Naivasha basin was made. The subset was rectangular in shape and was based on lower left and upper right corners of the basin. Rainfall data used include the TRMM3B42v7 and CHIRPS ranging from 2001 to 2015. TRMM3B42v7 data was in the NetCDF format while CHIRPS data was in the TIFF format. CREST model uses the ASCII format. These products were processed and converted to ASCII format. Processes followed in obtaining the final data can be summarised as shown in *Figure 4-8*.
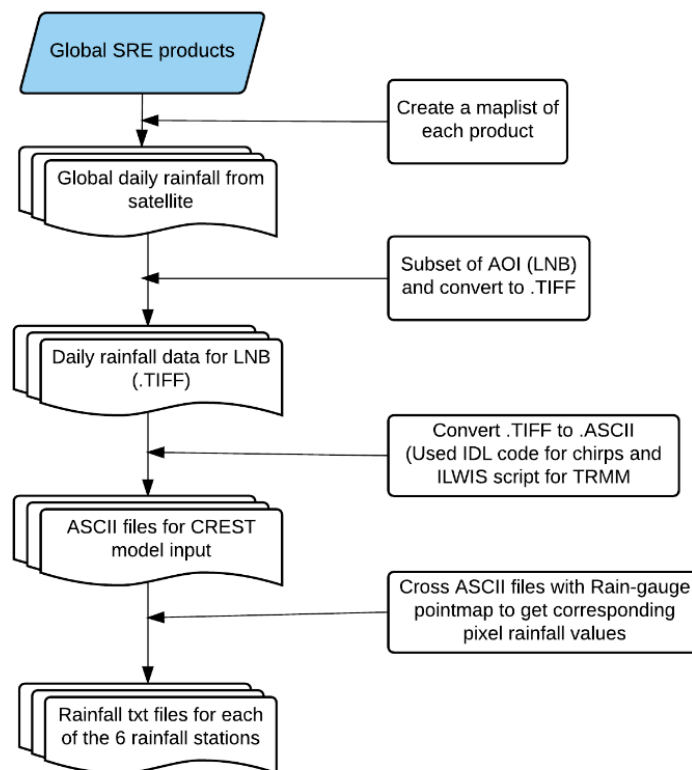
Figure 4-8: SREs processing flowchart

Resulting subset TIFF data was then converted to an ASCII format. These ASCII files have indications of the number of rows and columns in a file, the x-axis left corner, the y-axis left corner, the cell size and the no data values in the file. This information is used to determine the exact position and value of a pixel in the file. In naming of the file, the date is included since it is required in running the CREST model. The date is in the yyyymmdd format. The resulting ASCII files were populated and used as input in the Rain file of CREST model setup.

Since the pixels in the final ArcASCII files are geo-referenced, it is possible to extract the actual pixel that coincided with the rainfall recorded at the gauge station. The pixels matching and value extraction was done using an Interactive Data language (IDL) code. The extraction of the pixel value can also be done by crossing the rainfall maps with point maps of the gauging stations using ILWIS.

Performance analysis of the SREs rainfall was done by comparing the matching pixels to the corresponding locations to the rain gauges. A total of six gauging stations were plotted on a one to one line with the satellite retrieved rainfall. The x and y-axis limits of the graph were levelled to give equal scales for comparison. In plotting the graphs, cumulative of the satellite products were plotted against the cumulative of the gauge data. *Figure 4-9* shows the result of these comparisons.
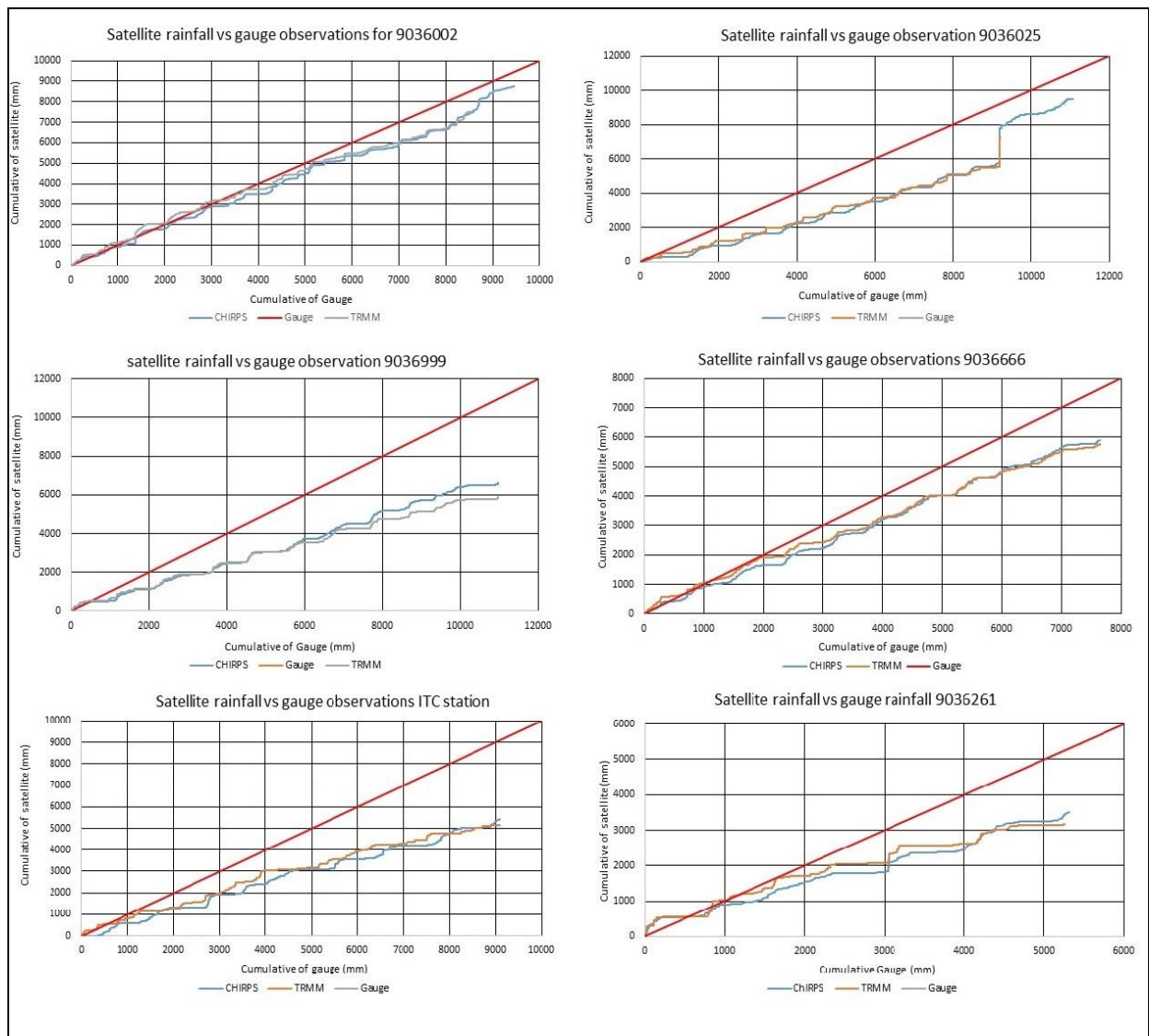
Figure 4-9: Comparison of cumulative of gauge, CHIRPS and TRMM rainfall.

The 9036002 station show the best correlation to the satellite data in the series of stations used from the basin. The level of deviation of the satellite retrieved rainfall products from the cumulative of the gauge at this station becomes significant after the year 2006. After this period, the satellite products tend to underestimate the amount of rainfall recorded. Underestimation of rainfall by both TRMM3B42v7 and CHIRPS is seen in all the stations even though the degree differs.

It should however be noted that the two satellite products tend to have highly correlated values. Both TRMM3B42v7 and CHIRPS tend to have the same trends in terms of peaks and troughs. CHIPRS has the highest resolution of the two products and it was expected it would give better correlation to the gauge as compared to TRMM3B42v7 but this was not the case. The two datasets showed very little trend variation from each other. Statistical analysis was done to gauge the performance of the two satellite retrieved rainfall products with the gauge data as a benchmark.

5. RESULTS AND DISCUSSION

5.1. Rainfall data analysis

A statistical analysis of performance of satellite rainfall products was done. Rain gauge data was used as a benchmark for this performance test. Satellites use algorithms to get the final value of rainfall amounts. These algorithms have limitations which introduce uncertainties to the final amount of rainfall recorded. The uncertainties are due to random and systematic errors. These errors can be termed as biases. Statistical analysis of both TRMM3B42v7 and CHIRPS was done to ascertain their performance. The bias and statistical error analysis was done for each rainfall station.

Probability of detection (POD) is one of the performance indicators that were analysed in this research. It refers to the ratio of the detected instances to that of all the possible blibs on the satellite. It can be further elaborated as a ratio of the correct values of rainfall recorded by a satellite to the total number recorded by the rain gauge. A value of 1 denotes a perfect detection probability with zero being a poor performance scenario. The Frequency Alarm Ratio (FAR) is a performance indicator used to evaluate the performance of a satellite by finding the ratio of a satellite instances of rainfall detection to those instances where the rain gauge does not detect the rainfall. A 0 value is considered to be a perfect FAR (Moazami, Golian, Kavianpour, & Hong, 2014). The equation used to derive the POD and FAR can be shown in (Eq 5.1) and (Eq 5.2) respectively.

$$POD = \frac{t_H}{t_H + t_M} \qquad \text{(Eq 5.1)}$$

$$FAR = \frac{t_F}{t_H + t_F} \qquad \text{(Eq 5.2)}$$

where H refers to a rain instance detected correctly, M is the instance where the rain recorded by the gauge was not detected while F is rainfall detected by the satellite but was not observed at the gauge. The terms t_H, t_F and t_M refers to times of occurrence of each instance. In this research, the threshold used for rainfall detection occurrence is a value of 0. The Relative bias (*RBias*) gives a percentage of the systematic difference between the rainfall values given by the satellite and those read from the rain gauge. It can be obtained by applying (Eq 3.33).

$$RBias = \frac{\sum_{i=1}^{N}(P_{oi} - P_{si})}{\sum_{i=1}^{N} P_{oi}} \times 100\% \qquad \text{(Eq 5.3)}$$

where the P_{oi} refers to the rain gauge values while P_{si} refers to the values given by the satellite for an event *i*. N denotes the daily event number. The Bias for the rainfall products was also calculated. Bias gives an indication of the systematic differences of the rainfall values obtained by the satellite and the values observed at the gauge. It can be calculated using (Eq 3.34)

$$Bias = \frac{\sum_{i=1}^{N}(P_{oi} - P_{si})}{N} \qquad \text{(Eq 5.4)}$$

The Root Mean Square Error is given by (Eq 5.5). This statistical analysis indicates the differences between the satellite retrieved rainfall values and those observed in the *in-situ* data. The residual errors between the two datasets can be obtained by this statistical analysis.

$$RMSE = \left[\frac{\sum_{i=1}^{N}(P_{si} - P_{oi})^2}{N} \right]^{1/2}$$
(Eq 5.5)

The resulting statistics of rainfall data analysis on a daily basis can be summarised in *Table 5-1*. Random errors in the amount of rainfall detected can be minimised when the satellite data is analysed for a longer duration. The performance of the rainfall data was also evaluated on a monthly basis to minimise the error that were evident when the daily performance of the satellite rainfall products are compared the *in-situ* data. The results of this assessment can be seen in *Table 5-2*.

Table 5-1: Daily performance analysis results summary

	Naivasha DO		North Kinangop		Gilgil Kwetu Farm		Kijabe Farm		New Holland Flowers		Nakuru Meteorological	
	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS
POD (-)	0.426	0.758	0.294	0.897	0.273	0.871	0.345	0.775	0.222	0.857	0.275	0.909
FAR (-)	0.764	0.783	0.588	0.663	0.434	0.534	0.538	0.623	0.519	0.561	0.437	0.520
R. Bias (-)	-0.108	-0.071	-0.383	-0.301	-0.459	-0.393	-0.246	-0.229	-0.435	-0.454	-0.466	-0.339
RMSE (mm/day)	7.481	7.525	7.904	6.634	8.121	7.372	6.970	6.115	7.319	6.783	7.896	7.105
Bias (mm/day)	-0.21	-0.136	-1.071	-0.869	-1.379	-1.181	-0.515	-0.48	-1.371	-1.430	-1.362	-0.992

Table 5-2: Monthly performance analysis results summary

	Naivasha DO		North Kinangop		Gilgil Kwetu Farm		Kijabe Farm		New Holland Flowers		Nakuru Meteorological	
	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS	TRMM	CHIRPS
POD (-)	0.993	1	0.949	1	0.986	1	0.979	1	0.951	1	0.986	1
FAR (-)	0.04	0.032	0.037	0.032	0	0	0	0	0.021	0.022	0.007	0.006
R. Bias (-)	-0.11	-0.07	-0.398	-0.319	-0.456	-0.391	-0.243	-0.228	-0.381	-0.534	-0.475	-0.338
RMSE (mm/month)	1.96	2.333	2.77	2.646	2.682	2.5	1.609	1.56	2.957	3.759	2.747	2.584
Bias (mm/month)	-0.215	-0.136	-1.153	-0.949	-1.366	-1.172	-0.509	-0.478	-1.41	-1.977	-1.387	0.986

From *Table 5-1*, it can be seen that CHIRPS rainfall has the highest probability of detection compared to TRMM3B42v7 at all the stations. The highest probability value for CHIRPS rainfall was at 0.909 and this was seen at the Nakuru Meteorological station while TRMM had 0.426 as the highest probability. This was a huge difference and could be attributed to the algorithms used in derivation of the satellite rainfall or the data obtained from the ground station in preparing the CHIRPS rainfall. The huge differences seen in POD performance is not seen in FAR analysis of the different rainfall stations. CHIRPS rainfall has a higher FAR then TRMM3B42v7 and this means the latter shows more detection of rainfall recorded at the gauge. The difference of the FAR between two products per station is also low and this shows that the two products are able to detect rainfall at almost the same accuracy when compared to the amount recorded at the gauge.

Systematic differences between the gauge and the satellite data represented by the relative bias show very low values between the two readings. It is negative value in all the stations suggesting that the amounts recorded by the satellites are generally below those once recorded by the rain gauges. The RMSE difference between the two satellite products is less than 2 millimetres for daily performance of all the stations in the basin with Kijabe Farm stations having the lowest difference.

The results of monthly aggregation shows an improved correlation of satellite data to the *in-situ* data. Because of the monthly aggregation, the POD and FAR gave near perfect results of almost 1 and 0 respectively. The RMSE of both TRMM3B42v7 and CHIRPS improved significantly for all the stations. CHIRPS had the lowest monthly RMSE values of the two satellite rainfall products except at the New Holland station. This can be attributed to the high spatial resolution of the CHIRPS rainfall. It is noted that the monthly relative bias and the bias did not change significantly.

5.2. FEWSNET PET

For this study, global evapotranspiration from FEWSNET was used. The FEWSNET PET downloaded was prepared using ILWIS where the global product was subset to the area of interest (AOI) which is the Lake Naivasha Basin. This data was then converted to the required ASCII format for input into the model. There was no representative *in-situ* evaporation data from the basin that could be used for performance analysis of FEWNET PET. The only *in-situ* data available was from the lake station which was not representative enough hence the use of satellite products.

5.3. HydroSHEDS

Topographic data was extracted from the HydroSHEDS platform of the USGS. The data used is the Shuttle Radar Topography Mission (SRTM) digital elevation model with a resolution of 90 m. Other elevation products used included the Flow Accumulation Map and Flow Direction maps which are inputs of the CREST model.

ArcMap was used in processing these elevation products. A subset of these products for the Lake Naivasha basin was created and these were later converted to ASCII for use in the CREST model. The Flow direction Map gives the basis of the routing in the basin. Water flow assumes the deepest areas in a cell and water will flow from the highest to the lowest points in the cell. Based on the elevation difference of the cells, the routing of the flow is created. The Flow accumulation map gives the model the ability to simulate the saturation scenario of interconnected cells. In the CREST model, cells that are on the lower ends of the routing channel are saturated first and this simulation moves up the routing channel to cells located away from the ones in the lowest elevations. These files are placed in the Basics folder during the setup of the model.

5.4. CREST model results

5.4.1. CREST model calibration results

The CREST model was calibrated for the three sub-basins of Lake Naivasha basin i.e. Malewa, Karati and Gilgil. A CSV file is generated after each model run. This is stored in a results folder which is created during the model setup stage. A sample of the contents of CREST model results file contents is shown in *Figure 5-1*.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	DateTime	Rain	PET	EPot	EAct	W	SM	RS	RI	ExcS	ExcI	R	RObs
2	1/1/2002 0:00	0.463	4.51	3.934	0	0	0	0.037	0	0	0	0.008	0.918
3	1/2/2002 0:00	2.622	3.62	3.158	0.253	0.012	0	0.036	0	0.001	0	0.041	1.024
4	1/3/2002 0:00	0.473	4.01	3.498	0.002	0.011	0	0.035	0	0	0	0.103	1.136
5	1/4/2002 0:00	0.291	3.87	3.376	0.001	0.009	0	0.034	0	0	0	0.166	1.117
6	1/5/2002 0:00	0.534	3.76	3.28	0.001	0.009	0	0.032	0	0	0	0.22	1.117
7	1/6/2002 0:00	0.398	3.92	3.42	0.001	0.008	0	0.03	0	0	0	0.262	1.234
8	1/7/2002 0:00	0.007	3.93	3.428	0.001	0.007	0	0.027	0	0	0	0.291	1.234
9	1/8/2002 0:00	1.698	4.11	3.585	0	0.006	0	0.025	0	0	0	0.307	1.243

Figure 5-1: Sample of CREST model results file

Important CREST outputs for this research include rain (Rain), simulated runoff (R) and the observed runoff (Robs). A comparison of the model simulated runoff was done to the runoff observed at the outlet stations of each sub-basin. In analysing the results of the CREST model, the simulated flows were plotted against observed flows and the rainfall received in the sub-basins. The calibration period was 2001 to 2010 with 2001 being used as the warmup period for each basin. The validation was done for the years 2011 to 2013.

5.4.1.1. Malewa basin CREST model results
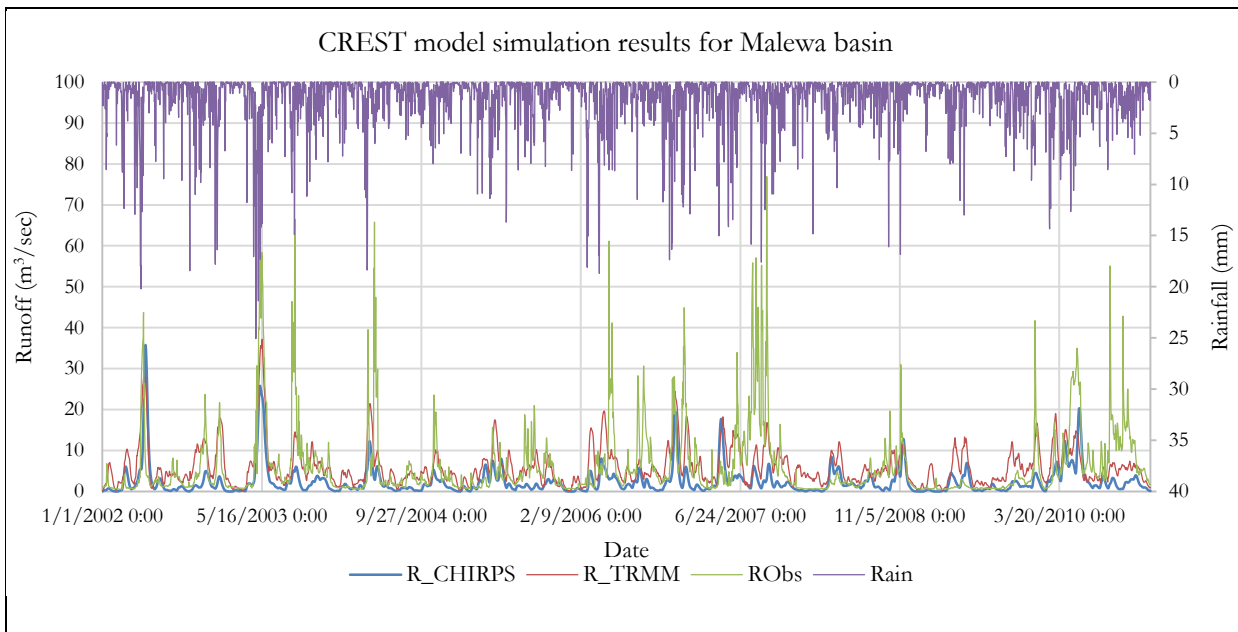
Figure 5-2: Malewa CREST model results

A visual inspection of the CREST model calibration results for Malewa basin (*Figure 5-2*) show that the two satellite products perform differently in simulating the discharges. In the first 3 years of simulation, there is a good correlation between satellite products simulations and the observed discharges. The highest peaks in the rainfall hydrograph are well simulated by the peaks of both satellite products and the observation data as well. However, TRMM3B42v7 shows higher peaks compared to the CHIRPS data.

The difference between the trends of the peaks between the three hydrographs increased as the simulation period increased. Between the 4th and the 8th year of simulation, there are mismatches of peaks projected by the two satellite products when compared to the observed values. This is attributed to the low quality of the streamflow data.

The two satellite products are able to adequately simulate the base flow but CHIRPS has a lower performance compared to TRMM3B4v7. Generally, the two satellite products fail to capture the high peaks of the observed flows during the simulation period. The stream flow observations trends do not adequately capture the peaks of the rainfall hydrograph after the 4th year of simulation. The time taken to rise for simulation of the CHIRPS data also seems to be low most of the time as the peaks it simulates come a little later after the peaks simulated by the TRMM3B42v7 hydrograph. A summary of the objective functions performance of the two satellite forcing is as shown in *Table 5-3*.

Table 5-3: Summary of performance of TRMM3B42v7 and CHIRPS in Malewa basin

Objective function	TRMM	CHIRPS
NS	0.283	0.036
Bias (%)	-2.374	-59.306
Correlation Coefficient	0.533	0.465

Statistically, TRMM3B42v7 had a better *NS* efficiency performance than CHIRPS. CHIRPS also had a lower bias and correlation coefficient value compared to TRMM3B42v7 Gilgil sub-basin CREST model results.
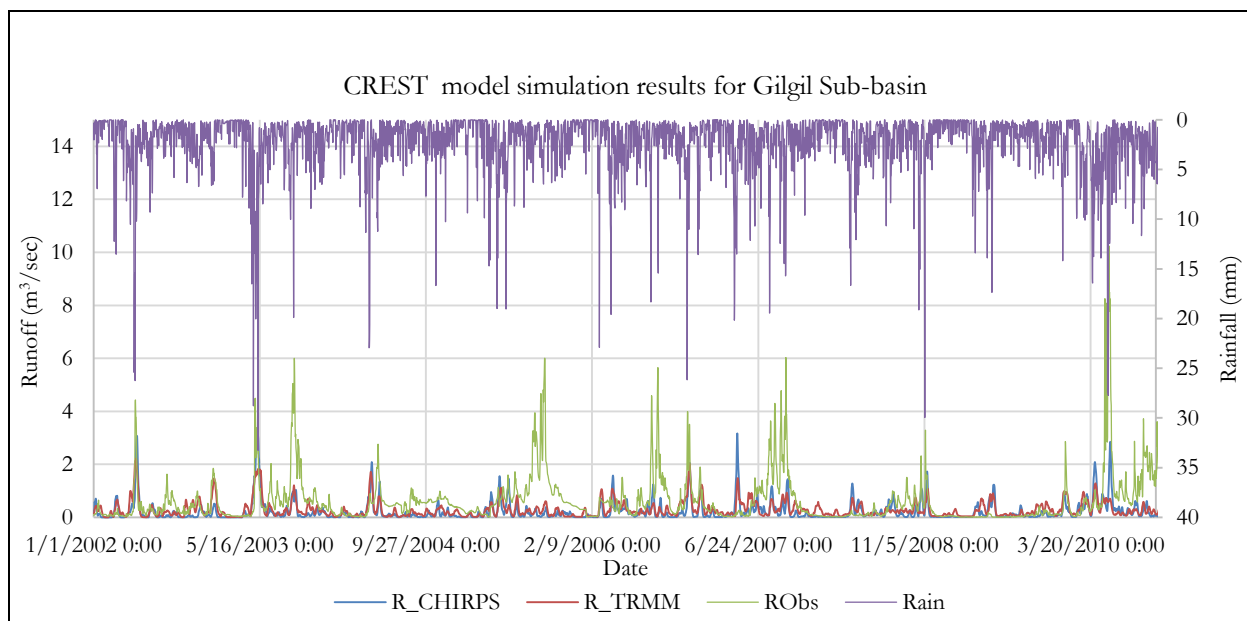
Figure 5-3: Gilgil CREST model results

Figure 5-3 shows the CREST model graphs of the Gilgil sub-catchment. The graph shows that the two satellite products were unable to simulate the observed discharges adequately. The mismatches between the observed and simulated flows are evident along the whole time series but more pronounced after the 3rd year of simulation. The peaks of both TRMM3B42v7 and CHIRPS showed good correlation in the first 3 years of simulation even though they were still below the peaks of the observed discharges. Both the products are able to simulate the base flow adequately but they underperform when simulating the peak flows.

A correlation of the observed discharges and the rainfall data also show major discharges since the effects of bigger rainfall events are not evident in the observed discharges. A visual inspection of the correlation between the rainfall hydrograph and the discharge simulated by the satellite products however, show there is a close match between their peaks. They correlate better compared to the correlation between the observed runoff and the rainfall hydrograph. TRMM3B42v7 and CHIRPS also have a better correlation in terms of peaks and trough in this basin. It can then be concluded that the observed stream flows being used in this basin are of poor quality and caution should be taken when applying them for running a model. *Table 5-4* shows a summary of performance of the objective functions of the CREST model.

Table 5-4: Summary of performance of TRMM3B42v7 and CHIRPS in Gilgil basin

Objective function	TRMM	CHIRPS
NS	0.111	0.102
Bias (%)	-7.378	-18.987
Correlation Coefficient	0.340	0.364

It can be seen from *Table 5-4* that TRMM3B42v7 performs better than CHIRPS in terms of Nash Sutcliff efficiency indicator. The bias of TRMM3B42v7 is also lower compared to the bias shown by CHIRPS. However, in terms of correlation coefficient, CHIRPS simulated discharge has a higher correlation coefficient compared to the TRMM3B42v7 values.

5.4.1.2. Karati sub-basin CREST model results
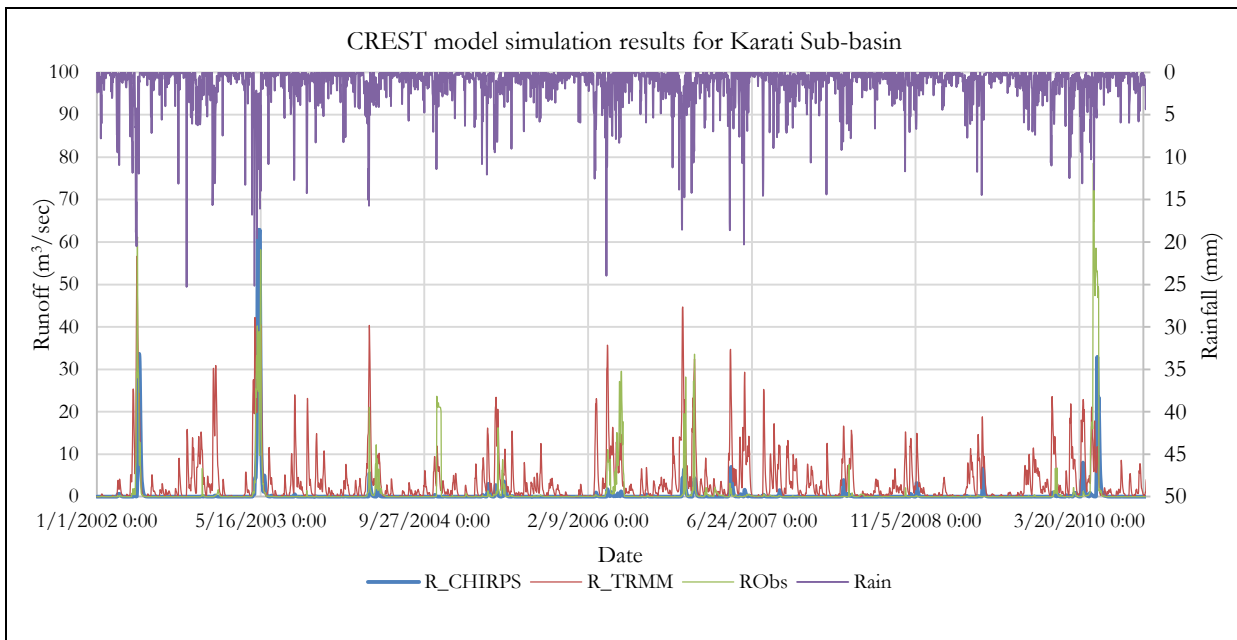
Figure 5-4: Karati CREST model results

From the CREST model simulation of the Karati sub-basin (*Figure 5-4*), it can be seen that the basin experiences very low discharges most of the time. Even though there are a few peaks seen in the trends, these are only seen in durations of high rainfall received in the basin. The results of the simulation also suggests there are a number of missing peaks from the rainfall hydrographs that are not seen in the simulated runoffs. There are instances where TRMM3B42v7 simulated very high peaks that are not simulated by CHIRPS. The TRMM3B42v7 product also shows high runoff compared to CHIRPS too. A statistical summary of the calibration is shown in *Table 5-5*.

Table 5-5: Summary of performance of TRMM3B42v7 and CHIRPS in Karati basin.

Objective function	TRMM	CHIRPS
NS	-0.490	0.175
Bias (%)	240.546	-50.509
Correlation Coefficient	0.405	0.464

Summary of optimum calibration parameters values for the two rainfall SREs is as shown in *Table 5-6*.

Table 5-6: Optimum parameter values after PEST calibration

	CHIRPS			TRMM		
Parameter	Malewa	Gilgil	Karati	Malewa	Gilgil	Karati
W0	0.0001	0.5	0.001	0.0001	0.5	0.001
SS0	0.1	1.0	0	0.1	1.0	0
SI0	2500	1.68	0	2500	1.68	0
Ksat	39.179	44.075	0.471	39.179	44.0752	0.471
RainFact	0.924	0.937	0.531	0.536	0.781	0.444
WM	26.671	25.386	59.847	26.671	25.386	59.847
B	0.102	1.450	0.336	0.102	1.450	0.336
IM	0.042	0.051	0.001	0.041	0.051	0.001
KE	0.872	1.021	0.924	0.092	0.701	0.123
CoeM	85.059	79.112	96.238	85.059	79.112	96.238
ExpM	0.534	0.798	0.855	0.867	0.81	0.855
CoeR	1.00	1.897	2.004	2.120	2.0	2.937
CoeS	0.045	0.131	0.5	0.050	0.010	0.050
KS	0.366	0.864	0.521	0.757	0.7	0.533
KI	0	9.0E-4	0	0	9.0E-4	0

In further analysis of the performance of CREST rainfall forcing inputs, *in-situ* rainfall values were averaged and used as model input to assess their performance when compared to satellite rainfall data as model input. The *in-situ* data was prepared by summing the daily values of all the stations in the basin and averaging them so as to get a single daily value for model input. This data was then converted to ASCII format for use in the CREST model. The resulting model performance statistics can be summarised as shown in *Table 5-7*.

Table 5-7: Results of using different rainfall data sources as CREST model inputs.

Basin	Malewa			Gilgil			Karati		
Performance indicator	*In-situ*	CHIRPS	TRMM	*In-situ*	CHIRPS	TRMM	*In-situ*	CHIRPS	TRMM
NS	0.172	0.184	0.263	-0.299	0.048	-0.227	0.049	0.171	0.246
Bias (%)	-38.303	-34.058	-6.680	-88.226	-31.236	-79.757	-90.025	-76.143	-32.922
CC	0.586	0.511	0.518	0.437	0.310	0.333	0.4604	0.463	0.502

The results show that the model performance improved when satellite data is used as model input in all the basins. The *NS* efficiency performance indicator of the model improved for all the sub-basins when the daily average rainfall is substituted with the satellite rainfall for model input. This concludes that the good spatial-temporal attributes of satellite retrieved rainfall estimates is an added advantage for their use in running hydrological models, even though the improvement was not very large in this case.

5.4.1.3.   Short term simulation

The results of the 10 year calibration period were not satisfactory as they resulted to low *NS* efficiency values and this has an impact on the simulated discharged of the CREST model. This can be attribute to the low quality of the streamflow observations that were used which had a negative impact on the calibration results. The model calibration period was reduced to small time periods as a means to contain the uncertainties brought about by using a longer duration for calibration.

In this instance, the shorter calibration windows considered spanned between 2001-2004 (4yrs), 2005–2008 (3yrs) and 2009-2010 (2yrs) with the first twelve months of each window being used as a warmup period for the model in the first two windows and the first six months for the 2009-2010 window. The performance of this simulations can be summarised in *Table 5-8* for TRMM3B42v7 inputs and *Table 5-9* for CHIRPS input.

Table 5-8: Performance results of reduced model simulation period (TRMM3B42v7). Values in bold show the best performance.

Basin	Malewa			Gilgil			Karati		
Performance indicator/Period	2001-2004	2005-2008	2009-2010	2001-2004	2005-2008	2009-2010	2001-2004	2005-2008	2009-2010
NS	**0.511**	0.073	0.204	**0.300**	0.025	-0.090	-0.101	-3.462	**0.043**
Bias (%)	4.219	-10.044	-21.688	-27.634	-34.711	-65.749	181.961	477.447	47.707
CC	0.722	0.323	0.493	0.592	0.305	0.331	0.655	0.336	0.289

Table 5-9: Performance results of reduced model simulation period (CHIRPS). Values in bold show the best performance.

Basin	Malewa			Gilgil			Karati		
Performance indicator/Period	2001-2004	2005-2008	2009-2010	2001-2004	2005-2008	2009-2010	2001-2004	2005-2008	2009-2010
NS	**0.224**	-0.183	-0.077	**0.152**	-0.082	0.039	**0.189**	0.067	0.172
Bias (%)	-55.724	-63.695	-63.398	-49.908	-52.538	-66.867	-29.373	-66.702	-70.952
CC	0.597	0.255	0.516	0.544	0.306	0.523	0.588	0.304	0.468

From the results, it can be seen that the model performance improved with shorter simulation periods. The shorter simulation periods have an effect on the systematic errors. The shorter period tend to reduce the effects of accumulated systematic errors.

However, the period 2005-2008 showed a dismal performance compared to other windows and this can be attributed to the poor rainfall and discharge data during this period. The lack of quality streamflow input data had an impact on the final performance of the model discharges.

5.4.2.   CREST model validation results

Validation is done to assess the performance of the model calibration results. During this process, all the model inputs and calibrated parameters are left intact and only the meteorological forcing are replaced with those of the validation period. Meteorological forcing data include the rainfall and potential evapotranspiration. The validation period for this study was done for 3 years after the calibration period i.e. (2011-2013). Validation hydrographs are as shown in *Figure 5-5* and the resulting statistical summary is as seen in *Table 5-10*.
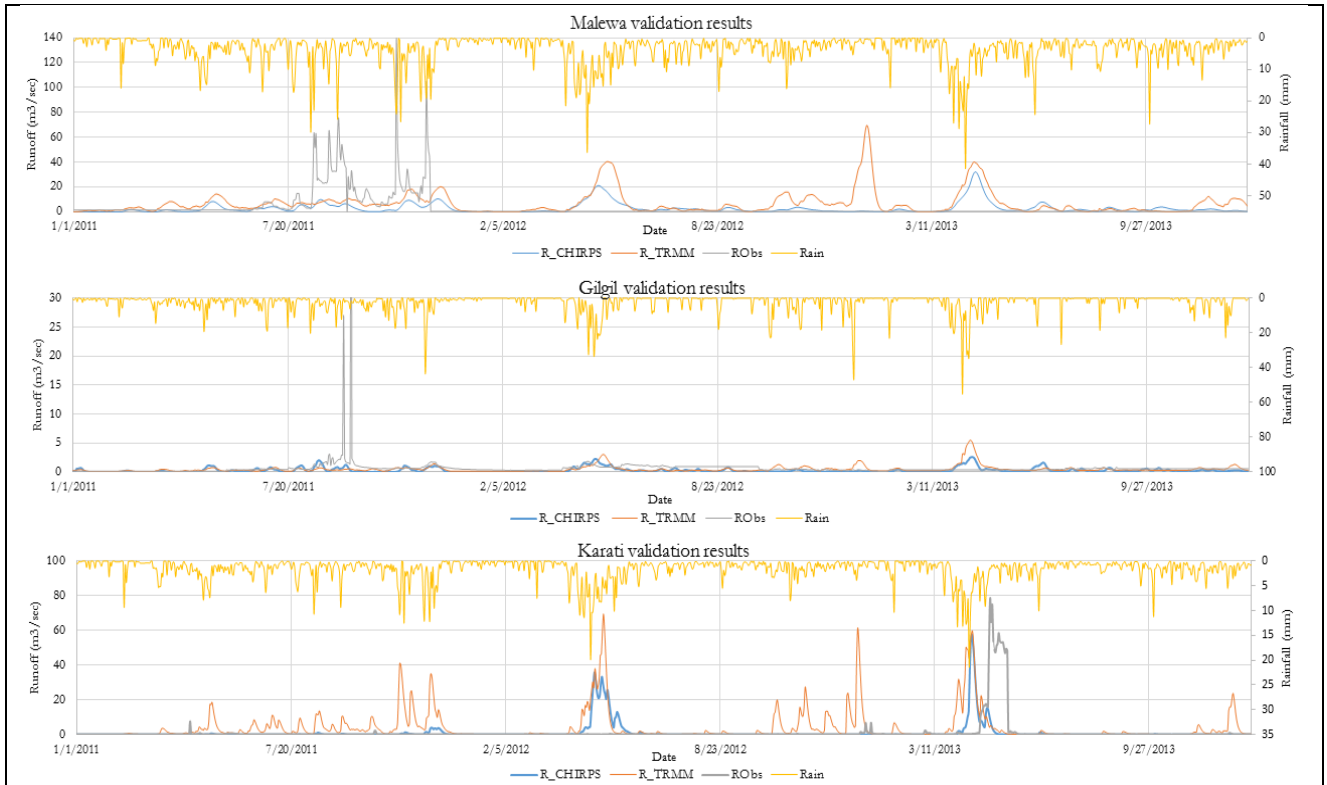
Figure 5-5: Validation hydrographs.

Visual inspection of the validation hydrographs of Malewa sub-basin show a good correlation between the satellite products simulation and the rainfall data but only for a few months into the simulation period. This performance reduces as the simulation continues and there are mismatches between the rainfall hydrographs and the observed hydrographs. The two satellite products, however, show good correlation in simulating the base flows and some peaks. The mismatches evident in the graphs are attributed to the low quality streamflow and *in-situ* rainfall inputs used in validating the model.

In the Gilgil sub-basin, there is a good correlation between the base flows of all the products in the first one year of simulation. It is also seen that there are peaks simulated by the observed data with no causative rainfall match. The two satellite products show good correlation to each other with both of them simulating the peaks simultaneously in mod instanced in this sub-basin. The peaks seen in the rainfall hydrograph are also seen in the satellite rainfall hydrographs.

In the Karati sub-basin, the performance of CHIRPS simulated and observed stream flows is very low. CHIRPS does not simulate the peaks well as compared to TRMM3B42v7. TRMM3B42v7 rainfall tends to give a good simulation of the peaks in this basin even though there are delays between the peaks time in some parts during the simulation period. As the last year of simulation, all the products miss the peaks shown by the rainfall hydrograph. A summary of the performance statistics is shown in *Table 5-10*.

Table 5-10: Validation results summary.

	TRMM3B42v7			CHIRPS		
	Malewa	Gilgil	Karati	Malewa	Gilgil	Karati
NS	-0.87	-0.067	-1.702	-0.096	-0.026	-0.405
Bias (%)	94.118	-27.726	254.958	-17.324	-45.835	-0.758
CC	0.053	0.017	-0.004	0.114	0.063	0.06

The Gilgil sub-basin shows a better performance of the validation with the highest values of *NS* efficiency in both TRMM3B42v7 (-0.067) and CHIRPS (-0.026) data. In the Malewa basin, CHIRPS shows a better

NS efficiency as compared to TRMM342v7. TRMM3B42v7 has the lowest performance of all the simulations in terms of Nash-Sutcliff efficiency. All the statistics obtained after the validation period all show differences from those obtained during the calibration period. There are also bigger differences between the bias validation results of all the products when compared to the same bias obtained during the calibration period with bias values of validation being higher than those obtained during calibration.

The differences seen between the calibration and the validation results can be attributed to the low quality discharge data obtained for this study. The amount of data was also limited with huge gaps evident in the streamflow data and this was more pronounced in the validation period.

5.5.   Abstraction data integration and scenario analysis

5.5.1.   Abstraction data integration

An analysis of the abstraction records retreived from WRMA PDB shows that the Malewa basin has the most number of surface water abstraction in LNB. The abstractions from the Lake and its aquifer were catered for in the Lake balance model (Van Oel et al., 2013). Abstractions from the river channels can be summarised as shown in *Figure 5-6*, *Figure 5-7* and *Figure 5-8*.

In this study, only the legal streamflow abstractions from the WRMA PDB were considered. The cancelled and deferred permits were also eliminated since they cannot be ascertained with accuracy. It can be seen that most of the abstraction by volume are in category C and D permits. This means that a bulk of the river flows are allocated for either for water supply or irrigation.
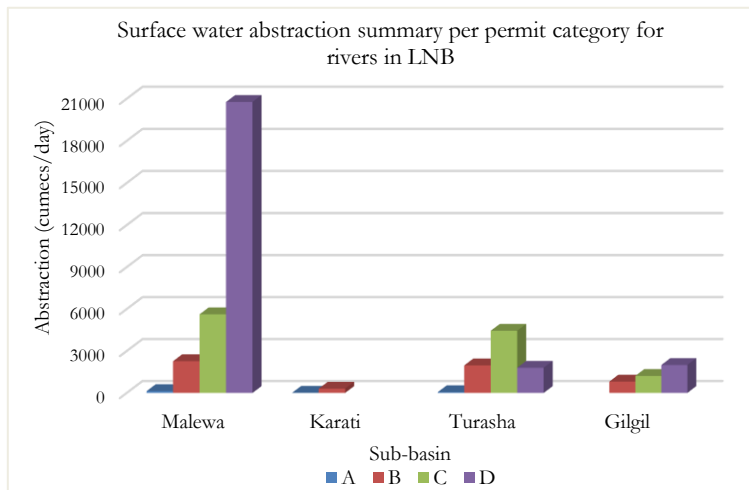
Figure 5-6: Surface water allocations summary for LNB rivers

Figure 5-6 shows a summary of surface water abstractions in the LNB rivers. Malewa sub-basin has the most number of surface water abstractions. Class D permits are the leading allocations in this basin. Karati sub-basin has the lowest allocation followed with only class A and B permits. Turasha sub-basin rivers which drains into Malewa river has more abstractions compared to Gilgil sub-basin with class C leading in Turasha compared to class D in Gilgil sub-basin. It can also be noted that the Karati river does not have class C and D permits allocated and this could be because of its ephemeral nature

Figure 5-7: Ground water allocation summary for LNB Sub-basins

Figure 5-7: shows a bar graph presentation of legal boreholes in the Lake Naivasha basin. Malewa sub-basin has the most number of boreholes with class D boreholes dominating the abstractions. There are no class A boreholes in the basin since all class A abstractions are done from the rivers. It is noted that Karati sub-basin does not have ground water allocations in the records received but there are othe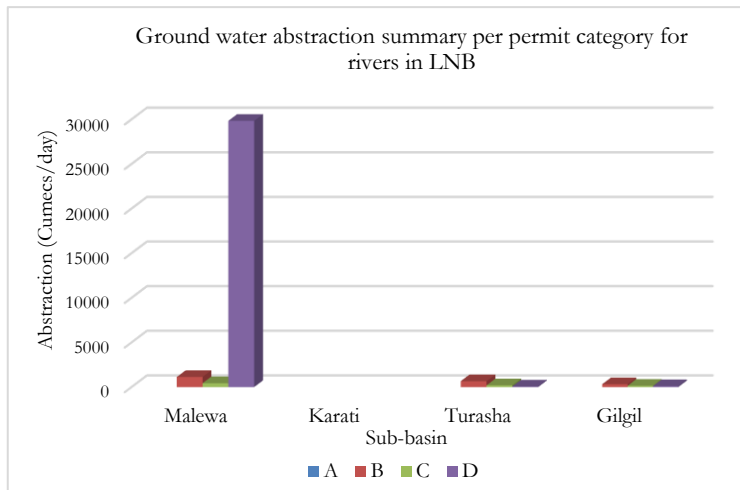r records which show old boreholes whose daily abstraction limits are not indicated so they were not considered for this study.
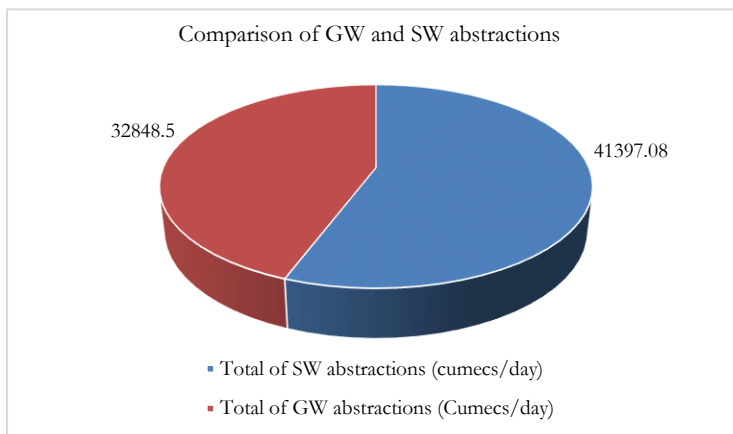
Figure 5-8: Difference of GW and SW abstractions in LNB in cumecs per day.

Figure 5-8 shows a comparison of the Ground water and river water abstractions. It shows surface water abstraction supersede ground water abstractions by 12% per day. This difference is a total of 8909.5 m³/day. It can be concluded that the surface water abstractions are more important to the inhabitants the 3 sub-basins.

The results of the legal daily limit of permitted abstractions for the three basins can be summarized in *Table 5-11*. Only the surface water abstractions from the rivers were used to assess the effects of abstractions from rivers draining into it.

Table 5-11: Summary of legal abstractions and their daily limits.

	Permit category	Malewa	Karati	Turasha	Gilgil	Totals (m³/day)	Monthly (30 days) (m³/month))
Ground	A					0	0
	B	1124.8		653.65	313	2091.45	4182.9
	C	433.5		194.55	143	771.05	1542.1
	D	29826		60	100	29986	59972
Surface	A	131.9	12.6	51.95		196.45	5893.5
	B	2273.43	305.79	1967.83	807.31	5354.36	160630.8
	C	5620.79		4444.65	1213.45	11278.89	338366.7
	D	20767.38		1800	2000	24567.38	737021.4

Ground water abstractions from the lake aquifer and the surface water abstractions from the lake itself have been included in the Lake balance model. A scenario analysis of the different abstraction limits that can be applied to ensure sustainable lake levels can be shown in section 5.5.2.

5.5.2. Scenario analysis

A relationship between the lake level, its volume and the amount of inflow into the lake was established in the lake balance model. This was based on a bathymetric survey of the lake. In this study, this relationship will be used to assess the performance of the CREST model simulated stream flows and how this can be used for effective river abstractions management. An analysis of the resilience of the lake levels to abstraction scenarios is conducted here.

Firstly, a correlation and *NS* efficiency analysis of the volumes of TRMM3B42v7 and CHIRPS monthly streamflow values was done with the observed monthly stream flows. In the lake balance model, the simulations are done on a monthly basis and the values of stream flows used in calibrating the lake model were also compared to CREST model calibration streamflow results. Secondly, the best performing satellite product streamflow simulation was used for the resilience test. The comparison period is between the years 2002 and 2010. *Table 5-12* shows the results of the correlation analysis.

Table 5-12: Table showing monthly correlation coefficient and NS efficiency values for different streamflow sources

Streamflow source correlation	R value	NS
TRMM3B42v7 to CHIRPS	0.845	0.002
TRMM3B42v7 to Observed	0.658	-1.23
TRMM3B42v7 to Lake balance model	0.580	-6.35
CHIRPS to Observed	0.639	-3.71
CHIRPS to Lake balance model	0.669	0.06
Observed to Lake balance model	0.869	0.17

The correlation analysis shows TRMM3B42v7 has the lowest correlation of monthly stream flow to those used to prepare the lake balance model with an *R* value of 0.580 while CHIRPS had a better *R* value of 0.669. The simulated stream flows from the two satellite products had a monthly correlation value of 0.845. The *R* value for TRMM3B42v7 to Observed stream flow was also lower than that of CHIRPS to the observed stream flows. CHIRPS also had a better *NS* efficiency value of 0.06 comapred to TRMM3B42v7 which had a NS efficiency value of -6.35 when compared against those values used for calibrating the Lake balance model.

The observed stream flows had the best correlation to the values used for the lake balance model at 0.869. CHIRPS had a correlation value of 0.699 to the lake balance model stream flows and this was better that that of the TRMM3B42v7 correlation which was at 0.580. With this, CHIRPS mean monthly stream flow for the years 2002-2010 was used to test the resilience of Lake Naivasha levels to abstractions along the rivers draining into it. The average monthly inflow was 7086528.8 (m³/month).
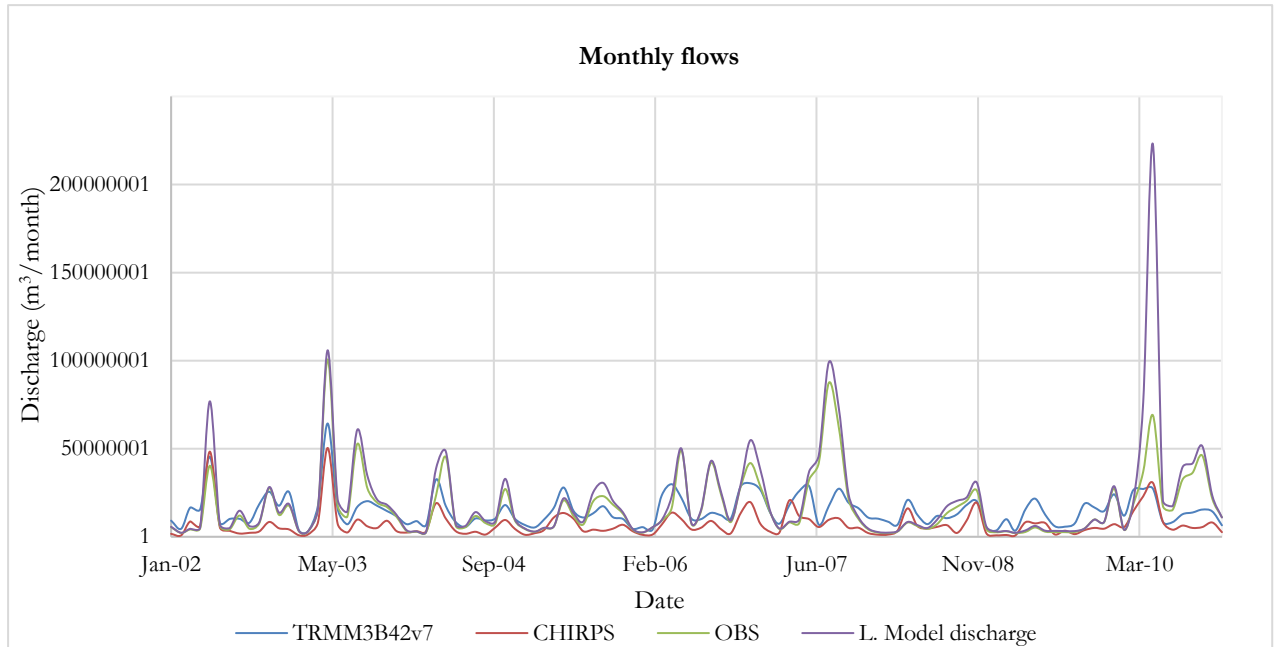
Figure 5-9: Monthly discharge comparison of satellite products, Observed and values used to calibrate the lake model

The monthly trends (*Figure 5-9*) show close relations between the peaks of all the discharges even though values used to calibrate the lake model discharge show higher peaks than the other three. The correlation of the peaks is good in the first three years of simulation but this changes between the periods 2005 to 2008. A comparison of the two satellite products in the periods 2005 to 2008 show marching peaks even though the peaks produced by CHIRPS are lower than those produced by TRMM3B42v7. The correlation between the peaks of the four discharges improves as the simulation continues but they are not satisfactory. The discharge values used to calibrate the lake balance model are still higher than those shown by the other three discharges.

Variation of Class A permits was capped at 50%, class B at 75% while class C and D were varied up to 100%. In carrying out the resilience test, the December, 2010 lake level was used as the initial lake level while the initial inflow were the monthly average inflows during the wet year, the dry year and the CHIRPS monthly discharge as simulated by the CREST model. The results of the permit variation by reducing the percentage of abstraction by class is as is shown in *Table 5-13* while the results of increases the permit allocations by varying the abstraction limits by class is as shown in *Table 5-14*. A definition of the scenarios is also shown below *Table 5-14*. The total inflows (m³ month⁻¹) and the area of the lake (m²) using the monthly wet year, monthly dry year and monthly average inflows of CHIRPS is also shown below *Table 5-14*.

Table 5-13: Permit variation results for reducing the % of abstraction based on the permit class, wet year, dry year and Average CHIRPS values.

	Permit Class and % reduction				Total Abstraction (m³/month)	Wet year (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l)	Dry year (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l.)	Average (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l)
	A	B	C	D										
Scenario 1	0	0	0	25	1057656.25	9016332	+5.21	1886.58	2792781.35	+5.37	1886.54	6028873	+5.29	1886.56
Scenario 2	0	0	25	50	788809.23	9285179	+3.89	1886.58	3061628.38	+4.01	1886.54	6297720	+3.94	1886.56
Scenario 3	0	25	50	75	399489.7	9674498	+1.97	1886.59	3450947.9	+2.03	1886.54	6687039	+2	1886.57
Scenario 4	25	50	75	100	335563.4	9738425	+1.65	1886.59	3514874.2	+1.7	1886.54	6750965	+1.68	1886.57
Scenario 5	50	75	100	100	43104.25	10030884	+0.21	1886.59	3807333.35	+0.22	1886.55	7043425	+0.22	1886.57

Table 5-14: Permit variation results for increasing the % of abstractions based on the permit class, wet year, dry year and Average CHIRPS values.

	Permit Class and % increase				Total Abstraction (m³/month)	Wet year (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l)	Dry year (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l)	Average (m³/month)	Change in Area (ha.)	Lake level (m.a.s.l)
	A	B	C	D										
Scenario 6	0	0	0	25	1426166.95	8647821	-7.04	1886.56	2424270.65	-7.25	1886.54	5660362	-7.64	1886.56
Scenario 7	0	0	25	50	1695013.98	8378974	-8.37	1886.56	2155423.63	-8.63	1886.53	5391515	-8.49	1886.56
Scenario 8	0	25	50	75	2004018.5	8069970	-9.9	1886.57	1846419.1	-10.21	1886.53	5082510	-10.05	1886.55
Scenario 9	25	50	75	100	2314496.4	7759492	-11.45	1886.57	1535941.2	-11.8	1886.53	4772032	-11.61	1886.55
Scenario 10	50	75	100	100	2440718.95	7633269	-12.08	1886.57	1409718.65	-12.44	1886.53	4645810	-12.25	1886.55

where

Scenario 1	-	Reduce Class D by 25%
Scenario 2	-	Reduce Class C by 25% and D by 50 %
Scenario 3	-	Reduce Class B by 25%, C by 50% and D by 75%
Scenario 4	-	Reduce Class A by 25%, B by 50 %, C by 75% and D by 100%
Scenario 5	-	Reduce Class A by 50%, B by 75 %, C by 100% and D by 100%
Scenario 6	-	Increase Class D by 25%
Scenario 7	-	Increase Class C by 25% and D by 50 %
Scenario 8	-	Increase Class B by 25%, C by 50% and D by 75%
Scenario 9	-	Increase Class A by 25%, B by 50 %, C by 75% and D by 100%
Scenario 10	-	Increase Class A by 50%, B by 75 %, C by 100% and D by 100%

Total inflow and lake area of the years considered CHIRPS

	Inflow (m³/month)	Area (m²)
Wet year	10073988	147215285.1
Dry year	3850438	146904535.2
Average	7086529	147067285.7

Initial Lake level: 1886.56 m.a.s.l.

From *Table 5-13* it can be seen that varying the abstractions has an effect in both the lake area and its volume. Varying class D permits has significant effects on the area of the lake. Reducing Class D abstraction by 25% causes an increase of the lake area by 5.21 ha in a wet year, 5.37 ha in a dry year and 5.29 ha when the CHIRPS monthly average inflow are used. The change in lake levels is not as rapid as the change in the size of the lake. A change of 0.01 m.a.s.l is seen during the first 5 reduction scenarios. The reduction of the lake area is also related to the class permit reduction. The impact of reducing abstractions is more pronounced in a dry year than in a wet year. As the lake level increase and reach to higher altitudes in each variation scenario, the corresponding magnitude of change in area is evidently low. This can be attributed to the lake bathymetry where higher lake contours have a larger surface area than the lower contours hence little change in lake level as the inflow increases.

A reverse variation of the permit Classes was done by increasing the amount daily limits i.e. scenarios 6-10, while using the same monthly flows as those used during the reduction of the daily abstractions. The increase of daily limits of the permits resulted to drastic reduction of the lake levels and the volume. The effects of increased allocation is more pronounced during the dry year compared to the wet year and when the monthly average of CHIRPS values are used for the resilience analysis. This means that the allocations should only be varied upwards only when the amount of inflow is increased. Increase in flows is more likely to be experienced during the wet years. A summary of total change in area of lake during the increase and reduction of percentage of abstractions is as shown in *Table 5-15*.

Table 5-15: Summary of change in area due to permit class limit variations.

	Wet year (ha)	Dry year (ha)	Average
Reducing permit limits	5	5.15	5.07
Increasing permit limits	5.04	5.19	4.61

The relationship between the lake area and the volume of water it contains is influenced by the lake bathymetry where if the lake level is low, the corresponding volume is also low but when the lake level is high, the corresponding volume is also high. Since the contours are not perfect cylinders, the lake areal increase and decrease will be evident at different directions. A bathymetric map of Lake Naivasha from a survey carried out in 1983 (Ase, Sernbo, & Syren, 1986) is shown in *Figure 5-10*.
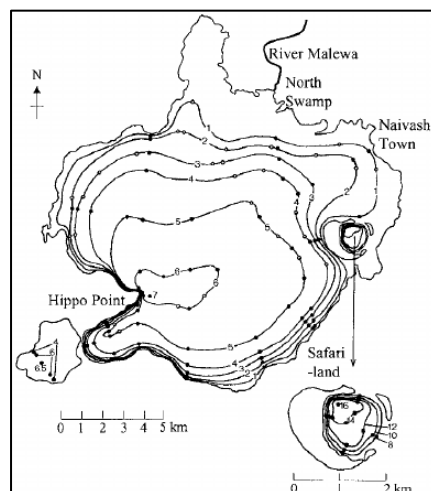

Figure 5-10: Bathymetric map of Lake Naivasha (Ase et al., 1986).

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

The study was aimed at assessing the effect of surface water abstractions to the amount of stream flow into Lake Naivasha and how this affects its volumes and water level. This was done using a distributed hydrological model (CREST) and a lake balance model previously developed for Lake Naivasha. Daily river abstractions limits were obtained from WRMA permit database.

In assessing the abstractions from rivers, the amount of streamflow needs to be quantified on regular basis. This is done by carrying out streamflow gauging using either manually or by use of automatic techniques. Abstraction thresholds can be set after establishing the Q_{80} of stream flows. The Q_{80} is used as a basis for distinguishing between normal flows and flood flows. The flow duration curves were prepared to analyse the flows and it was seen that the Q_{80} values differed significantly in the basin with the highest values being seen in Malewa sub-basin. The lowest discharges were seen in the Karati-sub-basin.

Stream flow data obtained showed gaps and this increased the uncertainties of applying them for use in this study. The presence of gaps called for use of interpolation methods such as the use of rating curves which employ the river stage levels to predict the amount of discharge. The resulting discharges were used to produce double mass curves for the different gauging stations in the basin. The double mass curves showed that the discharge data performed well in three out of four stations in Gilgil sub-basin and five out of nine stations in the Malewa sub-basin. Double mass curves obtained from Karati sub-basin were erratic and less reliable.

In this study, TRMM3B42v7 and CHIRPS rainfall were used as input to CREST model. STRM elevation data was used for DEM and for preparing the flow accumulation and direction maps of the basin. FEWSNET PET was also used as potential evaporation input for the model. The performance of the satellite rainfall products was evaluated using *in-situ* data from six stations spatially distributed in the study area. Cumulative plots of satellite rainfall showed they both underestimated the amount of rainfall as compared to the rain gauges even though the satellite products showed high correlation when compared against each other. Generally, CHIRPS underestimated the total amount of rainfall more than TRMM3B42v7. Daily bias analysis was also carried out for satellite products using *in-situ* data as benchmark. The daily values showed high RMSE for the two products (TRMM3B42v7=7.616 mm/day and CHIRPS=6.922 mm/day). This may be due to the residual errors that were seen in the *in-situ* data. Monthly average RMSE values reduced the bias of both products (TRMM3B42v7=2.454 mm/month and CHIRPS=2.564 mm/month). This represented an average RMSE change of 5.161 mm for TRMM3B42v7 and 4.359 mm for CHIRPS.

These rainfall and discharge data were used for calibration and validation of the CREST model. The calibration of the model for a period of 10 years yielded low performance and this was attributed to the data that was used as input. There were contrasts between hydrographs as the simulated peaks and troughs of the satellite products did not complement peaks produced by the *in-situ* data. Visual inspection showed that the mismatches were more evident after the third year of simulation for Malewa and Gilgil sub-basins. The mismatches were more pronounced in the entire Karati sub-basin calibration period. This influenced the resulting *NS* efficiency with TRMM3B42v7 outperforming CHIRPS in Malewa and Gilgil sub-basin.

CHIRPS outperformed TRMM3B42v7 in Karati sub-basin. CREST model was calibrated for shorter windows to investigate the deteriorating performance seen as the calibration period increased. This procedure showed that the *NS* efficiency values improved with shorter simulation periods but the period 2005-2008 produced low *NS* efficiency values. This can be attributed to the low credibility of discharge and rainfall data obtained for this period. Even though satellite rainfall produced low *NS* efficiency values, a comparison between calibrating the model using *in-situ* data and satellite data showed improvements in the final *NS* efficiency when satellite rainfall is used. Validation of the model gave lower values of objective functions and this was attributed to the huge gaps in the input data during the period 2011-2013.

The volumetric difference between the stream flows produced by the satellite products was used as the basis for choosing the product that produced, "near-real" volumes of monthly runoff. The observed stream flow volumes and those used in calibrating the lake balance model were used as benchmarks for this purpose. CHIRPS produced the best correlation to the observed discharges and those discharged used to calibrate the lake balance model. CHIRPS had a correlation of 0.639 to observed streamflow and 0.669 to the Lake balance model volumes while TRMM3B42v7 had a correlation of 0.658 to observed stream flows and 0.580 to the lake balance model volumes. The correlation between the observed monthly volumes obtained in this study to those used to calibrate the Lake balance model was 0.869. Average monthly streamflow volumes of CHIRPS were used as the total stream flow into the Lake. The resilience was also done using CHIRPS monthly stream flows for a wet and dry year.

The resilience test of the lake area and volume show that it is sensitive to the amount of inflow. The intensity of change in volume and area is also dependent on whether it is a wet or dry year. During the dry years, the impact of varying the permit class limits upwards or downwards is more than during the wet year and when monthly average of CHIRPS inflows are used. The intensity of varying the daily permit limits upwards is more than when the permit limits are varied downwards. The rate of change in the lake area and volume is higher when the permit limits are varied upwards. This was seen when a comparison of the total change in reduction scenarios were compared to total change in increase scenarios was done. The reduced flows can be mitigated by enforcing the 3-month-storage recommended by the Water Resources rules and only allow abstractions during the flood flows for Category C and D permits who are major abstractors. This will aid in easing the pressure on the inflow caused by abstractions during the dry periods.

6.2.   Recommendations

To improve the reliability of the model results, the stream flows data can be improved by updating the rating curves of the stations to include the most recent data that were collected after the calibration period. The rating curve were produced using interpolation methods and this was also a cause of the differences between the observed and simulated stream flows.

The model performance may also be improved by calibrating it for smaller basins sections since using calibration parameters on large basins increases the uncertainties of obtaining the correct values. This can be because different basins have different characteristics of shape and amount of meteorological forcing received.

In testing the resilience of the lake, integration of a code in the CREST model that embeds the abstractions to their actual locations may improve the predictive nature of the effects of the abstractions. The flow durations curves showed that streams flows along sections of rivers had different Q_{80} values and this can be used as a basis of increasing the volume of water that can be abstracted at this sections hence

easing the pressure on those sections with lower Q_{80} values. This can aid in reducing the conflicts between different water users in the basin.

In testing the resilience of the lake to abstractions, only legal abstractions were used as outflows from the streams. There are illegal abstractions with unknown or estimated amounts of daily abstractions. Knowing the actual abstractions of these abstractions will greatly influence the outcome of the resilience of the lake to abstractions.

CREST model has a real-time simulation component and this can used to simulate stream flows at a rapid rate based on accurate inputs. This can improve the timing for variations since varying the inflows has been seen to have an effect on the lake levels and doing it on time before the lake levels are adversely low can help mitigate the challenges that arise when the lake levels are low.

In notifying the public of the variation of the permits, billboards can be erected along river sections. The billboards will be important in notifying the public about abstraction limits based on stream flow amounts in that section of the river.

LIST OF REFERENCES

Ahmed, A. (1999). *Estimating Lake Evaporation Using Meteorological Data & Remote Sensing(A Case Study of Lake Naivasha Central Rift Valley, Kenya).* MSc Thesis, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC)

Alfarra, A. (2004). *Modelling water resource management in Lake Naivasha. International Institute for Geo-information Science and Earth Observation.* MSc Thesis, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC) Retrieved from ftp://192.87.16.103/pub/naivasha/ITC/Alfarra2004.pdf

Armstrong, M. F. (2002). *Water Balance of the Southern Kenya Rift Valley Lakes.* (MSc Thesis), University of Twente Faculty of Geo-Information Science and Earth Observation (ITC). Retrieved from ftp://ftp.itc.nl/pub/naivasha/ITC/Amstrong2002.pdf

Ase, L. E., Sernbo, K., & Syren, P. (1986). *Studies of lake naivasha, Kenya, and its drainage area.* Stockholm Universitet. Stockholm. Retrieved from ftp://ftp.itc.nl/pub/naivasha/Ase1986.pdf

Awange, J. L., Forootan, E., Kusche, J., Kiema, J. B. K., Omondi, P. A., Heck, B., … Gonçalves, R. M. (2013). Understanding the decline of water storage across the Ramser-Lake Naivasha using satellite-based methods. *Advances in Water Resources*, *60*, 7–23. http://doi.org/10.1016/j.advwatres.2013.07.002

Becht, R., & Harper, D. M. (2002). Towards an understanding of human impact upon the hydrology of Lake Naivasha, Kenya. *Hydrobiologia*, *488*, 1–11.

Becht, R., Odada, E., & Higgins, S. (2005). Lake Naivasha: experience and lessons learned brief. *Initiative: Experience and Lessons Learned Briefs* , *5*, 22. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Lake+Naivasha+Experience+and+Lessons+Learned+Brief#0

Bengtsson, L., Bonnet, R. M., Calisto, M., Destouni, G., Gurney, R., Johannessen, J., … Rast, M. (2014). *The Earth's Hydrological Cycle*, 46. Retrieved from http://doi.org/10.1007/978-94-017-8789-5

Bergström, S. (2006). Experience from applications of the HBV hydrological model from the perspective of prediction in ungauged basins. *IAHS-AISH Publication*, 307, 97–107. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-33845590699&partnerID=tZOtx3y1

Beven, K. (2001). How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences [Hydrol. Earth Syst. Sci.].* *5*(1), 1–12. http://doi.org/10.5194/hess-5-1-2001

Bhatti, H., Rientjes, T., Haile, A., Habib, E., & Verhoef, W. (2016). Evaluation of Bias Correction Method for Satellite- Based Rainfall Data. *Sensors*, *16*(6), 884. http://doi.org/10.3390/s16060884

Chen, S., Liu, H., You, Y., Mullens, E., Hu, J., Yuan, Y., … Hong, Y. (2014). Evaluation of high-resolution precipitation estimates from satellites during July 2012 Beijing flood event using dense rain gauge observations. *PloS One*, *9*(4), e89681. http://doi.org/10.1371/journal.pone.0089681

Climate Hazard Group. (2015). *CHG - Data - CHIRPS.* Retrieved December 8, 2016, from http://chg.geog.ucsb.edu/data/chirps/

De Jong, T. (2011). *Water Abstraction Survey in lake naivasha Basin, kenya.* Internship Report, Wagenigen University. Retrieved from ftp://ftp.itc.nl/pub/naivasha/DeJong2011.pdf

Dinku, T., Asefa, K., Hilemariam, K., Grimes, D., & Connor, S. (2011). Improving availability, access and use of climate information. *WMO Bulletin*, *60*(2), 80–86. Retrieved from http://public.wmo.int/en/bulletin/improving-availability-access-and-use-climate-information

Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., & Ropelewski, C. F. (2007). Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing*, *28*(7), 1503–1526. http://doi.org/10.1080/01431160600954688

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., & Ropelewski, C. F. (2008). Validation of high resolution satellite rainfall products over complex terrain. *International Journal of Remote Sensing*, *29*(14), 4097–4110. http://doi.org/10.1080/01431160701772526

Doherty, J., Brebber, L., & Whyte, P. (2005). PEST: Model-independent parameter estimation. *Watermark Computing, Corinda, Australia*, 4, 122. http://doi.org/10.1016/B978-141600119-5.50003-2

Dourojeannni, A. (2001). Water management at the river basin level: challenges in Latin America. *Serie Recursos Naturales E Infraestructura*, ISBN: 72. 92-1-121321-5

FAO. (2003). Review of world water resources by country. *Water Reports 23: Review of World Water Resources by Country*, (23), 127. ISSN: 1020-1023.

FAO. (2011). *The State of the World's land and water resources for Food and Agriculture, Managing systems at risk.*

*Food and Agriculture Organization.* London. http://doi.org/978-1-84971-326-9

Fayos, C. B. (2002). Competition over water resources: analysis and mapping of water-related conflicts in the catchment of Lake Naivasha (Kenya). MSc Thesis, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC)

FEWS-NET. (2006). *Famine Early Warning Systems Network.* Retrieved from https://www.fews.net/

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., … Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data, 2,* 150066. http://doi.org/10.1038/sdata.2015.66

Gathecha, H. M. (2015). *Reconstruction of Streamflow Into Lake Naivasha Using Crest Model and Remote Sensed Rainfall and Evapotranspiration.* MSc Thesis, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC)Retrieved from http://www.itc.nl/library/papers_2015/msc/wrem/gathecha.pdf

Global Finance. (2017). *Kenya GDP country report.* Retrieved from https://www.gfmag.com/global-data/country-data/kenya-gdp-country-report

Government of Kenya. The Water Act. (2002). *Kenya.* Retrieved from http://www.kenyalaw.org/lex//actview.xql?actid=CAP. 372

Government of Kenya. Water Resources Management rules. (2006). *Kenya.* Retrieved from http://www.wrma.or.ke/images/jdownloads/water_resources_rules.pdf

Government of Kenya. Water Allocation Guidelines (2009). *Water Resources Management Authority.* Retrieved from http://www.wrma.or.ke/downloads/FINAL_WATER_ALLOCATION_GUIDELINES.pdf

Harper, D., Nic, P., Caroline, U., Ed, M., & Richard, F. (2013). Water cooperation for sustainable utilization: Lake Naivasha, Kenya. *In Free Flow: Reaching Water Security Through Cooperation,* 256–259.

Molle, F. (2006). Water for Food, Water for Life. *Natures Sciences Sociétés* 16, 274-275. Retrieved from http://doi.org/10.1051/nss:2008056

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology, 5*(3), 487–503. http://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Kenya Flower Council. (2016). Industry Statistics Flower Export Volumes of the Kenya Flower Council. Retrieved August 10, 2016, from http://kenyaflowercouncil.org/?page_id=94

Khan, S. I., Hong, Y., Wang, J., Yilmaz, K. K., Gourley, J. J., Adler, R. F., … Habib, S. (2011). Satellite Remote Sensing and Hydrologic Modeling for Flood Inundation Mapping in Lake Victoria Basin: Implications for Hydrologic Prediction in Ungauged Basins. *IEEE transactions on geoscience and remote sensing, 49*(1). http://doi.org/10.1109/TGRS.2010.2057513

Knoche, M., Fischer, C., Pohl, E., Krause, P., & Merz, R. (2014). Combined uncertainty of hydrological model complexity and satellite-based forcing data evaluated in two data-scarce semi-arid catchments in Ethiopia. *Journal of Hydrology, 519(B),* 2049–2066. Retrieved from http://doi.org/10.1016/j.jhydrol.2014.10.003

Loucks, D. P., van Beek, E., Stedinger, J. R., Dijkman, J. P. M., & Villars, M. T. (2005). Water Resources Systems Planning and Management and Applications: An Introduction to Methods, Models and Applications. *United Nations Educational, Scientific and Cultural Organisation, 51,* 57. http://doi.org/92-3-103998-9

Maathuis, B., Mannaerts, C., Shouwenburg, M., Retsios, B., Lemmens, R., & Nkepu, M. R. (2014). In Situ and Online Data Toolbox Installation, Configuration and User Guide. Retrieved from http://52north.org/files/earth-observation/userguides/isod/ISOD_Toolbox_V1.4UG.pdf

Mason, S. A. (2004). *From Conflict to Cooperation in the Nile Basin.* Doctoral Thesis, ETH Zurich. Retrieved from http://www.css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/From_Conflict_to_Cooperation.pdf

Mbaga, D. P. (2015). *Impact of hydro- meteorological drivers on the streamflow of the usangu catchment impact of hydro-meteorological drivers on the streamflow of the usangu catchment,* MSc Thesis, University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC).

Md Ali, A., Solomatine, D. P., & Di Baldassarre, G. (2015). Assessing the impact of different sources of topographic data on 1-D hydraulic modelling of floods. *Hydrology and Earth System Sciences, 19*(1), 631–643. http://doi.org/10.5194/hess-19-631-2015

Meins, F. M. (2013). *Evaluation of spatial scale alternatives for hydrological modelling of the Lake Naivasha basin, Kenya,* (MSc Thesis). University of Twente Faculty of Geo-Information Science and Earth Observation (ITC).

Miralles, D. G., Gash, J. H., Holmes, T. R. H., De Jeu, R. A. M., & Dolman, A. J. (2010). Global canopy interception from satellite observations. *Journal of Geophysical Research Atmospheres*, 115(16), D16122. Retrieved from http://doi.org/10.1029/2009JD013530

Moazami, S., Golian, S., Kavianpour, M. R., & Hong, Y. (2014). Uncertainty analysis of bias from satellite rainfall estimates using copula method. *Atmospheric Research*, *137*, 145–166. Retrieved from http://doi.org/10.1016/j.atmosres.2013.08.016

NASA. (n.d.). *TRMM Mission Overview | Precipitation Measurement Missions*. Retrieved August 16, 2016, from https://pmm.nasa.gov/TRMM/mission-overview

NASA. (2016). *Precipitation Measurement Missions, Explanation of Data Products*. Retrieved January 28, 2017, from https://pmm.nasa.gov/data-access/data-products

Njuki, S. M. (2016). *Assessment of Irrigation Performance by Remote Sensing in the Naivasha Basin , Kenya,* MSc Thesis. University of Twente Faculty of GeoInformation Science and Earth Observation (ITC). Retrieved from http://www.itc.nl/library/papers_2016/msc/wrem/njuki.pdf

Odongo, V. O., Mulatu, D. W., Muthoni, F. K., van Oel, P. R., Meins, F. M., van der Tol, C., … van der Veen, A. (2014). Coupling socio-economic factors and eco-hydrological processes using a cascade-modeling approach. *Journal of Hydrology*, *518*(PA), 49–59. http://doi.org/10.1016/j.jhydrol.2014.01.012

Plummer, N., Allsopp, T., & Lopez, J. A. (2003). *Guidelines on Climate Observations Networks and Systems*. Retrieved from http://www.wmo.int/pages/prog/wcp/wcdmp/documents/WCDMP-52_000.pdf

Prakash, S., Mahesh, C., & Gairola, R. M. (2013). Comparison of TRMM Multi-satellite Precipitation Analysis (TMPA)-3B43 version 6 and 7 products with rain gauge data from ocean buoys. *Remote Sensing Letters*, *4*(7), 677–685. http://doi.org/10.1080/2150704X.2013.783248

Rientjes, T. (2015). *Hydrologic modelling for Intergated Water Resource* [Lecture Notes]. Retrieved from https://blackboard.utwente.nl/webapps/blackboard/content/listContent.jsp?course_id=_20973_1 &content_id=_920207_1&mode=reset

Shaw, E. (1994). *Hydrology in Practice* (3rd ed.). London: CRC Press.

Smith, T. M., Arkin, P. A., Bates, J. J., & Huffman, G. J. (2006). Estimating Bias of Satellite-Based Precipitation Estimates. *Journal of Hydrometeorology*, *7*(5), 841–856. http://doi.org/10.1175/JHM524.1

Sorooshian, S., Hsu, K.-L., Coppola, E., Tomassetti, B., Verdecchia, M., & Visconti, G. (2008). *Hydrological Modelling and the Water Cycle. Water Science and Technology Library* (Vol. 63). Springer.

Stockholm International Water Institute. (2005). *Making water a part of economic development: The Economic Benefits of Improved Water Management and Services. Stockholm International Water Institute (SIWI)*. Retrieved from http://hdl.handle.net/10535/5146

Tarekegn, T. H., Haile, A. T., Rientjes, T., Reggiani, P., & Alkema, D. (2010). Assessment of an ASTER-generated DEM for 2D hydrodynamic flood modeling. *International Journal of Applied Earth Observation and Geoinformation*, *12*(6), 457–465. http://doi.org/10.1016/j.jag.2010.05.007

The Ramsar Convention on Wetlands. (2014). *The annotated Ramsar List: Kenya.* Retrieved from http://ramsar.rgis.ch/cda/en/ramsar-documents-list-anno-kenya/main/ramsar/1-31-218%5E16536_4000_0__

Toté, C., Patricio, D., Boogaard, H., van der Wijngaart, R., Tarnavsky, E., & Funk, C. (2015). Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. *Remote Sensing*, *7*(2), 1758–1776. http://doi.org/10.3390/rs70201758

Treut, L., Somerville, R., Cubasch, U., Ding, Y., Mauritzen, C., Mokssit, A, … Tignor, M. (2007). Historical Overview of Climate Change Science. *Earth*, *Chapter 1*(October), 93–127.

United States Geological Society. (2016). *What is hydrology and what do hydrologists do? The USGS Water Science School*. Retrieved from http://water.usgs.gov/edu/hydrology.html

Van der Velde, R. (2010). Linking Pest to a Model. [Lecture Notes].

Van Oel, P. R., Mulatu, D. W., Odongo, V. O., Meins, F. M., Hogeboom, R. J., Becht, R., … Van der Veen, A. (2013). The Effects of Groundwater and Surface Water Use on Total Water Availability and Implications for Water Management: The Case of Lake Naivasha, Kenya. *Water Resources Management*, *27*(9), 3477–3492. http://doi.org/10.1007/s11269-013-0359-3

Vorosmarty, C. J., Leveque, C., Revenga, C., Bos, R., Caudill, C., Chilton, J., … Reidy, C. A. (2005). Ecosystems and Human Well-Being: Current State and Trends: Findings of the Condition and Trends Working Group: Chapter 7: Fresh Water. *Ecosystems and Human,* 165–207.

Wang, J., Hong, Y., Li, L., Gourley, J. J., Khan, S. I., Yilmaz, K. K., … Okello, L. (2011). The coupled routing and excess storage (CREST) distributed hydrological model. *Hydrological Sciences Journal*, *56*(1), 84–98. http://doi.org/10.1080/02626667.2010.543087

Water Resources Management Authority. (2010). *Water Allocation Plan - Naivasha Basin.* Retrieved from ftp://ftp.itc.nl/pub/naivasha/PolicyNGO/WRMA2010.pdf

Yihdego, Y., Reta, G., & Becht, R. (2016). Hydrological analysis as a technical tool to support strategic and economic development: A case study of Lake Navaisha, Kenya. *Water and Environment Journal*, 9. http://doi.org/10.1111/wej.12162

Zhang, X. C. W. C., & Luo, X. M. L. (2015). *Water Cycle Management. Water S*. Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-662-45821-1

APPENDICES

Appendix 1: Changes in 2GA06 over the years.
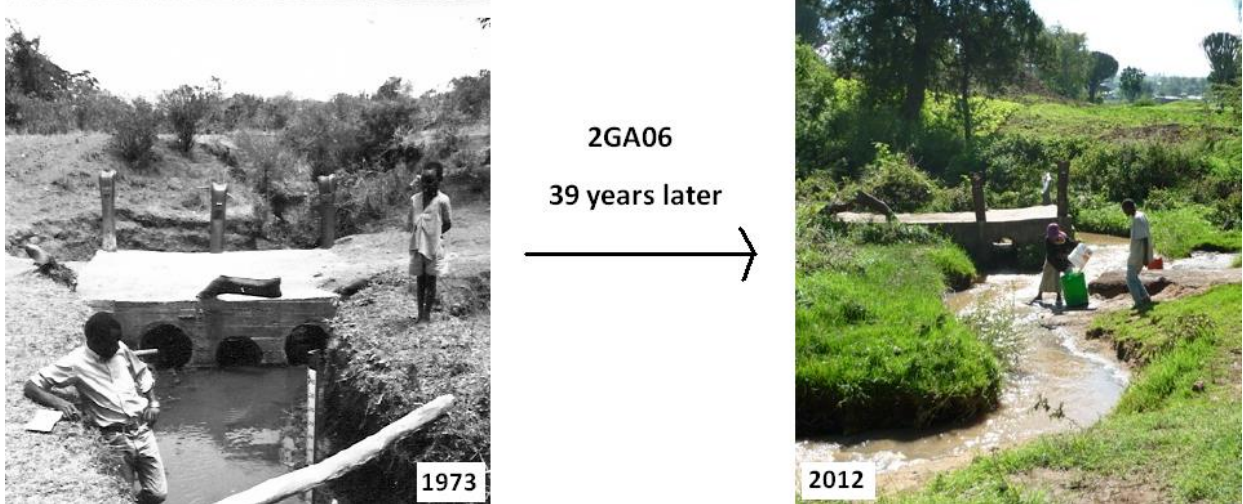
Appendix 1 shows some of the changes that have been experienced in the LNB RGS stations over the years and this has caused the changes exhibited in the rating curves of the station that have been affected such as 2GA06 above.

Appendix 2: An Example of an ArcASCII file

```
chirp_20010101.ASC
 1  ncols         74
 2  nrows         69
 3  xllcorner     33.50
 4  yllcorner     -3.25
 5  cellsize      0.05
 6  NODATA_value  -9999
 7  1.348322 1.261566 1.207108 1.158713 1.148396 1.199744 1.122372 1.265593 1.241363
 8  1.272580 1.191434 1.142193 1.052384 1.117859 1.162653 1.083099 1.206409 1.183068
 9  1.341410 1.350722 1.299771 1.163176 1.087239 1.184578 1.192863 1.204634 1.174012
10  1.451650 1.465812 1.320985 1.231061 1.336866 1.306639 1.211119 1.356400 1.280783
11  1.473657 1.399408 1.495641 1.361009 1.438356 1.454262 1.408309 1.219003 1.282190
12  1.572510 1.643202 1.605022 1.536043 1.737620 1.717943 1.596615 1.519689 1.485505
13  1.566898 1.694613 1.717205 1.611982 2.045246 1.906488 1.905185 1.794232 1.763652
14  1.532149 1.665727 1.795855 1.723608 2.065058 1.931793 1.924137 1.889625 1.815880
15  1.577345 1.714758 1.745791 1.778923 2.079023 1.995787 2.042059 1.906664 1.894965
16  1.538002 1.635510 1.669267 1.681127 2.043163 1.923081 2.009806 1.839053 1.827744
17  1.144066 1.101349 1.198123 1.182078 2.999348 2.929566 2.897478 2.595377 2.648908
18  1.144468 1.099814 1.160647 1.177258 2.911982 2.510527 2.578753 2.605654 2.499010
19  1.138950 1.026310 1.163613 1.142882 2.622906 2.496785 2.488459 2.424511 2.409177
20  1.023816 1.021554 1.047577 1.037805 2.630125 2.507372 2.415295 2.424886 2.412054
21  1.056383 1.054212 1.041909 1.033768 2.630649 2.510131 2.585283 2.603873 2.588636
22  1.116633 1.078184 1.029932 1.022472 2.511955 2.399960 2.396784 2.660086 2.650507
23  1.077417 1.114162 1.031101 1.019456 2.424819 2.404651 2.401684 2.571107 2.642765
24  1.074612 1.076407 0.959027 0.954065 2.422422 2.405895 2.400440 2.496779 2.549845
25  1.065649 1.031301 0.954724 0.951716 2.419796 2.410628 2.409236 2.509318 2.562270
26  0.954343 0.962070 0.958693 0.947357 2.417344 2.414500 2.441964 2.528572 2.507750
```

Appendix 3: CREST model parameters

| Module	Description	Symbol
CREST parameters (v2.0)		
Initial condition	Initial value of soil moisture	*W0*
	Initial value of overland reservoir	*SS0*
	Initial value of interflow reservoir	*SI0*
Physical parameters	Soil saturated hydraulic conductivity	*Ksat*
	Multiplier on the precipitation field	*RainFact*
	Mean water capacity	*WM*
	Exponential of the variable infiltration curve	*B*
	Impervious area ratio	*IM*
	Factor to convert PET to local actual	*KE*
	Overland runoff velocity coefficient	*CoeM*
Conceptual parameters	Overland flow speed exponent	*ExpM*
	Multiplier used to convert overland flow speed to channel flow speed	*CoeR*
	Multiplier used to convert overland flow speed to interflow flow speed	*CoeS*
	Overland reservoir discharge parameter	*KS*
	Interflow reservoir discharge parameter	*KI*

Appendix 4: Limits of CREST model parameters

Parameter	Lower Limit	Upper limit
RainFact	0.5	1.2
Ksat	18	313.92
WM	0.1	232.5
B	0.05	1.5
IM	0.0	0.2
KE	0.1	1.5
coeM	1.0	150
expM	0.1	2.0
coeR	1.0	3.0
coeS	0.0	1.0
KS	0.0	1.0
KI	0.0	1.0

Appendix 5: CREST model setup and outputs
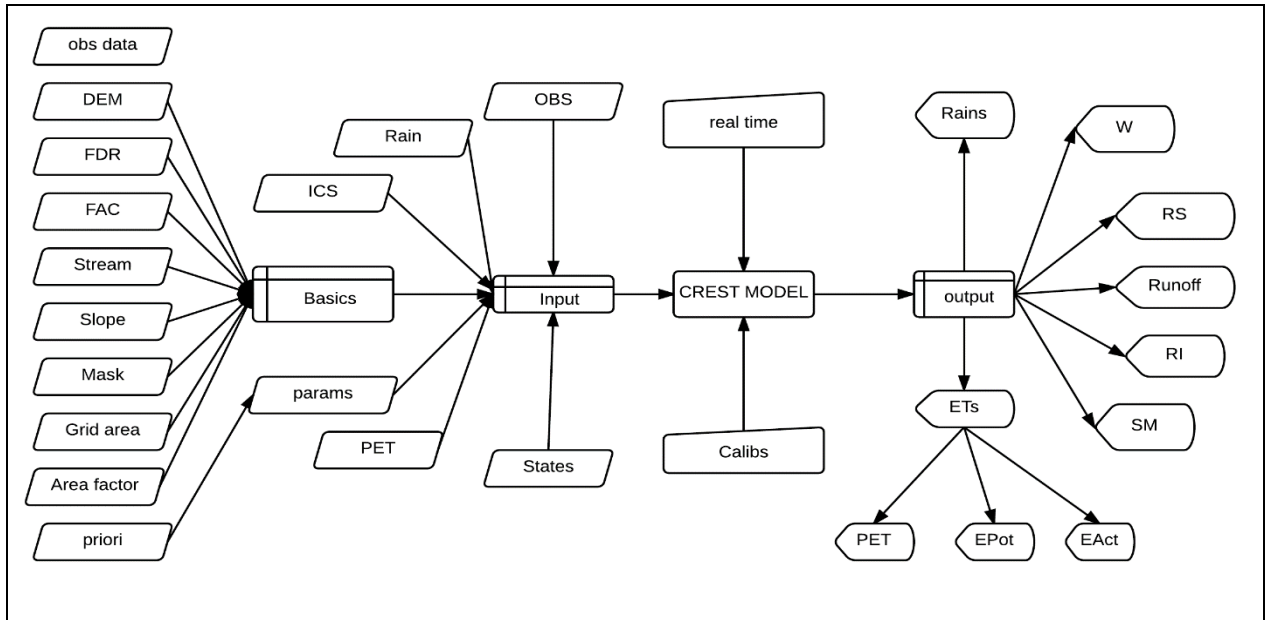
Appendix 6: Abstraction restriction indicators
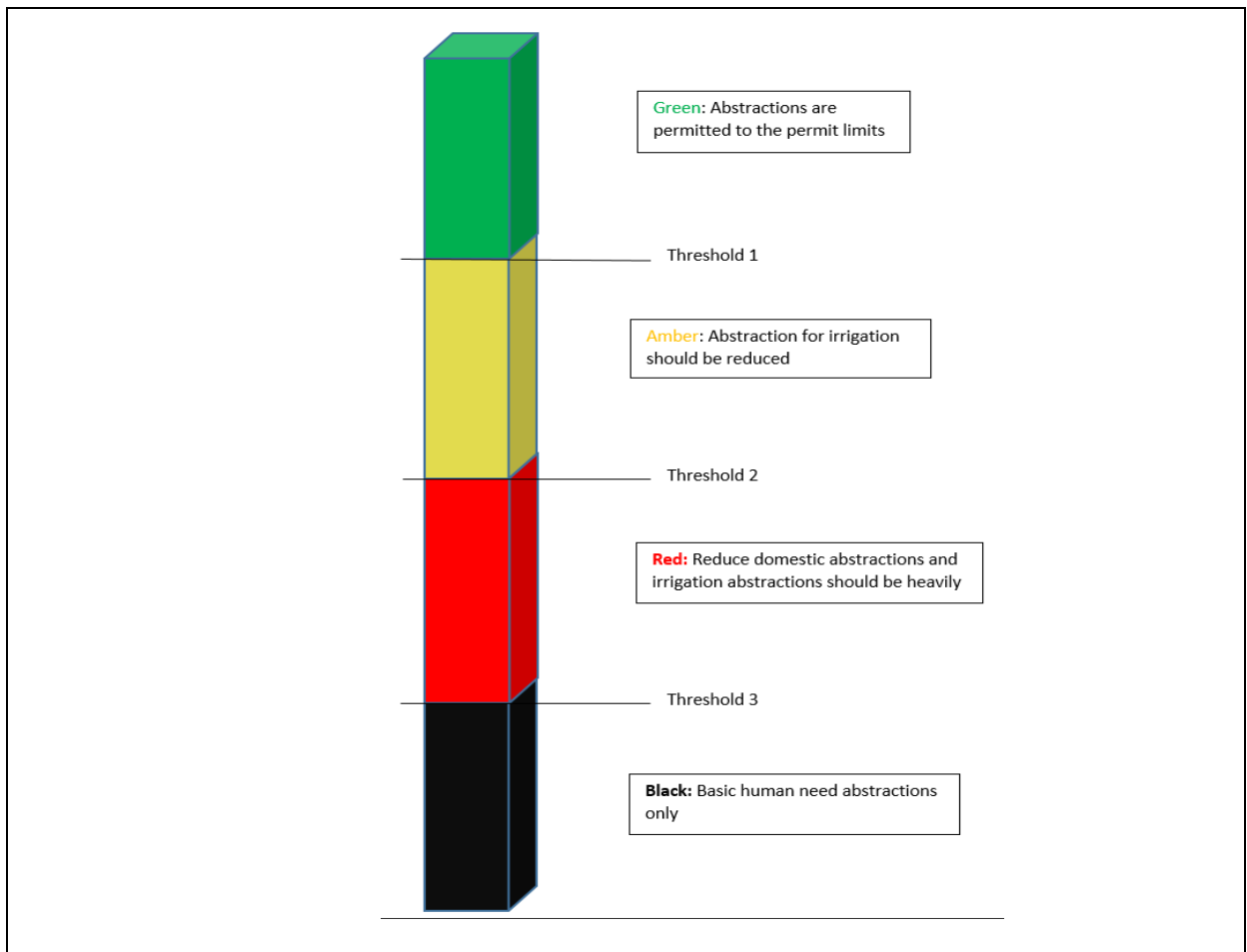
Appendix 7: Gauging at the Lake Naivasha basin

Appendix 8: Water Allocation Plan billboard for Lake Naivasha abstractions.