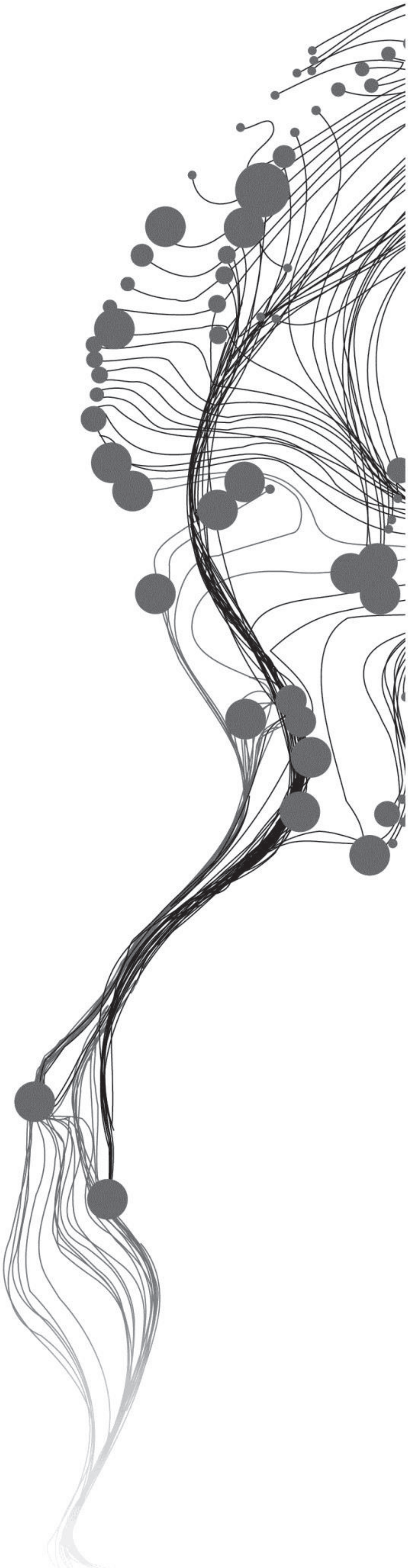
# RUNOFF MODELLING OF THE MARA RIVER USING SATELLITE OBSERVED SOIL MOISTURE AND RAINFALL

JONATHAN MUITHYA MWANIA
March, 2014

SUPERVISORS:
Dr. Ir. Rogier van der Velde
Dr. Zoltan Vekerdy



# RUNOFF MODELLING OF THE MARA RIVER USING SATELLITE OBSERVED SOIL MOISTURE AND RAINFALL

JONATHAN MUITHYA MWANIA

Enschede, The Netherlands, March, 2014

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Water Resources and Environmental Management

SUPERVISORS:
Dr. Ir. Rogier van der Velde
Dr. Zoltan Vekerdy

THESIS ASSESSMENT BOARD:
Prof. Dr. Ing. W. Verhoef (Chair)
Dr. H. van der Kwast (External Examiner, UNESCO-IHE Delft)

# ABSTRACT

Hydrological models are necessary tools in water resource management. However modeling in poorly gauged catchments is a big challenge. Recent research has shown that satellite based hydrological and meteorological data has the potential of being part of the solution towards overcoming this challenge. In this research we use the conceptual lumped rainfall-runoff model by Meier et al. (2011) to model runoff in the Mara River Basin. The model simulates runoff as a function of soil moisture with runoff as forcing data. It is built on the basis established between satellite observed soil moisture and rainfall, and the measured runoff. Reliability of the model is evaluated over three sub-catchments namely Mara mines, Nyangores and Amala in the Mara river basin using correlation coefficient (*r*) and Root Mean square Errors (RMSE) Mean Absolute Error and bias. The *r* for Mara mines Nyangores and Amala during calibration and (validation) were 0.54 (0.77), 0.67 (0.74), 0.125 (0.48) respectively. The model showed great potential in simulating dry season runoff. It needs further improvement to be able to fairly simulate wet season runoff.

Key words: Hydrological modeling, soil moisture, rainfall, runoff

**To God be the Glory**

# ACKNOWLEDGEMENTS

First I wish to express my appreciation to the Dutch government for offering me the NFP (Netherlands Fellowship Program) scholarship. If it was not for the scholarship, it would have been impossible to do the Master of Science Program in the Netherlands. On the same note I wish to express my appreciation to the ESA (European Space Agency) Tiger Initiative - Alcantara Project for partially funding the fieldwork of this research.

Second I wish to express my heartfelt gratitude to my supervisors, Dr. Ir. Rogier van der Velde and Dr. Zoltan Vekerdy for their unrelenting support and critical review of this work. Particularly I thank Dr. Ir. Rogier van der Velde for sharing with me the journal article containing the model concept used in this research and computer codes for data pre-processing. I thank Dr. Zoltan Vekerdy for facilitating my fieldwork and the extra efforts of providing leadership and capturing the fieldwork moments in pictures. I wish to thank Joseph Mtamba of the University of Dar Salaam, the Principal Investigator (PI) for being very supportive and listening during my many consultations with him. Gentlemen thank you for the excellent mentorship.

I wish to thank Drs Robert Becht for providing a link to Professor M. E. McClain of UNESCO – IHE (Institute for Water Education) from whom I obtained historical discharge data for the MRB. Kind regards to Mr Richard Kidd of TU Wien (Vienna University of Technology) for providing me with information on downloading the satellite observed soil moisture data used in this research. Kind regards also to the soil laboratory staff of the University of Dar Salaam Dar Salaam for analyzing the soil samples collected during fieldwork.

I take this opportunity to also thank the chairman of Amala river branch of the Mara River Water Users Association (MRWUA) Mr Joseph Chepusit and vice-chair Madam Jessica Tesot for showing us around the upper and middle parts of the MRB and for sharing information on the role of MRWUA as a key stakeholder in the management of the basin.

I wish to thank Mr Reuben Ngessa of Water Resources Management Authority (WRMA) for facilitating my access to recent discharge data of Amala and Nyangores rivers. On the same note I thank Kenya Meteorological Department (KMD) for providing measured rainfall data used in this research.

I wish to thank the staff at the Department of Water Resources at Faculty of Geo-Information Science and Earth Observation of the University of Twente for building me professionally. I am also very grateful to my classmates for the social and moral support during the period we have been together.

Last but not the least I wish to express my heartfelt gratitude to my family and in particular my wife Enessia for providing me with the much needed encouragement and moral support. She took excellent care of our two sons, Luckwell and Mwanga in my absence.

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLE

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AMSR-E | Advanced Microwave Scanning Radiometer for Earth Observation System |
| ASCAT | Advanced Scatterometer |
| ASCII | American Standard Code for Information Interchange |
| BWI | Basin Water Index |
| CERES | Cloud and Earth Radiant Energy Sensor |
| DEM | Digital Elevation Model |
| EGM96 | Earth Gravitational Model 1996 |
| ERS | European Remote Sensing |
| ESA | European Space Agency |
| EUMETSAT | European Organisation for the Exploitation of Meteorological Satellites |
| FDC | Flow Duration Curve |
| GeoSFM | Geospatial Stream flow Model |
| GeoTIFF | Geo-referenced TIFF |
| GLDAS | Global Land Data Assimilation System |
| GRACE | Gravity Recovery and Climate Experiment |
| GSFC | Goddard Space Flight Canter |
| GV | Ground Validation |
| HDF | Hierarchical Data Format |
| HWSD | Harmonized World Soil Database |
| IDL | Interactive Data Language |
| ISBA | Interaction-Soil-Biosphere-Atmosphere |
| KMD | Kenya Meteorological Department |
| LPRM | Land Parameter Retrieval Model |
| LVB | Lake Victoria basin |
| m.a.s.l | Metres above sea level |
| MetOp | Meteorological Operational |
| MFC | Mau Forest Complex |
| MRB | Mara River basin |
| MR | Mara River |
| MRWUA | Mara River Water Users Association |
| NASA | National Aeronautics and Space Administration |
| NetCDF | Network Common Data Format |
| NRB | Nile River Basin |
| PI | Principal Investigator |
| PRI | Polarization Ratio Index |
| PRI | Precipitation Radar |
| RfGS | rainfall gauging stations |
| RGS | River gauging stations |
| SMAP | Soil Moisture Active Passive |
| SMOS | Soil Moisture and Ocean Salinity |
| SMOSMANIA | Soil Moisture Observation System - Meteorological Automated Network Integrated Application |
| SRTM | Shuttle Radar Topography Mission |

| | |
|---|---|
| SSM | Surface soil moisture |
| SWAT | Soil Water Assessment Tool |
| SWI | Soil Water Index |
| TIFF | Tagged Image File Format |
| TMI | TRMM Microwave Imager |
| TMPA | TRMM Multi-Satellite Precipitation Analysis |
| TRMM | Tropical Rainfall Measurement Mission |
| TU Wien | Vienna University of Technology |
| USDA | United States Department of Agriculture |
| WRMA | Water Resources Management Authority |

# 1. INTRODUCTION

## 1.1. Background

Recent studies on the use of satellite observed soil moisture estimates in hydrological models have shown that these products have great potential in contributing to the quality of hydrological modelling results especially in poorly gauged catchments (Bolten et al., 2010; Brocca et al., 2010; Draper et al., 2011; Matgen et al., 2012; Pauwels et al., 2002; Scipal et al., 2008). Khan et al. (2012) in their study on microwave satellite modelling in Okavango basin (South Africa) found a Pearson correlation coefficient of 0.9 between measured runoff and Advanced Microwave Scanning Radiometer for Earth Observation System (AMSR-E) observed soil moisture. Scipal et al. (2005) in their study on soil moisture-runoff relationship at the catchment scale as observed with coarse resolution microwave remote sensing demonstrated that there is relevant hydrological information in course resolution satellite data. They used regression equation of the best fit relationship between ERS observed soil moisture and measured runoff to simulate runoff. Meier et al. (2011) in their study on hydrological real-time modelling in the Zambezi river basin further developed the concept by introducing rainfall as a forcing data. They built the model on the basis of the relationship found between satellite observed soil moisture, rainfall and in-situ measured runoff. From the soil moisture estimates, the catchments averaged profile soil moisture is calculated and expressed as Basin Water Index (BWI). Its values range from 0-1 with 0 signifying a completely dry basin with all the rainfall infiltrating, and 1 a completely saturated basin with constant infiltration. In this concept, BWI is used to partition rainfall into surface runoff and infiltration.

Some of the microwave remote sensing instruments which have provided soil moisture estimates at global scale include European Space Agency's (ESA) Soil Moisture and Ocean Salinity (SMOS) mission, the European Remote Sensing (ERS) scatterometer and the Advanced Scatterometer (ASCAT) and National Aeronautics and Space Administration's (NASA) Advanced Microwave Scanning Radiometer for Earth Observation System (AMSR-E). NASA is scheduled to launch the Soil Moisture Active Passive (SMAP) mission in 2014-2015. The instrument will have the capability of differentiating between frozen from thawed land surfaces (Entekhabi et al., 2010). Previous studies comparing retrievals from ASCAT and AMSR-E and ASCAT and SMOS indicate that ASCAT retrievals have better correlation with in-situ measurements (Brocca et al., 2011; Parrens et al., 2012). (Brocca et al., 2011) compared the soil moisture estimates generated from ASCAT and AMSR-E sensors with in-situ measurements in over 17 sites in Italy, Spain, France and Luxembourg. The authors used the Land Parameter Retrieval Model (LPRM), the Polarization Ratio Index (PRI) and the standard NASA algorithm to retrieve moisture data from the AMSR-E product. They used the Vienna University of Technology (TU Wien), change detection algorithm for retrieval from the ASCAT product. Out of the three sets of soil moisture estimates retrieved from the AMSRE-E product, the estimates by the LPRM had the highest correlation with in-situ measurements. Estimates from the ASCAT product had the best correlation results compared to estimates from the AMSRE-E products for approximately 5 cm soil layer. The average correlation coefficients were 0.71 and 0.62 for the ASCAT and the AMSR-E (retrieved using the LPRM), respectively. Parrens et al. (2012) compared ASCAT and SMOS Surface Soil Moisture (SSM) products with Interaction-Soil-Biosphere-Atmosphere (ISBA-A-gs) model simulations and in-situ measurements from the Soil Moisture Observation System - Meteorological Automated Network Integrated Application (SMOSMANIA) network. From their study, they found out that the significant anomaly correlation coefficients between in situ measurements and the SMOS (ASCAT) product was in the range of 0.23 to 0.48 (0.35 to 0.96).

## 1.2. Problem statement

MRB is a sub catchment of the Lake Victoria basin (LVB) and the larger Nile River Basin (NRB). In the upper parts of the basin is the Mau Forest Complex (MFC) where the Mara River (MR) originates from. The forest is a key water tower, a source for other rivers including Sondu, Njoro and Ewaso Ng'iro rivers. In the middle part of the catchment is the tropical savannah vegetation supporting the unique Mara-Serengeti ecosystem, famous with the scenic large scale seasonal migration of the wilder beast migration. In the south western parts is the Mara Wetlands ecosystem. MR and its two main tributaries, Amala and Nyangores, are the only perennial rivers in the basin. The ecosystems, thriving tourism industry, agriculture and pastoral farming depend on these rivers especially during the dry seasons (Dessu et al., 2014; Gereta et al., 2009). According to Dessu and Mellesse (2012), a third of available arable land in MRB is under small scale farming.

Previous studies show that there has been change to the MR flow regime. A study on the impacts of land use/cover on the hydrology of MR by Mati et al. (2008) using Geospatial Stream Flow Model (GeoSFM) found out that the peak flows have increased by 7%, occurring 4 days early for the period between 1973 and 2000. Using Landsat images, they also found out change in land cover/use over the same period. Notably, agricultural and wetland areas had increased by 203% and 387% while the savannah vegetation and forest areas were found to have reduced by 79% and 32% respectively. Mango et al. (2011) used the Soil Water Assessment Tool (SWAT) to investigate the impact of land use and climate on the hydrology of the upper MRB. Their results showed that conversion of forest areas to agriculture and grasslands areas was most likely reducing dry season flows while increasing quick peak flows. Human activity in MRB is affecting both the flow regime and the water quality of MR, (Gereta et al., 2009), (see also figure 1). Juston et al. (2013) used 44 year historical data to study the rating curve uncertainty and change in discharge time series of the Nyangores River. From 4 Flow Duration Curves (FDC) of 8 year data intervals, they detected a reduction in the lowest base flow.

Figure 1: The picture on the left shows a man cutting trees in the upper Mara River Basin. The picture on the right shows men washing motorbikes and few steps down stream, cattle drinking the turbid water. Human activities are contributing to degradation of the basin. Pictures by Vekerdy Z. (2013)

There is need for integrated management of the basin's water resources for it to meet the demand of the competing users especially during drought. To achieve this, hydrological models can be very useful to the managers. However as noted by Dessu and Mellesse (2012) in their study on modelling rainfall-run off processes in the MRB using SWAT, performance of the model in the basin depends on the quality and quantity of discharge data. They noted uncertainties in the discharge data. Previous studies indicate that lack of sufficient in-situ data, especially in developing countries, is a challenge to researchers, (Khan et al., 2012; Sivapalan, 2003). Globally, available in-situ data lacks homogeneity in the quality of individual measurements as seen with for example, global soil moisture measurements, (Dorigo et al., 2011; Dorigo et al., 2013). In-situ measurements are in particular difficult and time consuming (Brocca et al., 2007; Engman & Chauhan, 1995; Robock et al., 2000). This has motivated more research on derivation of hydrological information from satellite derived products.

This research builds from success of previous studies in runoff simulation models based on satellite observed soil moisture and rainfall products. The approach has shown to be very promising in addressing the problem of modelling data scarcity in poorly gauged basins. The satellite observed soil moisture and rainfall products used in this research are from ASCAT and Tropical Rainfall Measurement Mission (TRMM) respectively. The products are readily available in open source internet databases. The research is linked to the ESA Tiger Initiative - Alcantara Project No. 12-A15. It is expected to contribute towards flow estimation in the Lower Mara basin for wetlands hydrodynamic modelling by Joseph Mtamba – the PI.

## 1.3. Objectives

The main objective of this thesis is to use satellite based soil moisture and rainfall products for quantifying the runoff of the Mara River basin.

The specific objectives of the thesis are as follows:

- To develop an empirical model simulating Mara river runoff as function of the soil moisture using satellite observed rainfall as forcing data;
- To calibrate and validate the model for three gauging stations along the Mara river using satellite observed rainfall and soil moisture;
- To investigate the performance of the calibrated model for the different sub-catchments.

## 1.4. Research questions

The objectives defined above led to the following questions which this research sought to answer:

- Is there a relationship between satellite observed soil moisture and rainfall, and measured runoff of the Mara River?
- How do the model parameters behave for the different sub-catchments?
- How does the model performance compare for the different sub-catchments?

# 2. STUDY AREA AND DATA COLLECTION

## 2.1. Study area

MRB covers an area of 13750 $km^2$ in south western Kenya and north western Tanzania. The MR originates from the MFC at an attitude of about 3000 metres above sea level (m.a.s.l). The river flows south westwards over a stretch of 395 km before draining into Lake Victoria at Musoma in Tanzania. It has two main perennial tributaries in the upstream part, namely the Nyangores and Amala rivers. Analysis of historical (1970 to 1996) discharge data from for Mara river at Mara mines, Nyangores at Bomet and Amala at Mulot shows a mean of 33.9 $m^3s^{-1}$, 8.4 $m^3s^{-1}$and 9.9 $m^3s^{-1}$with standard deviation of 60 $m^3s^{-1}$ , 7.1 $m^3s^{-1}$ and 19.9 $m^3s^{-1}$ respectively (Dessu & Mellesse, 2012). The MRB has two rainy seasons. The long rainy season is between March and June and the short season is between November and December. The mean annual rainfall varies from 1000 mm to 1750 mm, 900 mm to 1000 mm and 300 mm to 800 mm in the upper, middle and lower parts of the basin respectively (Dessu et al., 2014; Dessu & Mellesse, 2012; Krhoda, 2005). Figure 2 is a map showing the major land covers/uses in the basin. The map also shows the discharge gauging stations along Mara, Nyangores and Amala rivers. The middle areas of the catchment are dominantly savannah vegetation supporting the Mara-Serengeti ecosystem. On the lower parts of the basin are the Mara wetlands.

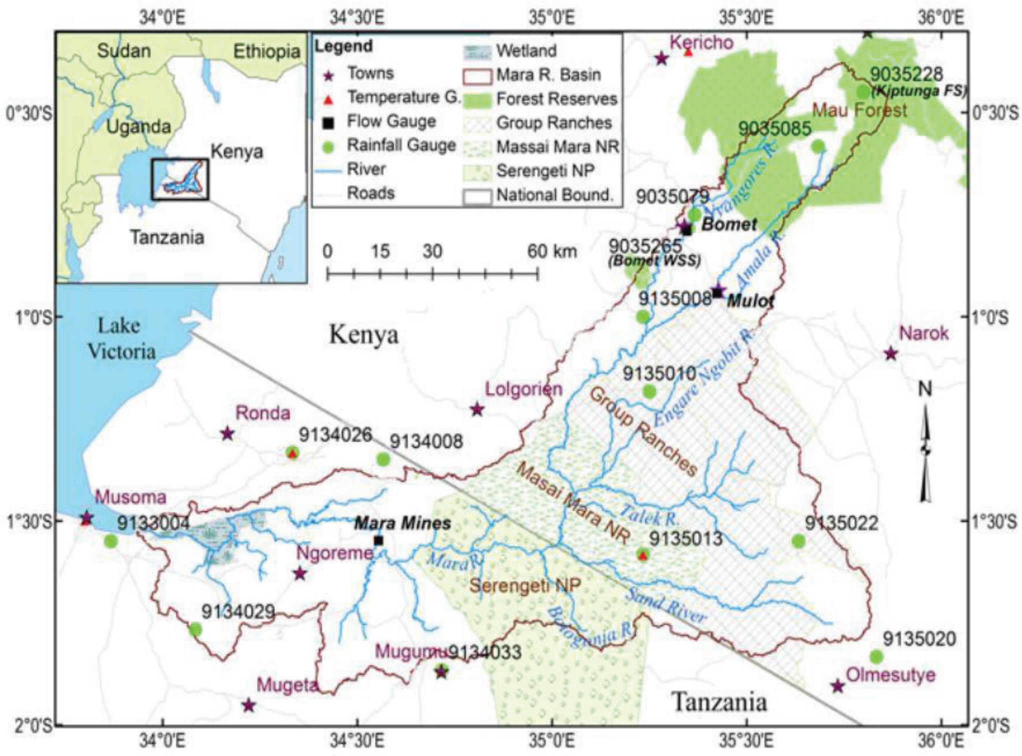



Figure 2: Map of the MRB. The major land covers/uses in the basin include forest reserves on the upper northern parts, in the middle parts are the savannah vegetation supporting the Mara-Serengeti ecosystem. In the lower southern parts is Mara wetlands. Source: Dessu & Mellesse (2012)

## 2.2. Ground measurements and field observation

Fieldwork was undertaken between September 15 and October 8, 2013. The activity was partly funded by ESA Tiger Initiative. The objectives of the field work were to:

- Contact a reconnaissance of the study area
- Obtain runoff data

- Collect rainfall data
- Collect soil samples for soil characterisation
- Vegetation mapping in the Mara wetland for the ESA Tiger Initiative - Alcantara Project No. 12-A15.

The fieldwork team was composed of:

1. Jonathan Mwania - Master of Science student
2. Joseph Mtamba - PI
3. Dr Zoltan Vekerdy - Supervisor

The team traversed the basin collecting soil samples and making observations of the basins characteristics including land cover/use, topography and the RGSs beginning from the source of Nyangores and Amala and then downstream all the way to the Mara wetlands and to the outlet to Lake Victoria.

### 2.2.1. Runoff data

The runoff data is needed for validation and calibration of the soil moisture-runoff model. There are six river gauging stations (RGS) along the MR and its two main tributaries, Nyangores and Amala (see tables 1). Out of the six RGSs, only three have got long time series of data with minimal gaps. These are Mara mines, Nyangores and Amala. Data for these three RGS was collected from WRMA regional office in Kisumu, Kenya. The data are daily averages expressed in m3s-1. Historical data for Mara mines, Nyangores and Amala Mara-Lalgorian Bridge and Kirumi Ferry RGSs was obtained from the UNESCO – IHE courtesy of Professor M. E. McClain and Joseph Mtamba of the University of Dar Salaam. Vandalism, negligence (see figure 3) and destruction of the RGS equipment by floods, were noted as the causes of gaps in the runoff data of the Mara River. The Nyangores and Amala RGSs were rehabilitated and installed with automatic gauges in 2012. However the data used in this research is for up to June 2011.

Figure 3: The picture on the left is Amala RGS at Kapkimolwa, Mulot. The station was rehabilitated and equipped with an automatic gauge with support from the World Bank Nile Basin Initiative (WB-NBTF), Swedish International Development Cooperation Agency (SIDA) and the Germany Society for International cooperation (GIZ). The picture on the right is neglected RGS at Emarti bridge along the Amala river. Picture by: Vekerdy Z.

At the Mara Mines, Nyangores and Amala RGS, two readings of the water level are taken daily, one in the morning and the other in the evening. Rating curves are then used to estimate daily average discharges. In Mara mines a new rating curve was developed in 2012. However all the data for this station used in this research is for the period before the new curve was developed. Inconsistency on the time for recording the morning and evening water levels by the gauge readers was noted as a possible source of uncertainty in the data. During fieldwork, it was also observed that the Mara mines RGS was incapable of capturing extreme flood events. This is because the water levels surpass the gauge staff height. The river cross sectional dimension at the location may change with time as they are not embanked.

Table 1: RGS along the Mara River. Only the Mara mines, Nyangores and Amala RGSs have relatively long historical data.

| Station Name | Station code | Longitude ($^0$E) | Latitude ($^0$S) | Altitude (masl) | Start Year |
|---|---|---|---|---|---|
| Nyangores | 1LA03 | 35.35 | -0.79 | 1899 | 1963 |
| Amala | 1LB02 | 35.43 | -0.89 | 1860 | 1955 |
| Lalgorian bridge | ILA04 | 35.04 | -1.23 | 1594 | 1970 |
| Mara Mine | 5H2 | 34.55 | -1.55 | 1181 | 1969 |
| Kirumi ferry | 5H3 | 33.86 | -1.51 | 1132 | 1969 |

#### 2.2.2. Rainfall data

The in situ rainfall data was used to investigate the reliability of the satellite rainfall product. There are forty four rainfall gauging stations (RfGS) within and around the basin (see also figure 2). Data for the stations on the Kenyan side of the basin was obtained from KMD, Nairobi. For the stations on the Tanzanian side, the data was provided by Joseph Mtamba of University of Dar Salaam. Out of the forty four stations, only six (shown in figure 6) had sufficient data falling within the span of satellite rainfall data used in this research. Sample RfGS visited during the field work were observed to have the tipping bucket type of gauges (see also figure 4 below).

Figure 4: Weather station at the Mara River Water Users Association (MRWUA) office in Mulot. The station has a tipping bucket and a manual rain gauge. The weather station automatic measuring instruments are not yet connected to the data logger. Picture by Vekerdy Z. (2013)

#### 2.2.3. Soil characterisation

Fifteen soil samples were collected from a depth of 0-20cm during the fieldwork. These samples, in addition to nine others collected in a previous fieldwork by Joseph Mtamba were used for particle distribution analysis. The analysis was conducted at the University of Dar Salaam in Tanzania. The results of the analysis were used to classify the soils as suggested by United States Department of Agriculture, USDA (1999). Additional soil data was extracted from the Harmonized World Soil Database (HWSD) (FAO et al., 2012) and used to identify the soil types with respect to the classifications (table 2). Figure 5 shows soil map of the MRB extracted from the 1:5 million HWSD raster map including the soil sampling points for this research. The results of the analysis helped in understanding the modelling results.

Table 2: Soil particle distribution analysis results, *S Pt.* is sample point, C is clay, Slt silt, Snd sand and G gravel. The samples were classified with the USDA Texture Class.

| *S Pt* | Long. $^0$E | Lat. $^0$N | *C* % | *Slt* % | *Snd* % | *G* % | USDA Texture Class | HWSD Soil type |
|---|---|---|---|---|---|---|---|---|
| 1 | 35.51 | -0.87 | 17 | 55 | 26 | 27 | loam | Humic Cambisols (CMu) |
| 2 | 35.50 | -0.87 | 17 | 49 | 34 | 0 | clay loam | Haplic Phaeozems (PHh) |
| 3 | 35.54 | -0.82 | 14 | 39 | 35 | 12 | clay (light) | Mollic Andosols (ANm) |
| 4 | 35.57 | -0.77 | 10 | 48 | 18 | 24 | clay (light) | Mollic Andosols (ANm) |
| 5 | 35.52 | -0.77 | 13 | 46 | 39 | 2 | clay (light) | Mollic Andosols (ANm) |
| 6 | 35.45 | -0.81 | 18 | 52 | 27 | 3 | clay (light) | Mollic Andosols (ANm) |
| 7 | 35.41 | -0.96 | 13 | 67 | 18 | 2 | silty clay | Vertic Luvisols (LVv) |
| 8 | 35.23 | -1.06 | 14 | 32 | 49 | 5 | silt | Eutric Vertisols (VRe) |
| 9 | 35.42 | -0.71 | 10 | 42 | 46 | 2 | clay (light) | Mollic Andosols (ANm) |
| 10 | 35.33 | -0.83 | 13 | 53 | 34 | 0 | sandy clay | Luvic Phaeozems (PHl) |
| 11 | 35.25 | -0.93 | 23 | 51 | 26 | 0 | silt | Eutric Vertisols (VRe) |
| 12 | 35.25 | -1.08 | 4 | 33 | 57 | 6 | sandy clay | Luvic Phaeozems (PHl) |
| 13 | 35.24 | -1.17 | 10 | 47 | 41 | 2 | silt | Eutric Vertisols (VRe) |
| 14 | 35.20 | -1.18 | 10 | 44 | 41 | 5 | silt | Eutric Vertisols (VRe) |
| 15 | 35.12 | -1.20 | 10 | 53 | 35 | 2 | silt | Eutric Vertisols (VRe) |
| 16 | 34.28 | -1.47 | 19 | 58 | 23 | 0 | loamy sand | Eutric Fluvisols (FLe) |
| 17 | 34.12 | -1.65 | 12 | 35 | 53 | 0 | silty clay loam | Eutric Planosols (PLe) |
| 18 | 34.26 | -1.58 | 15 | 63 | 22 | 0 | silty clay loam | Eutric Planosols (PLe) |
| 19 | 34.57 | -1.52 | 5 | 37 | 56 | 3 | silty clay loam | Eutric Planosols (PLe) |
| 20 | 34.51 | -1.50 | 13 | 59 | 28 | 0 | silty clay loam | Eutric Planosols (PLe) |
| 21 | 34.63 | -1.55 | 6 | 38 | 54 | 2 | silty clay loam | Eutric Planosols (PLe) |
| 22 | 34.87 | -1.57 | 6 | 46 | 45 | 4 | silty clay loam | Eutric Planosols (PLe) |
| 23 | 34.68 | -1.66 | 5 | 37 | 56 | 3 | sandy clay | Luvic Phaeozems (PHl) |
| 24 | 34.71 | -1.74 | 5 | 36 | 59 | 0 | sandy clay | Luvic Phaeozems (PHl) |

## 2.3. Satellite data sets

The satellite data sets used in this research were:

1. TRMM rainfall
2. ASCAT Soil Water Index (SWI).
3. Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM)

All these data sets were downloaded from open source internet data bases (see also table 3)

Table 3: Satellite data sources - the sources are open to the public.

| Data | Web Source |
|---|---|
| TRMM rainfall | http://mirador.gsfc.nasa.gov/cgi-bin/mirador/homepageAlt.pl?keyword=TRMM |
| ASCAT SWI | http://rs.geo.tuwien.ac.at/products/ |
| SRTM DEM | https://lta.cr.usgs.gov/SRTM2 |

### 2.3.1. Soil Water Index (SWI)

The satellite soil moisture product to be used in this model is the Soil Water Index (SWI). The product is derived from scatterometer generated SSM following the concept of a two-layer force-restore model as suggested by Wagner et al. (1999b). In this model, the profile soil moisture is calculated from previous SSM measurements as a function of time and expressed as the SWI as shown in equation 1 below. This model was developed at the Vienna University of Technology (TU Wien).

$$SWI = \frac{\sum_i^n SSM(ti)e^{\frac{-(t-ti)}{T}}}{\sum_i^n e^{\frac{-(t-ti)}{T}}} \quad for\ t = ti < t \tag{1}$$

Where SSM is the surface soil moisture from scatterometer at time $t_i$ and $T$ is the characteristic time length. The SSM measurements are generated using a change detection algorithm described in (Wagner et al., 1999a)

The SWI product used in this research is calculated from ASCAT sensor generated SSM data distributed by European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). ASCAT is a remote sensing instrument on board Meteorological Operational (MetOp) platform. The instrument is an active microwave with vertical polarization and C-band at 5.255 GHz (Brocca et al., 2010; Wagner et al., 2013). It is a follow up to the ERS scatterometer. (Brocca et al., 2011; Brocca et al., 2010; Wagner et al., 2013)

The product was developed under the framework of the Geoland2 project. It has a daily temporal resolution and 12.5 km spatial resolution. It is available for the period from 1st January 2007 up to date. It is in Hierarchical Data Format (HDF5) and compressed in .bz2. After downloading the product it was pre-processed using Interactive Data Language (IDL) codes (see appendix A and B). The data was extracted for further processing using BEAM VISAT software.

### 2.3.2. Tropical Rainfall Measurement Mission (TRMM) rainfall

TRMM rainfall data was used as a forcing data to the rainfall-runoff model in this research. TRMM is a joint mission of NASA and Japan Aerospace Exploration Agency for measuring tropical and subtropical rainfall. The mission employs various instruments including TRMM Microwave Imager (TMI), Cloud and Earth Radiant Energy Sensor (CERES), Precipitation Radar (PR) and Lightning Imaging Sensor (Liu et al., 2012; NASA, 2011). The TRMM satellite observations validation with ground observations is supported by the TRMM Ground Validation (GV) program at the NASA/ Goddard Space Flight Canter (GSFC) (NASA, 2011; Wolff et al., 2005). TRMM provides several rainfall products (see table 4). The data range of the TRMM product is from 1st January 1997 up to date.

Table 4: Gridded TRMM rainfall products have been spatial averaged with a resolution of 0.25° x 0.25° and 1° x 1°. 3B42 is also available on a daily temporal resolution. Source: NASA (2011)

| **Gridded TRMM Products** | |
|---|---|
| Product ID | Product Name |
| 3A11 | Monthly 5° x 5° Oceanic Rainfall |
| 3A12 | Monthly 0.5° x 0.5° mean 2A12, profile, and surface rainfall |
| 3A25 | Monthly 5° x 5° and .5° x .5° Space-borne Radar Rainfall |
| 3A26 | Monthly 5° x 5° Surface Rain Total |
| 3B31 | Monthly 5° x 5° Combined Rainfall |
| 3A46 | Monthly 1° x 1° SSM/I Rain |
| 3B42 | 3-hour 0.25° x 0.25° TRMM and Other-GPI Calibration Rainfall |
| 3B43 | Monthly TRMM and Other Sources Rainfall |
| CSH | Monthly 0.5° x 0.5° Convective & Stratiform Heating |

This research used the 3B42 version 7 product resampled to daily temporal resolution. The product is a result of the TRMM Multi-Satellite Precipitation Analysis (TMPA) (NASA, 2011). The data is available on a 0.25° x 0.25° resolution and at latitudes 50° N and 50° S.

The data was downloaded in Network Common Data Format (NetCDF). IDL codes (see appendix C and D) were used to convert the data into Tagged Image File Format (TIFF) and then into American Standard Code for Information Interchange (ASCII) file which was then opened in excel spread sheet for further processing.

#### 2.3.3. Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM)

This research used the SRTM DEM with a 90m spatial resolution (3arc seconds). The DEM has been resampled using cubic convolution and voids filled using interpolation algorithms and other sources of elevation data (USGS, 2012). DEM product specifications are elaborated in table 5. The DEM was downloaded via Earth explorer as Geo-referenced TIFF (GeoTIFF). It was processed using arc hydro toolbox in arc map. The delineation was done for Mara mines, Nyangores and Amala sub-catchments with respect to their corresponding RGSs. The area of the sub catchments were established as 11,280, 693 and 697km$^2$ for Mara mines, Nyangores and Amala sub-catchments, respectively. Figure 6 shows the delineated catchments and the drainage network. The sub catchment shape files were used in masking the satellite observed soil moisture and rainfall products for data extraction. The DEM indicates the highest and the lowest point in the basin to be 3063 m.a.s.l. and 1134 m.a.s.l. respectively, (see also figure 6).

Table 5: SRTM DEM specifications, source: USGS (2012)

| **Product Specifications** | |
|---|---|
| Projection | Geographic |
| Horizontal Datum | WGS84 |
| Vertical Datum | EGM96 (Earth Gravitational Model 1996) |
| Vertical Units | Meters |
| Spatial Resolution | 1 arc-second for the United States (~30 meters)<br>3 arc-seconds for global coverage (~90 meters) |
| Raster Size | 1 degree tiles |
| C-band Wavelength | 5.6 cm |

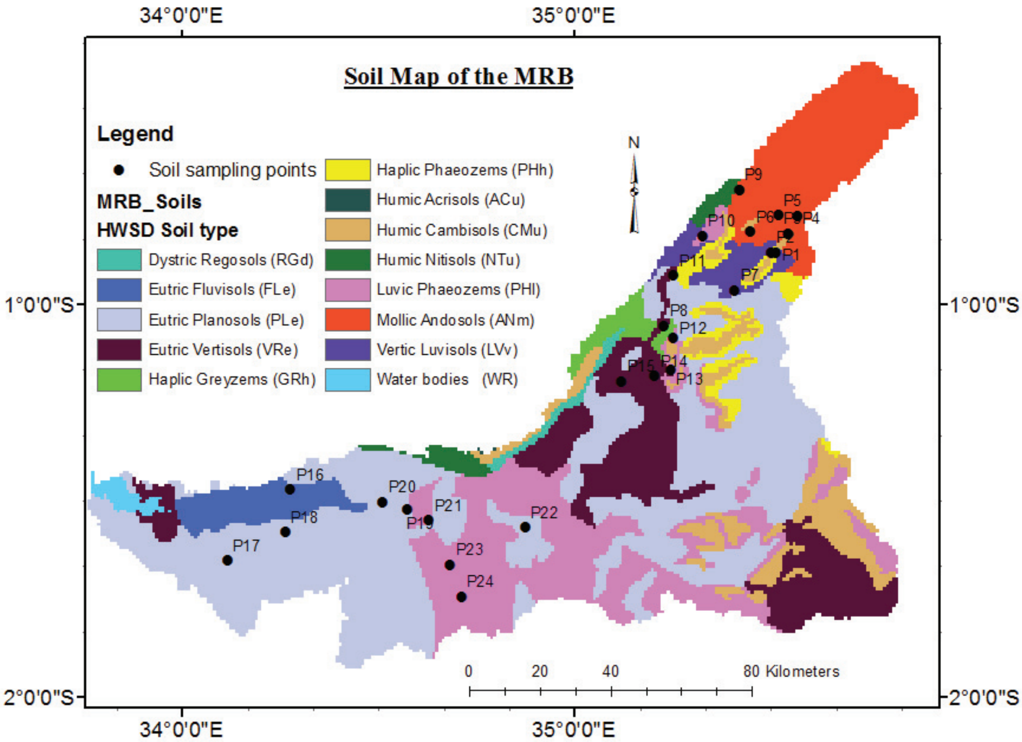



Figure 5: Soil Map of the MRB extracted from the HWSD raster map. The sampling points for this research are also shown in the map.

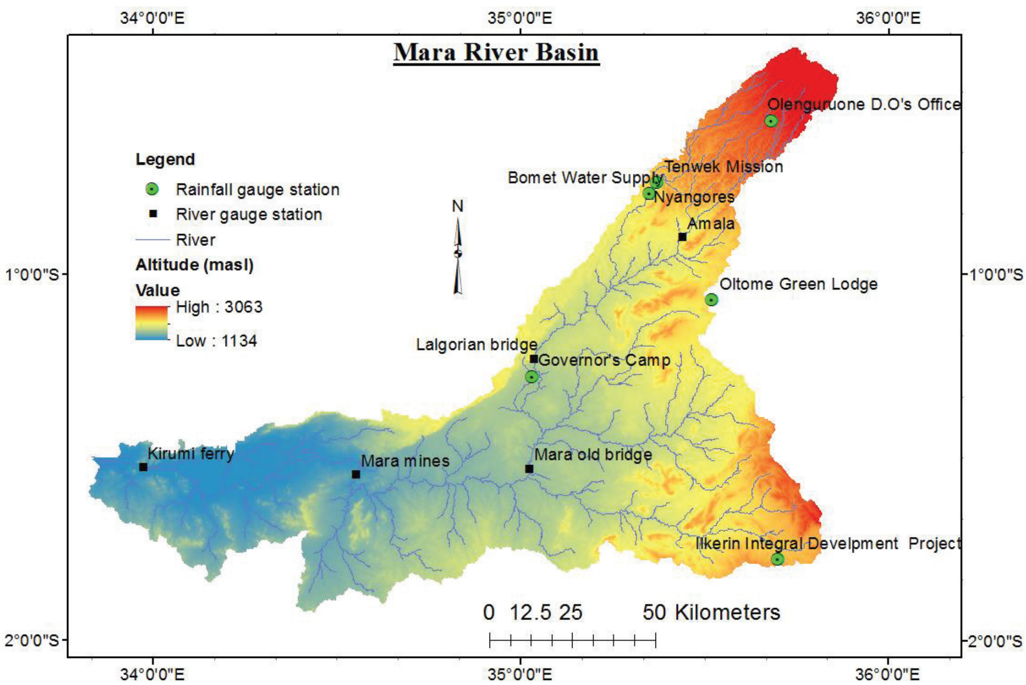



Figure 6: Processed SRTM DEM illustrating the elevation of MRB. The figure also shows the RfGS used in this research and RGS along the MR

# 3. SWI, RUNOFF, RAINFALL TIME SERIES ANALYSIS

The SWI used in this research has been derived from ASCAT generated SSM. From this SWI, Basin Water Index (BWI) index is calculated as suggested by Scipal et al. (2005). They defined BWI as SWI averaged over a given catchment, see equation 2.

$$BWI = \frac{\sum_{i=1}^{n} SWI_i}{n} \tag{2}$$

Where $n$ is the number of pixels in the catchment and $i$ refers to a pixel (-). BWI is dimensionless and varies from 0 to 1. A BWI value of 0 indicates completely dry catchment condition, while a value of 1 indicates a completely saturated condition. BWI daily time series data sets for the three sub-catchments namely Mara mines, Nyangores and Amala were calculated. Ground soil moisture measurements were not available for validation of the BWI. TRMM rainfall daily time series for each of the three sub-catchments were calculated by averaging the extracted TRMM data over the whole catchment. Figure 7 shows BWI, TRMM and discharge 30 day daily mean time series plots for the three sub catchments. The BWI and TRMM data sets are from January 2007 up to July 2013. There are a lot of gaps on the discharge data sets.

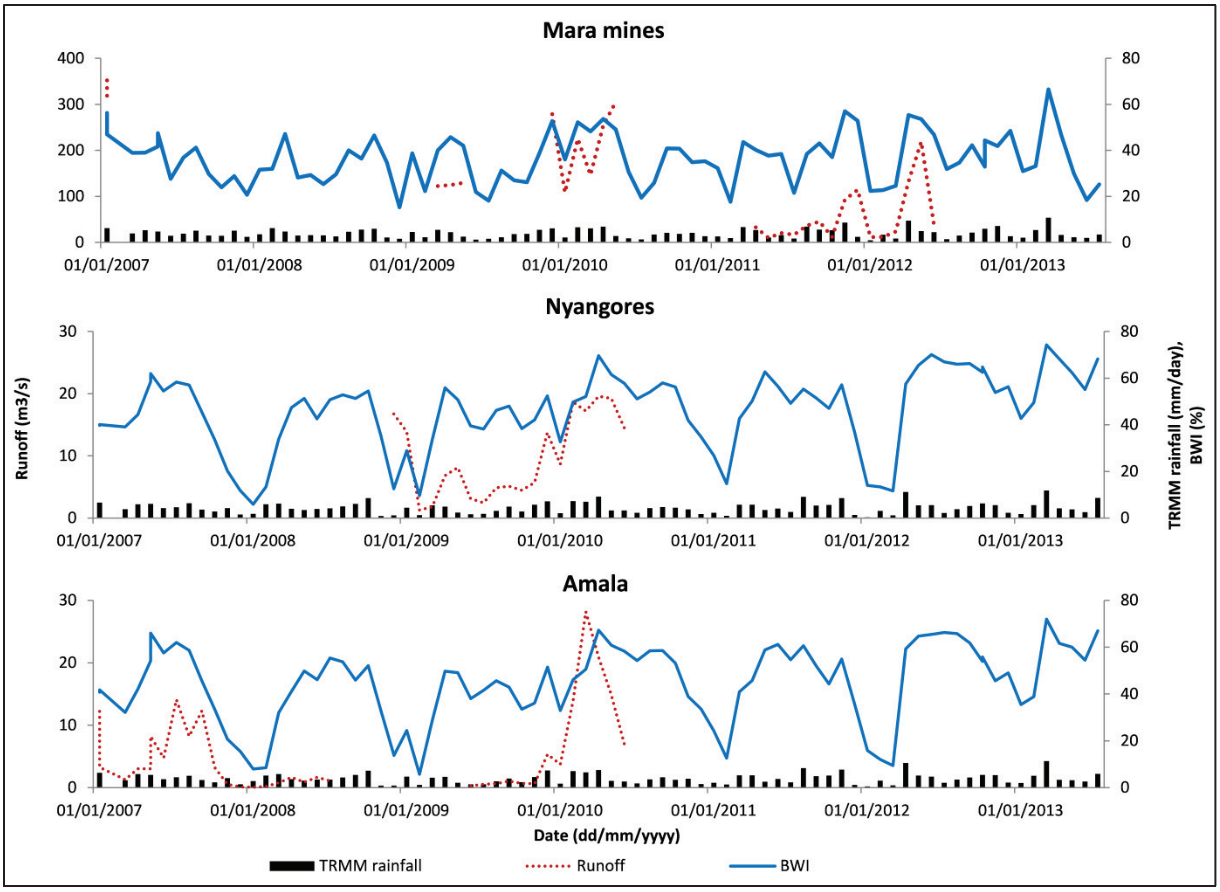



Figure 7: BWI, TRMM and discharge 30 day daily mean time series plots for Mara mines, Nyangores and Amala sub catchments. BWI is expressed as a percentage on the right Y axis together with rainfall in mm while runoff is on the left scale. Nyangores and Amala have relatively higher BWI compared to Mara mines. The peaks tend to follow a seasonal pattern. The peak rainfall seasons in MRB are March to May and November to December.

Also in this figure, the peaks of the BWI, TRMM and runoff are shown to coincide. However there is a notable exceptional case of the peak events in 2010 for Amala sub-catchment. In this case, the runoff is shown to peak before BWI. Nyangores and Amala have relatively high and low BWI peaks compared to Mara mines. For the period of under consideration, Nyangores had the highest mean BWI of 0.46 and a standard deviation of 0.2 followed by Amala with mean of 0.45 and a standard deviation of 0.2. Mara mines had the lowest mean of 37 with a standard deviation of 16. From figure 9, a seasonal trend is seen with the peaks coinciding with the rainfall seasons in the MRB. The basin has two rain seasons between March and June and between November and December.

### 3.1. BWI and TRMM rainfall relationship

The relationship between BWI and TRMM was further investigated quantitatively using the coefficient of determination ($R^2$). 30 day daily mean time series for the three sub-catchments were used in this investigation. Temporal averaging of the data sets was done to minimise noise. The period for the investigation was between January 2007 and July 2013. The best fitting trend line for BWI plotted against TRMM was found to be logarithmic (figure 8). The $R^2$ values for Mara mines, Nyangores and Amala sub-catchment were 0.54, 0.5 and 0.51 respectively. From figure 8, it can be seen that as BWI and rainfall increases, the scatter of the data points increases also. This is because as the rainfall increases, soil moisture continues increasing until the soil is completely saturated (BWI = 1). At this point, infiltration is at maximum capacity and any further increase in rainfall intensity leads to increase in contribution to surface runoff. This supports the findings of Meier et al. (2011) who however suggest that soil moisture is more correlated to occurrence of rainfall events than to the magnitude of the rainfall event. This latter was not investigated in this research.

Also in figure 8, a distinct difference in the slope of the trend line for Mara mines compared to Nyangores and Amala is observed. The Mara mines slope is gentle compared to the rest. .

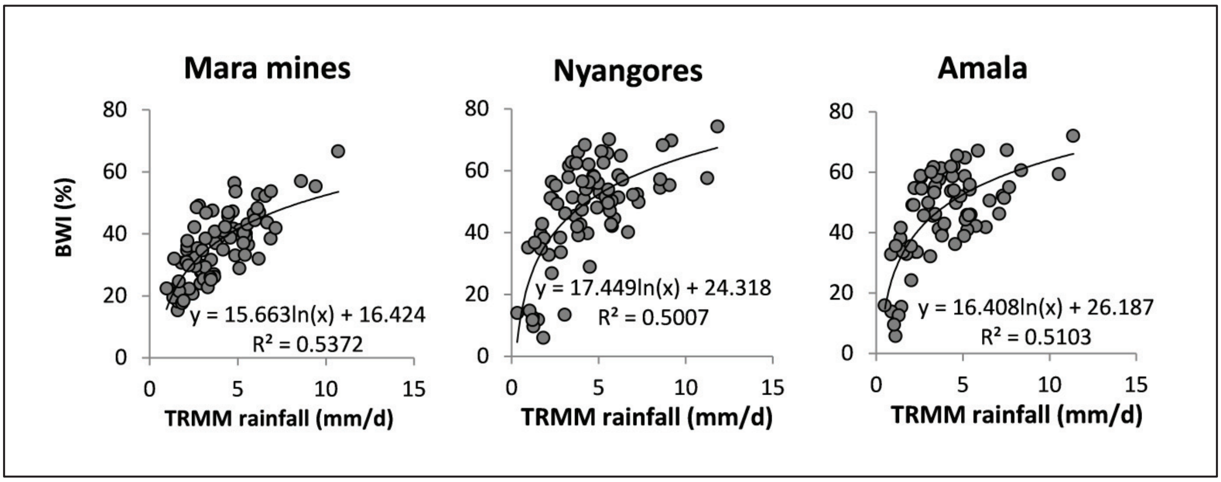



Figure 8: BWI versus TRMM scatter plots for Mara mines, Nyangores and Amala sub catchments. The plots are for 30 day daily means. The slope is taken as an indicator of the infiltration rate of the given catchment. A steep slope indicates higher infiltration rate and vice versa. Mara mines is shown to have the lowest infiltration rate compared to Nyangores and Amala. The study period was between January 2007 and July 2013.

The shape of the trend line can be assumed to be related to catchment characteristics influencing infiltration. From these curves, it can thus be deduced that there is higher infiltration in Nyangores than in Amala with Mara mines having the lowest. These results support the arguments by previous studies which attribute the high infiltration in Nyangores and Amala to their relatively higher forest cover compared with

Mara mines (Dessu et al., 2014; Dessu & Mellesse, 2012; Gereta et al., 2009; Mango et al., 2011; Mati et al., 2008).

### 3.2. BWI and runoff relationship

The relationship between BWI and runoff was investigated quantitatively using $R^2$. Runoff 30 day daily mean time series for the three sub-catchments were plotted against BWI. The best fitting trend line for this relationship was found to be exponential with $R^2$ values of 0.6, 0.68 and 0.67 for Mara mines, Nyangores and Amala respectively. The period for the investigation was between January 2007 and July 2013. From figure 9 it can be seen that the scatter of the data points increases with increase of BWI and runoff. The explanation to this trend is that, as the soil moisture increases surface runoff also increases and as the moisture content approaches saturation point, infiltration rate approaches optimal level with more rainfall being routed to surface runoff. Meier et al. (2011) and Scipal et al. (2005) in a similar analysis found similar behaviour in the relationship of BWI and runoff. Meier et al. (2011) attributed this decoupling of soil moisture from runoff as the moisture content approaches saturation point to rainfall.

From figure 9 it can be seen that the slope of the trend lines for Mara mines and Amala are steeper compared to that of Nyangores. The steepness can be taken to indicate the area related storage capacity with very steep slope indicating low storage capacity. With this assumption, Mara mine is shown to have lower storage capacity compared to the other sub catchments. Comparing Nyangores and Amala, Amala is shown to have a steeper slope as BWI increases. Amala is also shown to have a relatively gentle slope for lower BWI compared to Nyangores. This may be an indication of lower flow rates during dry seasons in comparison to Nyangores. These results support the findings by previous studies that Nyangores and Amala have higher storage capacities which serves to sustain Mara river during dry seasons, (Dessu et al., 2014; Dessu & Mellesse, 2012; Mango et al., 2011; Mati et al., 2008).

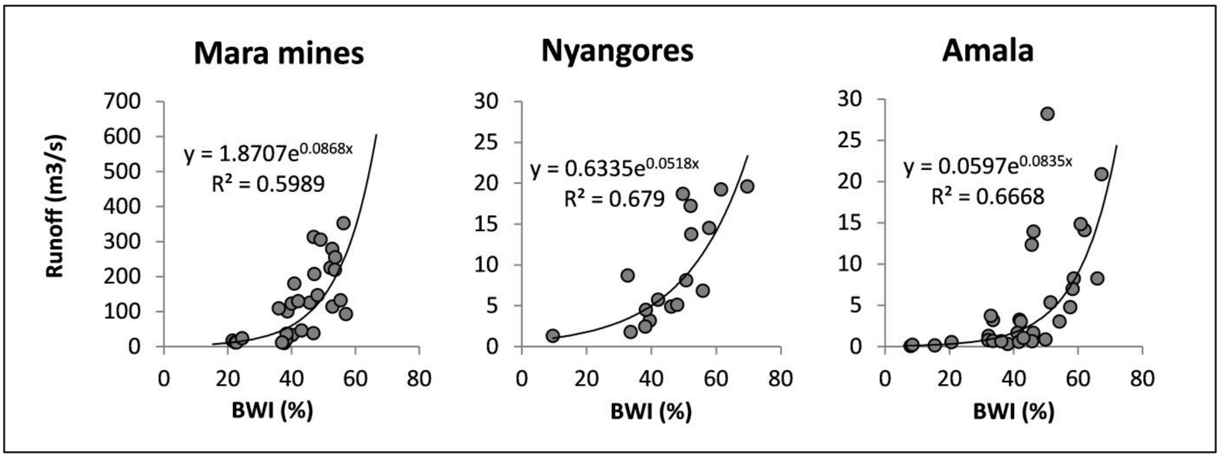



Figure 9: Runoff versus BWI scatter plots for Mara mines, Nyangores and Amala sub catchments. The plots are for 30 day daily means. The steepness of the slope is taken to indicate the water storage capacity of the catchment. A steep slope indicates low catchment storage capacity and vice versa. Nyangores is shown to have the highest storage capacity. The study period was between January 2007 and July 2013.

### 3.3. TRMM rainfall versus in situ measured rainfall

TRMM data was validated by investigating its relationship with the in-situ rainfall measurements. The investigation was performed through comparison of TRMM and in-situ measured 30 days summations on a pixel to point measurement basis for selected RfGSs in the MRB. The period for the investigation was between January 2007 and July 2013. $R^2$ and Root Meat Square Error were used to assess the relationship. The number of RfGSs with sufficient data and their spatial distribution (see also figure 6) was not

sufficient for conducting a conclusive spatially averaged comparison. Only six RfGSs were considered for this analysis. 30 day summations were used so as to smooth the high spatial-temporal variability. MRB being in the tropics has intense but less spatially distributed rainfall events. For all the six stations, a linear relationship was found. The $R^2$ and RMSE varied from 0.5 to 0.86, and 47.7 mm/30 days to 73.7 mm/30 days respectively (see table 6). Ilkerin Integral Development Project RfGS (09135025) had the highest $R^2$ (0.86) and lowest RMSE (29.8 mm/30 days). However the results for this RfGS were not considered conclusive since very few points were used for the analysis. Governor's Camp RfGS (09135026) had the highest RMSE of 73.7mm/30 days despite having a high $R^2$. The Olenguruone D.O's Office - Molo (09035085) and Tenwek Mission RfGS (09035079) were observed to be systematically underestimating rainfall with the rest overestimating with respect to a 1:1 linear relationship with TRMM estimates, (figure 10). These results indicate the uncertainty brought by lack of homogeneity in the quality of the RfGS.
Previous TMM validation studies have shown that TRMM fairly estimates rainfall. Wolff et al. (2005) in their research to analyse TRMM with tipping bucket rain gauges over central Florida, found out that the correlation was better on monthly and yearly scales than on shorter time scales. They attributed the low correlation on shorter time scales to the difference in spatial and temporal sampling modes.

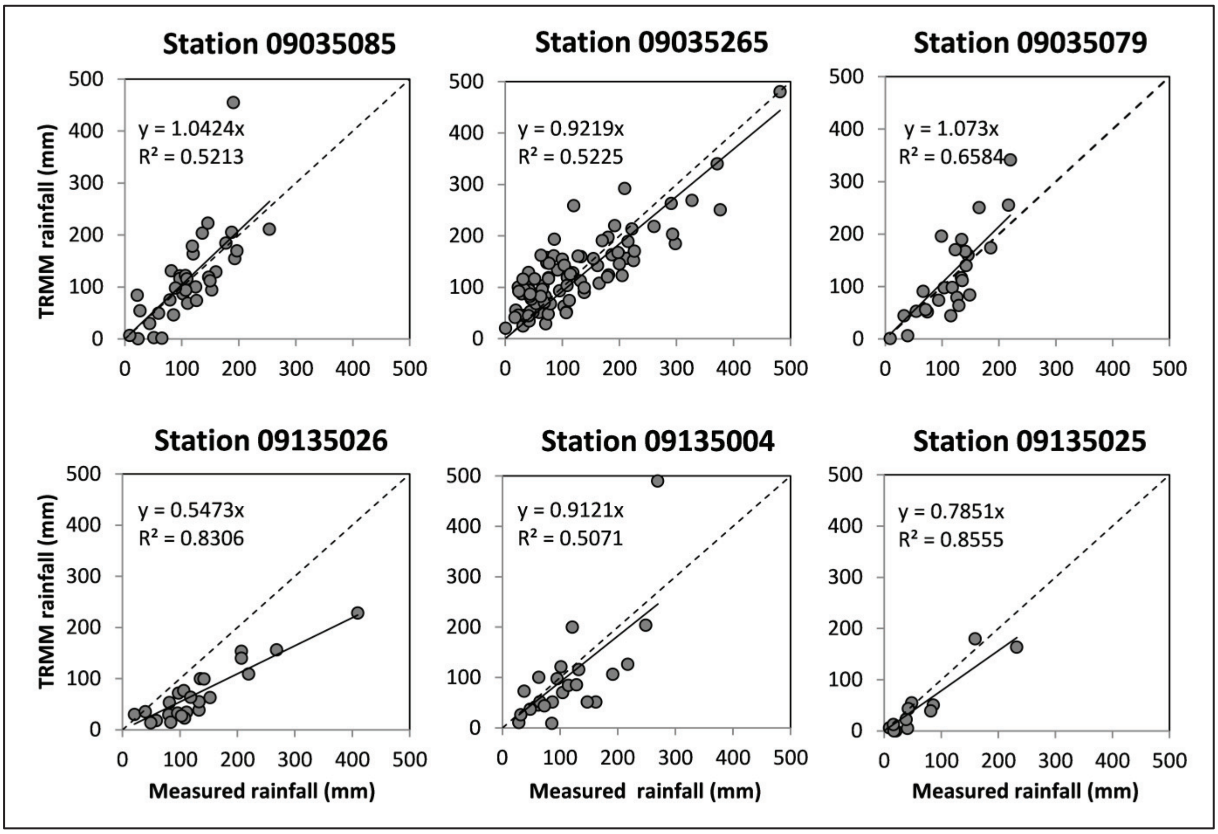



Figure 10: TRMM versus in-situ measured rainfall for selected RfGS scatter plots. The data plotted is pixel (TRMM) to in-situ measurements computed to 30 day summations. The solid line represents the linear relationship of TRMM and in-situ measurement while the dotted line represents a 1:1 linear relationship. The period for the investigation was between January 2007 and July 2013.

Prasetia et al. (2013) in their research on validation of TRMM estimates in the region over Indonesia conducted a similar investigation and found out that there was a medium correlation between the TRMM and in-situ measurements. Prakash and Gairola (2013) in their research on TRMM rainfall validation over the tropical Indian ocean with measured rainfall on a daily time scale also found good correlation with RMSE varying from 1 to 22 mm $d^{-1}$.

Table 6: TRMM rainfall validation with in-situ measured rainfall. The period for the investigation was between January 2007 and July 2013. There are gaps in the in-situ measurements for all the RFGS in the basin.

| Station Name | Station Code | Location | | No of points | $r^2$ | RMSE of 30 day summations (mm) |
|---|---|---|---|---|---|---|
| | | Latitude °S | Longitude °E | | | |
| Tenwek Mission – Sotik | 09035079 | -0.75 | 35.37 | 26 | 0.66 | 47.7 |
| Olenguruone D.O's Office – Molo | 09035085 | -0.58 | 35.68 | 37 | 0.52 | 57.8 |
| Bomet Water Supply | 09035265 | -0.78 | 35.35 | 93 | 0.52 | 51.7 |
| Oltome Green Lodge – Narok | 09135004 | -1.07 | 35.52 | 23 | 0.50 | 69.2 |
| Ilkerin Integral Development Project | 09135025 | -1.78 | 35.70 | 12 | 0.86 | 29.8 |
| Governor's Camp | 09135026 | -1.28 | 35.03 | 24 | 0.83 | 73.7 |

### 3.4. Water budgets at Mara mines, Amala and Nyangores

Annual water budgets at Mara mines, Amala and Nyangores sub catchments were calculated from 1998 up to 2012 (see table 7). The rainfall summations were calculated from the TRMM time series data sets. Evapotranspiration (ET) was calculated as residual from rainfall and runoff. A lot of gaps were noted in the runoff data as from 1990 especially on the Mara Mines RGS. This made it difficult to account for runoff in each year hence only a non-conclusive analysis could be performed. Historical data for Nyangores and Amala rivers were used to analyse the seasonal runoff trend. Figure 11 shows the monthly mean runoff for Nyangores and Amala rivers for the period 1964 to 1992. From this figure it is shown that Amala River has a higher and early peak runoff than Nyangores. It is also shown that Nyangores has higher base flow compared to Amala. Two distinct peak runoff seasons corresponding to the MRB wet season can also be observed.

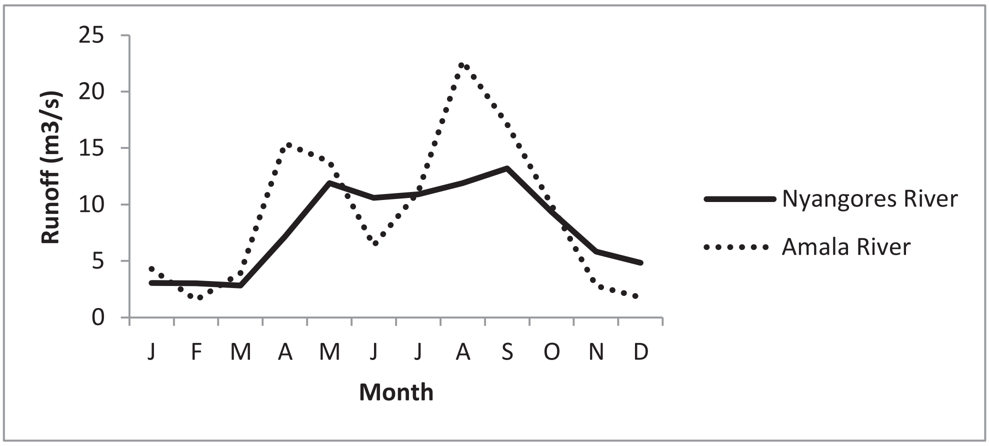



Figure 11: Monthly mean runoff for Nyangores and Amala rivers for the period 1964 to 1992. The Amala River has higher and early peak runoff compared to Nyangores. The latter however has higher base flow

Surprisingly, for the period under consideration, Mara mines has a higher mean annual TRMM rainfall of 1403 mm compared to Amala which has an annual mean of 1360 mm with a standard deviation of 148 mm and 178 mm respectively. Nyangores was found to have the highest mean annual TRMM rainfall with a standard deviation of 196mm.

Table 7: Water budgets for Mara mines, Nyangores and Amala sub catchments. There are many gaps in the runoff data hence these budgets are not conclusive

| | **Runoff annual summations (mm/year)** | | | **TRMM Rainfall annual summations (mm/year)** | | | **ET annual summations - Residue (mm/year)** | | |
|---|---|---|---|---|---|---|---|---|---|
| Year | Mara mines | Nyangores | Amala | Mara mines | Nyangores | Amala | Mara mines | Nyangores | Amala |
| 1998 | | 668 | | 1290 | 1405 | 1320 | 1290 | 737 | |
| 1999 | | 380 | 113 | 1374 | 1416 | 1250 | | 1036 | 1137 |
| 2000 | | 192 | 94 | 1243 | 1247 | 1083 | | 1055 | 989 |
| 2001 | | 632 | 406 | 1426 | 1581 | 1430 | | 949 | 1024 |
| 2002 | | 395 | 314 | 1473 | 1558 | 1416 | | 1163 | 1102 |
| 2003 | | 509 | | 1319 | 1448 | 1314 | | 939 | |
| 2004 | | 370 | 185 | 1294 | 1371 | 1216 | | 1001 | 1031 |
| 2005 | 254 | 415 | 392 | 1178 | 1270 | 1132 | 924 | 855 | 740 |
| 2006 | 355 | 465 | | 1681 | 1921 | 1682 | 1326 | 1456 | |
| 2007 | | 628 | 882 | 1581 | 1765 | 1580 | | 1137 | 698 |
| 2008 | | 413 | | 1422 | 1561 | 1402 | | 1148 | |
| 2009 | | 129 | | 1214 | 1273 | 1105 | | 1144 | |
| 2010 | 598 | | | 1459 | 1703 | 1518 | 861 | | |
| 2011 | | | | 1586 | 1661 | 1503 | | | |
| 2012 | | | | 1505 | 1611 | 1448 | | | |
| **Mean** | **403** | **433** | **341** | **1403** | **1519** | **1360** | **1100** | **1052** | **960** |
| **Max** | **598** | **668** | **882** | **1681** | **1921** | **1682** | **1326** | **1456** | **1137** |
| **min** | **254** | **129** | **94** | **1178** | **1247** | **1083** | **861** | **737** | **698** |
| **Std. dev.** | **177** | **165** | **270** | **148** | **196** | **178** | **242** | **182** | **173** |

# 4. RAINFALL-RUNOFF MODEL

For this research, the conceptual lumped soil moisture-runoff model by Meier et al. (2011) is used. The model includes two linear storage reservoirs, namely surface and subsurface storage layer. The inputs to this model are BWI and satellite rainfall data. This model does not account for catchment heterogeneity and land use/cover (Meier et al., 2011). The concept of the model is that BWI is the state variable directing rainfall into surface and groundwater runoff production pathways. This relationships is expressed as follows,

$$I_{GW} = k_1 A\, R(t)\big(1 - BWI(t)\big) \tag{3}$$

Where, $I_{GW}$ is the infiltration to the subsurface storage ($m^3\,d^{-1}$), A is the area of the catchment ($km^2$), R is the rainfall ($mm\,d^{-1}$), $k_i$ is a model parameter ($d^{-1}$) and $t$ is the model time step (d). As the soil becomes more saturated, more rainfall is routed through surface storage. Similarly as the soil becomes less saturated, less rainfall is routed through subsurface storage.

The change in surface and groundwater storage over a time step is calculated as follows,

$$\frac{\Delta S_s(t)}{\Delta t} = k_1 A\, R(t) BWI(t) - I_{GW} - k_2 S_s(t-1) \tag{4}$$

$$\frac{\Delta S_{GW}(t)}{\Delta t} = \max(I_{GW} + k_3(BWI(t) - BWI(t-1); 0) - k_4 S_{GW}(t-1) \tag{5}$$

Where, $S_S$ and $S_{GW}$ are the surface and subsurface storage components, respectively ($m^3d^{-1}$), $\Delta S_S$ and $\Delta S_{GW}$ are the change in surface storage and subsurface storage components, respectively ($m^3d^{-1}$).

These surface and groundwater storage change equations are linked to their respective water budget as follows,

$$S_s(t) = \frac{\Delta S_s(t)}{\Delta t} + S_s(t-1) \tag{6}$$

$$S_{GW}(t) = \frac{\Delta S_{GW}(t)}{\Delta t} + S_{GW}(t-1) \tag{7}$$

The runoff components are subsequently computed from the storage components as follows,

$$Q_s(t) = k_2 S_s(t-1) \tag{8}$$

$$Q_{GW}(t) = k_4 S_{GW}(t-1) \tag{9}$$

Where, $Q_S$ and $Q_{GW}$ are the surface and groundwater runoff components respectively. Summing the two components and routing provides the total runoff production,

$$Q(t) = Q_s(t - \Delta\tau_s) + Q_{GW}(t - \Delta\tau_{GW}) \tag{10}$$

Where, $Q$ is the total runoff ($m^3d^{-1}$), $\Delta\tau_S$ and $\Delta\tau_{GW}$ are the surface and subsurface time lags respectively (d).

This set of equations is schematically represented in Figure 12.

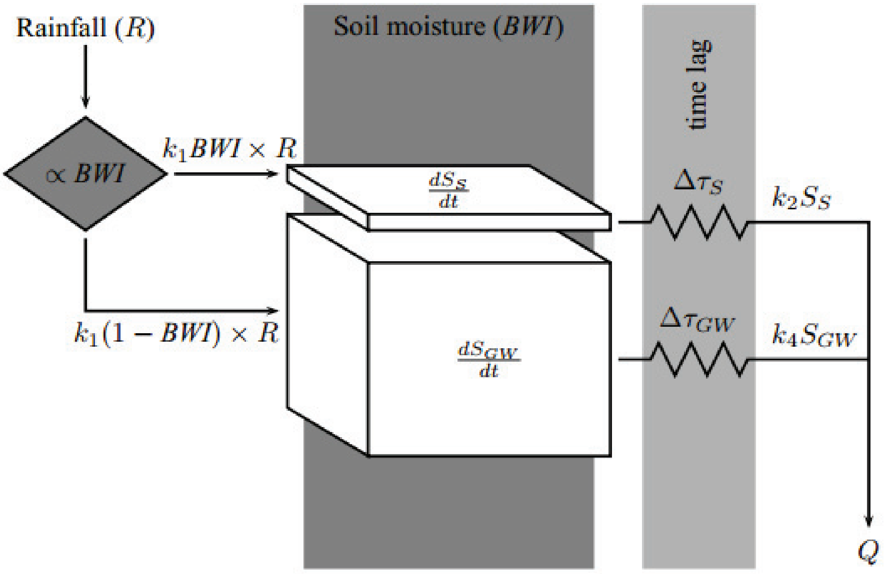



Figure 12: The input in this set up is rainfall. The level of BWI determines the distribution of rainfall between the surface and subsurface storage compartment. As the BWI increases, infiltration reduces and more rainfall is routed to the surface storage. The two storage compartments contribute to surface and groundwater runoff. Source: Meier et al. (2011)

The physical meaning of the empirical parameters as described by (Meier et al., 2011) are: $k_1$ parameterizes the initial loss of rainfall due to evaporation and interception ($d^{-1}$); $k_2$ parameterizes the retention of water in the surface storage before being routed to the river ($d^{-1}$); $k_3$ quantifies retention of water in the surface storage before being transported to the subsurface storage ($d^{-1}$); and $k_4$ is the rate of depletion from the subsurface storage to the river ($m^3d^{-1}$). The $k_i$ parameters are dependent on soil infiltration properties and catchment average retention time that is influenced by topography, geography and vegetation. According to Meier et al. (2011), $\Delta\tau_S$ and $\Delta\tau_{GW}$ are dependent on the size of catchment, but from equations (4-10) the uniqueness with respect to $k_2$ and $k_4$ is not clear.

In this research, the Meier et al. (2011) model has been modified by replacing the delay factor with a low pass filter approach. The low pass filter attenuates the storage components as a function of time before they are routed as runoff, as shown in equations 11 and 12. In this new approach, the contribution of previous rainfall events is factored. The reasoning is that contribution of a particular rainfall event is not instantaneous but rather increases exponentially over a given time before reaching a peak value. The total runoff is consequently computed as shown in equation 13.

Where:

$$Q_s(t) = \frac{\sum_{i=(t-1)}^{n} S_S e^{-(t-i)/\tau}}{\sum_{i=(t-1)}^{n} e^{-(t-i)/\tau}} \tag{11}$$

$$Q_{GW}(t) = \frac{\sum_{i=(t-1)}^{n} S_{GW} e^{-(t-i)/\tau_g}}{\sum_{i=(t-1)}^{n} e^{-(t-i)/\tau_g}} \tag{12}$$

$$Q(t) = Q_s(t) + Q_{GW}(t) \tag{13}$$

Where $\tau$ and $\tau_g$ are the characteristic catchment response times (d) related to the surface and groundwater runoff respectively.

# 5. MODEL PERFORMANCE

In this research, TRMM was converted from mm $d^{-1}$ to m $d^{-1}$. Consequently the storage and runoff components were also expressed in m $d^{-1}$. The $k_i$ parameters were expressed in $d^{-1}$. The model was built in excel spread sheet. The model simulations were done on a daily time step.

## 5.1. Calibration and Sensitivity Analysis

The model was initially calibrated for Mara mines by optimising by optimising $k_i$ parameter and the surface and groundwater runoff catchment response times $\tau$ and $\tau_g$ with respect to bias, RMSE, Mean Absolute Error (MAE) and correlation coefficient *r*. This was done with the solver tool in excel using least-square fitting. The initial calibration values for $k_1$, $k_2$, $k_3$ and $k_3$ were 0.16 $d^{-1}$, 0.991 $d^{-1}$, 0.005 $d^{-1}$ and 0.99 $d^{-1}$ respectively. The characteristic times $\tau$ and $\tau_g$ were 2.5 d and 180 d respectively. These results prompted the need to check the sensitivity of the model to these parameters. For a given parameter, the sensitivity was investigated by varying the parameter, while keeping the others at initial values. Increasing $k_1$ from the initial value it was observed that there was no change in *r* while RMSE, MAE and Bias increased accordingly. Reducing $k_1$, RMSE, MAE and Bias reduced slightly up to at $k_1$~0.08 then started increasing. Only a slight decrease of *r* was observed. Increasing or reducing $k_2$ and $k_4$ from their initial values very slight change in *r*, RMSE, MAE and Bias was observed. Increasing $k_3$ the initial value, it was observed that *r* very slightly reduced while RMSE, MAE and bias notably increased accordingly. Changing of the time lags $\tau_S$ and $\tau_{GW}$ was found to have negligible effect on RMSE, *r* and bias. However the initial $\tau$ and $\tau_g$ values had to be fixed to realistic values. From this analysis it shows that setting both $k_2$ and $k_4$ to 1 and $k_3$ to zero improves the overall performance of the model. The sensitivity analysis results are illustrated in figure 13. These results clearly indicate that $k_1$ is the only parameter that the model is sensitive to.

This research concluded that only the $k_1$ parameter and the surface and groundwater runoff characteristic time $\tau$ and $\tau_g$ were necessary. Subsequently, calibration of the $k_1$ parameter was done for Mara mines, Nyangores and Amala sub-catchment. The catchment response time $\tau$ and ($\tau_g$) for Mara mines, Nyangores and Amala were fixed at, 3, (200), 1, (100), 1, (100) respectively (table 8). These values were assumed taking into consideration the size of the given sub-catchment and field observations. For each of the three sub-catchments, different calibration periods and durations had to be used due to the large data gaps, as shown in figure 14 and 15. The periods were selected to capture at least one wet season. The calibrated $k_1$ for Mara mines, Nyangores and Amala were 0.074, 0.157 and 0.179 respectively. $k_1$ is a factor accounting for losses especially due to interception and evapotranspiration (Meier et al., 2011). $k_1$ of 1 indicates a situation with zero losses, a case where rainfall is equal to runoff. $k_1$ of 0 indicates a case where all the rainfall is lost hence no runoff. $k_1$ for Mara mines is notably lower compared with the other sub catchments. This indicates that there are more losses within Mara mines compared to Nyangores and Amala. Amala has the highest $k_1$ compared to Nyangores. The difference in $k_1$ for the two can be attributed to land cover. Amala is highly deforested and has higher peak flows compared to Nyangores (Mango et al., 2011; Mati et al., 2008). It can, thus, be assumed that few losses due to evapotranspiration occur in Amala as most of the rainfall is drained as quick runoff. Nyangores with big forest area has higher retention period consequently, it is expected to have higher ET than Amala. The latter has the poorest calibration results with respect to the RMSE, MAE, Bias and *r* compared to Mara mines and Nyangores. Overall, Nyangores had the best calibration results. However, it has higher bias error compared to Mara mines. As shown in figure 15, the model does not capture the peak flows very well, especially for the Amala and Mara mines sub catchments. This explains the high RMSE for these two sub catchments compared to Nyangores. The model is shown to simulate low flows fairly well in Mara mines and

Nyangores. From figure 15 it can be observed that the model initially overestimates runoff in all the three sub-catchments. It can be argued that initially the soil may be dry hence less runoff, a situation not taken into consideration by the model. It can also be observed that in some instances, the peaks for measured runoff and TRMM are not coinciding for example for Nyangores towards the end of the calibration period.

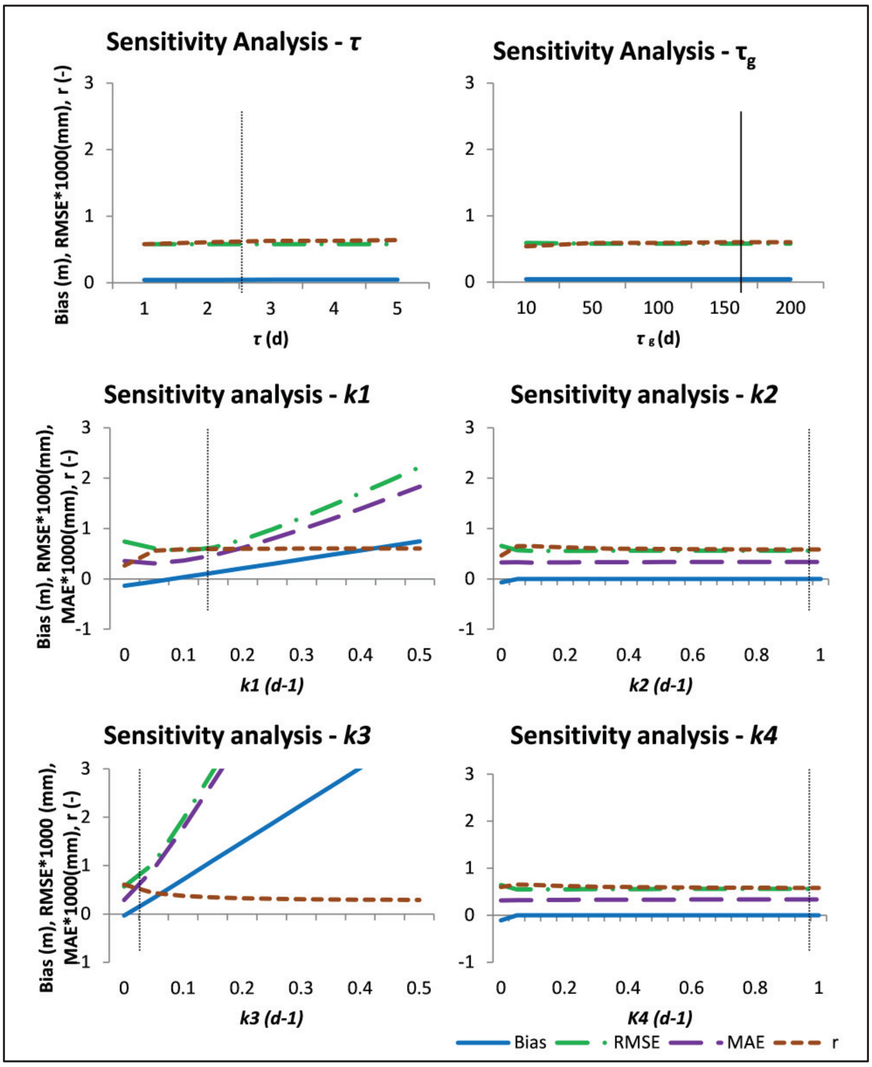



Figure 13: Sensitivity analysis results. The dotted vertical line indicates the initial values of the $k_i$ parameters and time lags $\tau$ and $\tau_g$. Only $k_1$ was found to be sensitive. $k_2$ and $k_4$ were equated to 1 and $k_3$ to zero.

Table 8: Calibration results for Mara mines, Nyangores and Amala sub catchments. High $k_1$ indicates high losses in the sub catchment and vice versa

| | units | Mara mines | Nyangores | Amala |
|---|---|---|---|---|
| $k_1$ parameter | $d^{-1}$ | 0.074 | 0.157 | 0.179 |
| $\tau_S$ | d | 3 | 1 | 1 |
| $\tau_{GW}$ | d | 200 | 100 | 100 |
| RMSE | m | 0.00034 | 0.00025 | 0.00081 |
| MAE | m | 0.0002 | 0.00018 | 0.00058 |
| Bias | m | 0.00052 | 0.01184 | 0.027 |
| $r$ | - | 0.54 | 0.67 | 0.125 |

Previous studies in the MRB have support our argument, for example Dessu and Mellesse (2012) in their study on modelling rainfall runoff processes in the MRB using Soil and Water Assessment Tool (SWAT), found that Amala sub catchment was consistently giving poor simulation results. They attributed the poor results to uncertainties in either measured rainfall or measured runoff input data.

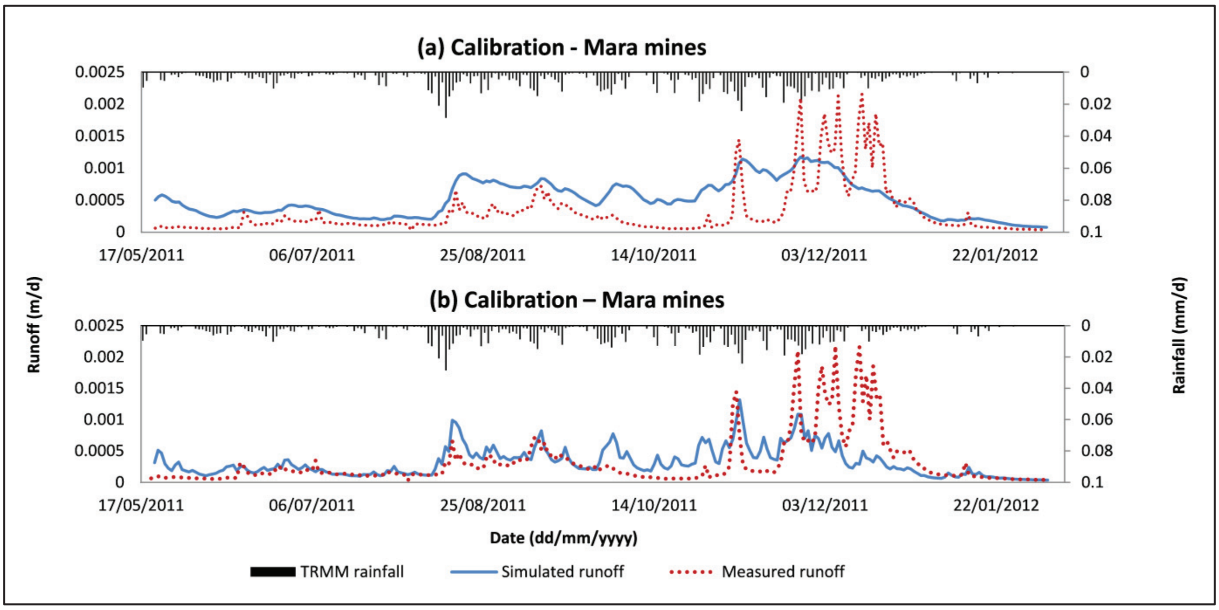



Figure 14: Comparison of simulation and measured runoff during calibration results at Mara mines. (a) is simulations with $k_1$, $k_2$, $k_3$ and $k_3$ parameters while (b) shows simulations taking into consideration only the $k_1$ parameter. The model simulations are shown to improve with reduction of the parameters.

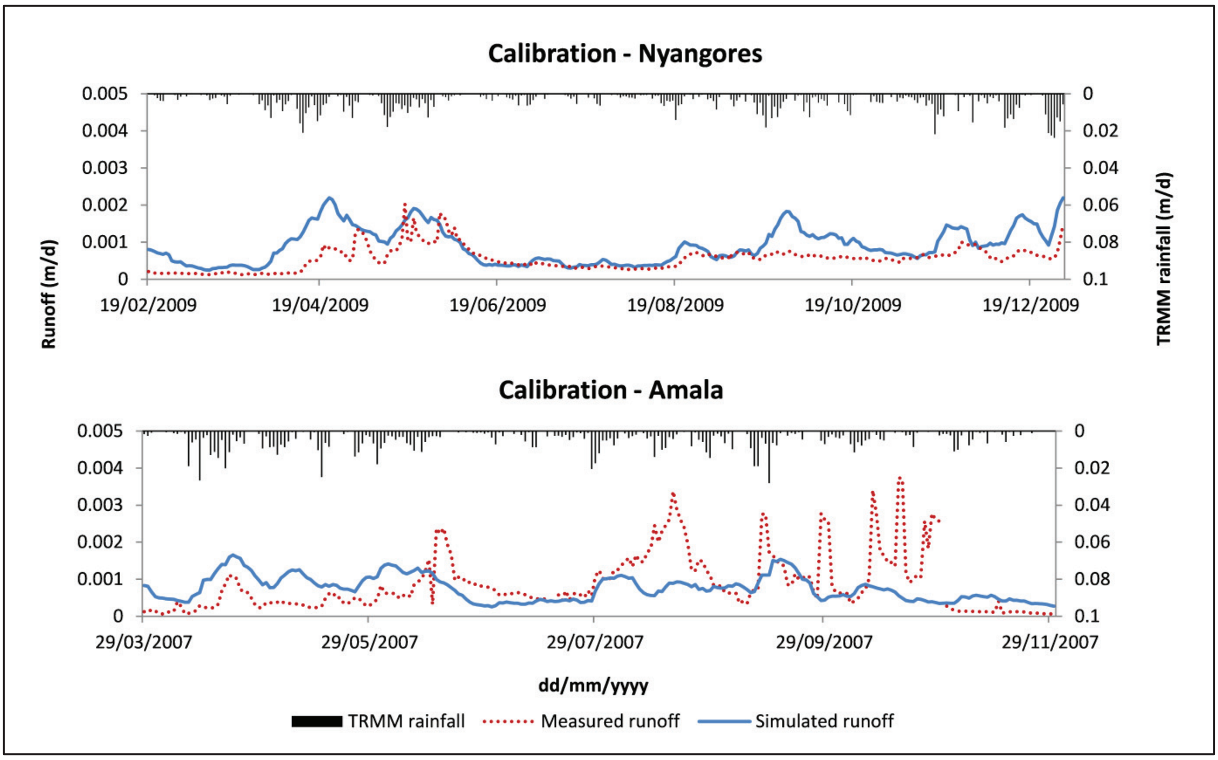



Figure 15: Measured and simulated runoff for Nyangores and Amala sub-catchments during calibration of the model with only the $k_l$ parameter. The model fairly simulates low flows than peak flows.

## 5.2. Validation

Validation was done by running the model with the respective optimised $k_1$ parameter and the $\tau$ and $(\tau_g)$ for each of the sub-catchments (see table 9). Different validation periods and durations had to be used for each sub-catchment, as shown in figures 16. The validation periods were selected to capture at least a wet season. The results show that there was an increase in RMSE, MAE and bias during validation compared to calibration results for all the three sub-catchments. However, $r$ notably improved. The $r$ for Mara mines, Nyangores and Amala sub-catchments were 0.77, 0.74 and 0.48 respectively. As shown in figure 16, the model poorly simulated peak flows especially in Amala and Mara mines. The model tends to overestimate the low flows notably in Amala and towards the end of the validation period for Nyangores. However from the figure, it can clearly be seen that the measured runoff is not responding to the low rainfall events during the dry periods as expected. This supports the early arguments that there are a lot of uncertainties in the measured runoff data.

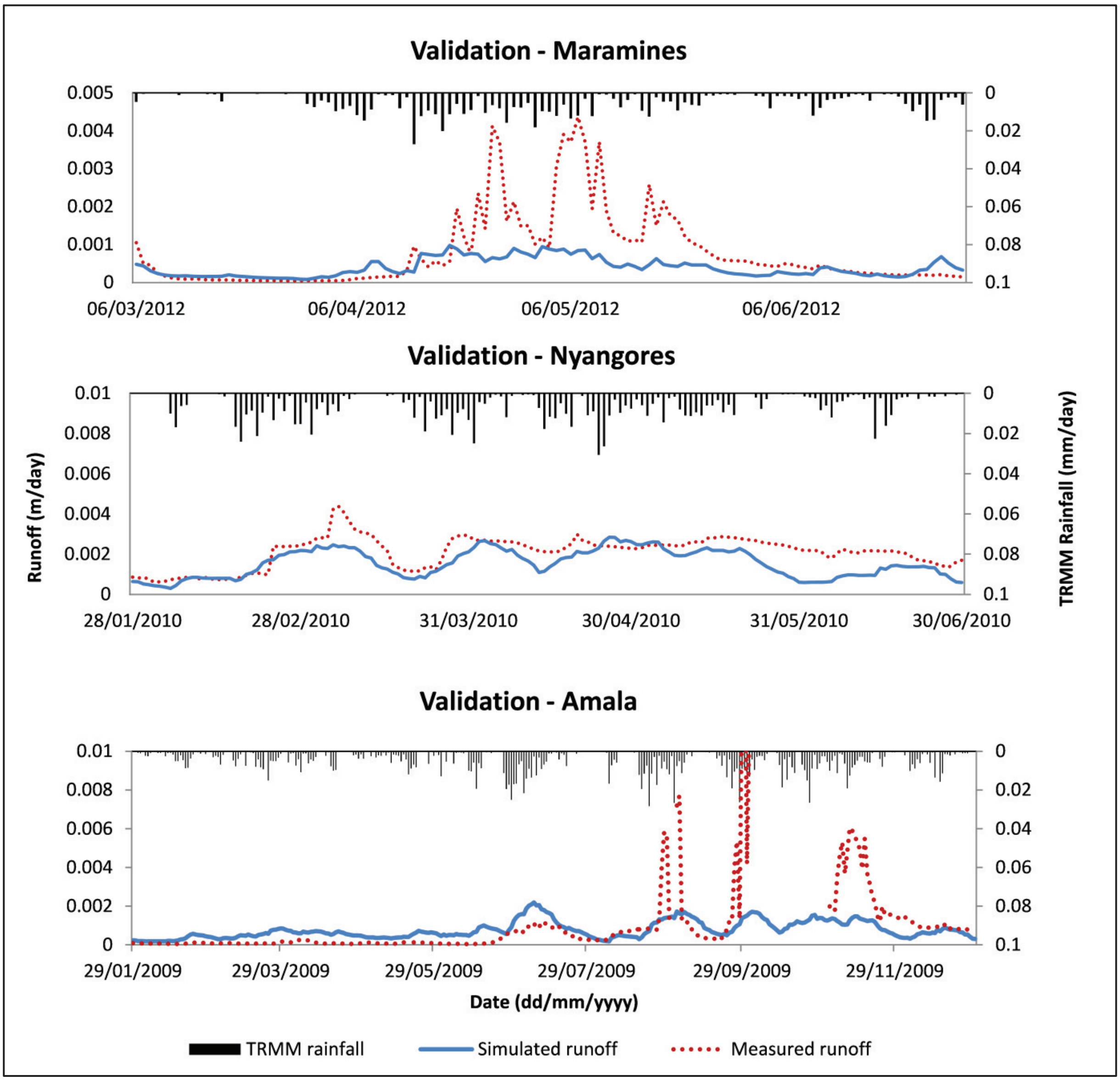



Figure 16: Measured and simulated runoff for Mara mines, Nyangores and Amala sub catchments during validation. The model poorly simulates quick peak runoff in Mara mines and Amala. It is also overestimating low flows in Amala

Table 9: validation results for Mara mines, Nyangores and Amala sub catchments. Mara mines has the lowest RMSE, MAE and bias and the highest *r*.

| | Units | Mara mines | Nyangores | Amala |
|---|---|---|---|---|
| RMSE | m | 0.00094 | 0.00132 | 0.00176 |
| MAE | m | 0.00051 | 0.00121 | 0.00079 |
| Bias | m | 0.048 | 0.186 | 0.062 |
| *r* | - | 0.77 | 0.74 | 0.48 |

## 5.3. Long term runoff simulations

Long term runoff simulations were done for the period from January 2007 to July 2013 on a daily time step for each of the three sub catchments. Figure 17 shows the measured and simulated total runoff for the three sub catchments. In this figure, at Mara mines and Amala, the model is observed to have poorly simulated the peak flows just as during calibration and validation. Notably, at Nyangores the model was able to fairly simulate the peak flows. This supports the arguments of Meier et al. (2011) that this model is not suitable for catchments with relatively low storage capacity and quick peak runoffs. For example, as mentioned earlier, Amala sub-catchment has been highly deforested hence it has high peak flows compared to Nyangores which has more forest areas. Nyangores is thus expected to have higher infiltration and storage capacity.

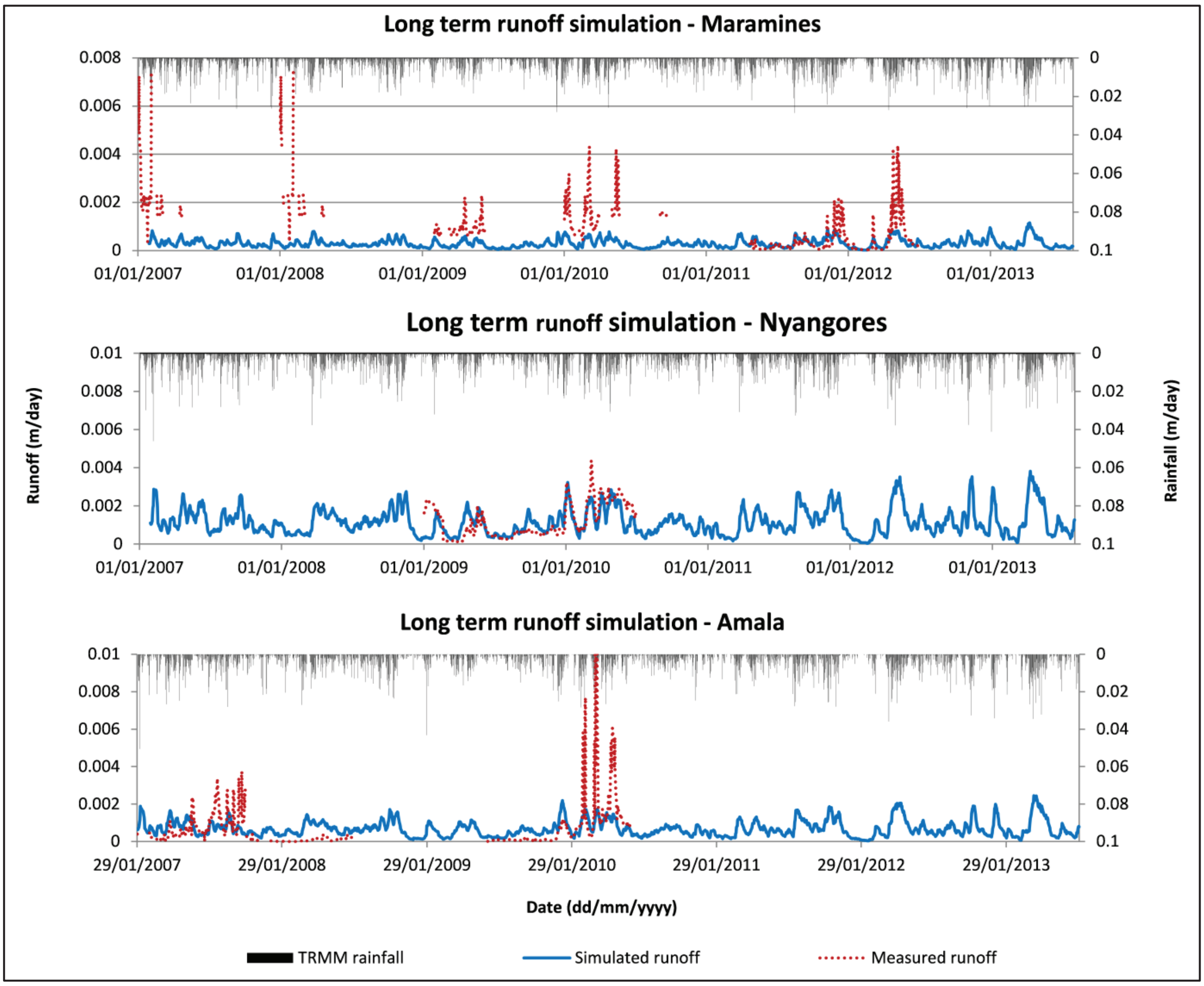



Figure 17: Long term runoff simulations on a daily model time step for Mara mines, Nyangores and Amala sub catchments. The simulation period is from January 2007 up to July 2013.

Analysis of soil samples collected during the fieldwork and data from Harmonized World Soil Database (HWSD)(FAO et al., 2012) indicate that the middle and lower area of Mara mines is dominantly sandy clay and clay loam. During field work, it was observed that the area was prone to flooding, proving that there is low infiltration and quick runoff. Figure 18 shows the monthly summation time series of simulated surface and groundwater runoff. From this figure it is observed that Nyangores generates more surface and groundwater runoff per unit area compared to the other sub catchments. Quantitative analysis of the simulations showed that Nyangores sub-catchment generates 54%, Amala 32% and Mara mines 14% of the total runoff in MR. This further supports our argument that there is apparently higher infiltration in Nyangores compared to the other sub catchments. This is also in agreement with previous studies which indicate that Nyangores has higher base flows than Amala (Dessu et al., 2014; Dessu & Mellesse, 2012; Mango et al., 2011; Mati et al., 2008). In this research, groundwater runoff is assumed to be the base flow. Mara mines has the lowest surface and groundwater runoff generation per unit area. This may be an indication of relatively higher losses in the Mara mines sub-catchment.

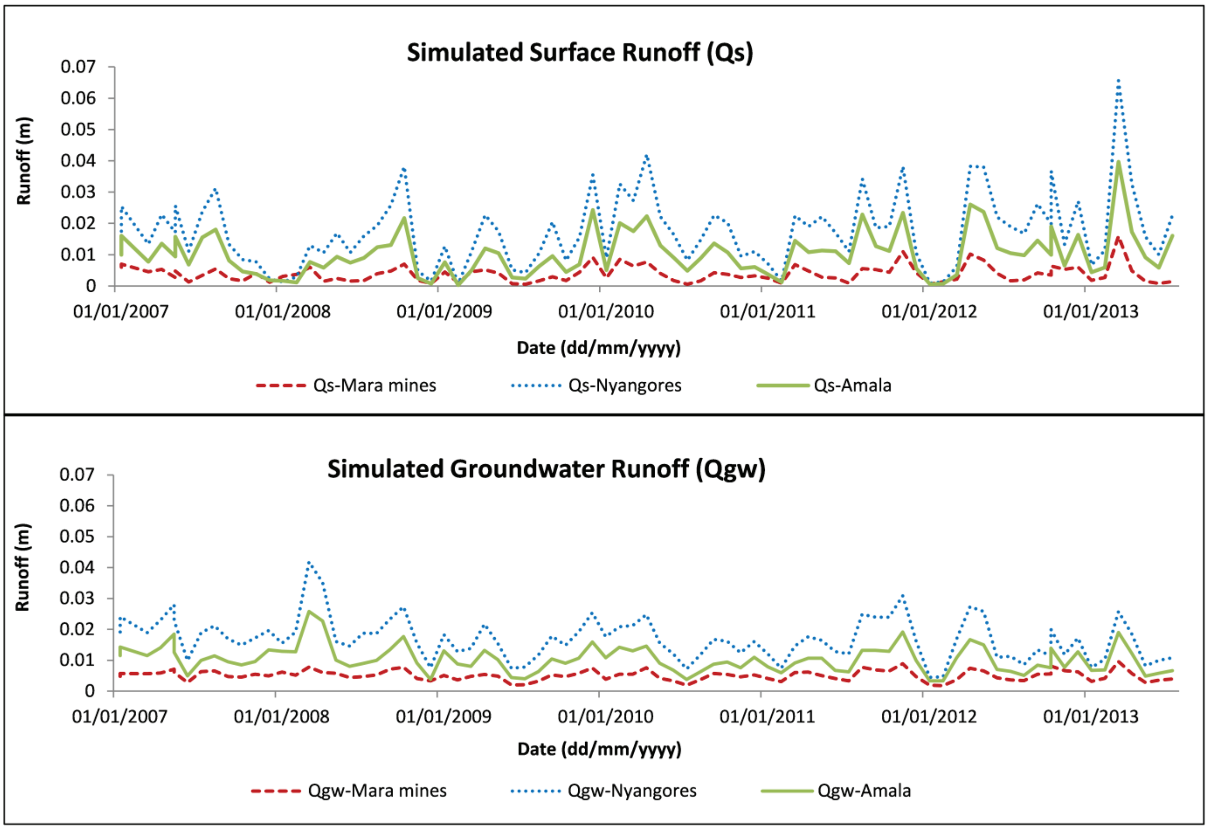



Figure 18: Comparison of the monthly summations of simulated surface and groundwater runoff components for Mara mines, Nyangores and Amala sub catchments. The simulation period is from January 2007 up to July 2013. Nyangores generates the highest surface and groundwater runoff while Mara mines generates the least.

## 5.4. Error analysis

Further analysis was done to investigate the behaviour of bias error with respect to rainfall. The analysis was done by plotting bias error with rainfall as shown in figure 19. The period of the analysis was different for each sub-catchment. The results showed that the bias of the simulations were high during high rainfall and low during low rainfall with an exception of Amala where this behaviour is seen in the last half of the analysis period. The exceptional behaviour in Amala further strengthens our argument that there are a lot of uncertainties in the Amala in-situ measurement runoff data. The results of this analysis are in agreement with the study of Meier et al. (2011) who in their analysis found that the absolute error of simulations was high during the wet seasons.

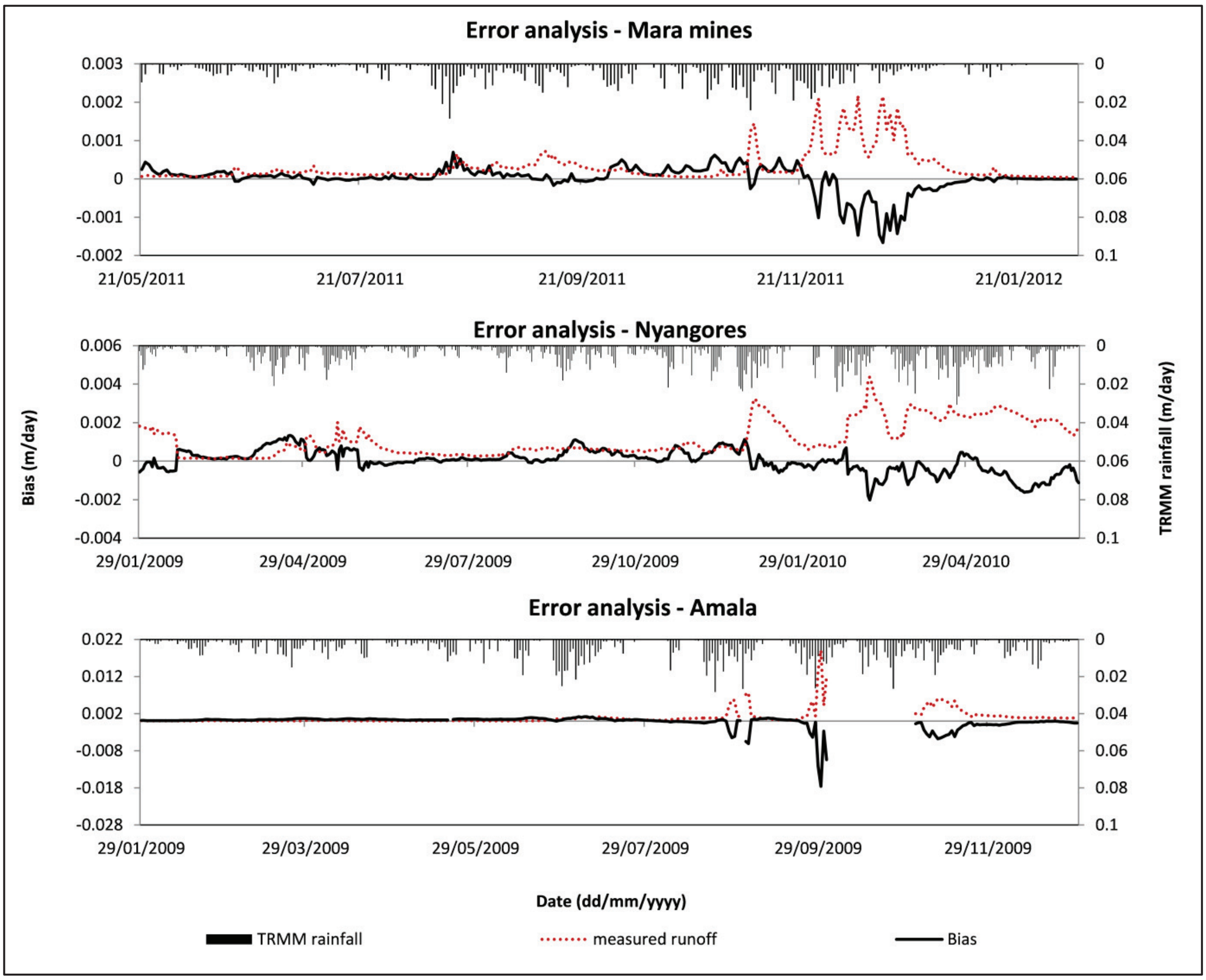



Figure 19: Error analysis at Mara mines, Nyangores and Amala. The bias error of the simulations tends to be propagated by intensity of the rainfall events.

# 6. DISCUSSION

A visual analysis of the BWI, TRMM and measured runoff time series curves indicated that these data sets were following a seasonal trend. A quantitative analysis of the time series data using scatter plots showed that there is indeed good relationship between the three. Surprisingly, even in the largely forested Nyangores and Amala sub-catchments, the moisture retrievals had good relation with TRMM and measured runoff. Previous studies on retrieval of soil moisture from scatterometer backscatter suggest that there are a lot of uncertainties when retrieving from tropical rain forests (Brocca et al., 2010; Wagner et al., 1999). This is an indication that, even though the satellite based soil moisture and rainfall products are independently observed, they both capture the seasonal runoff variations generated by rainfall. Previous studies on the relationship of satellite based soil moisture products and runoff have also shown a correlation.

This research has simplified the Meier et al. (2011) soil moisture-runoff concept further by reducing the parameters to be calibrated from four to only one. The research found that the model performance was only sensitive to the parameter linked to losses due to ET ($k_1$). As much as parameters may have physical meanings, over parameterisation of hydrological models does not necessarily improve their performance, but rather complicates them (Beven, 1996). Previous studies suggest that it is crucial to minimise the parameters to be calibrated as much as possible (Refsgaard, 1997; Refsgaard & Knudsen, 1996). Refsgaard (1997) in a study on parameterisation, calibration and validation of distributed hydrological models went further to suggest that parameterisation should be linked to specific hydrological conditions and data availability.

This research further improved the Meier et al. (2011) rainfall-runoff by introducing a low pass filter concept. In this concept, the generated storage components are passed through the filter before being routed to runoff. The concept makes the generation of runoff more realistic by factoring the contribution of previous rainfall events to runoff as a function of specified time lag. During sensitivity analysis, it was found out that the time lags could be fixed without affecting the overall performance of the model. However the time lags were found critical when it came to simulation of peak flows. Analysis of the three sub catchments pointed out that they have short response time consequently reasonable limits were set. For Mara mines, Nyangores and Amala, the surface runoff time lags were fixed at 3 days, 1 day and 1 day respectively. This was also supported by observations during the fieldwork where flooding was observed in the lower parts of the basin barely three days after a rainfall event in the upper basin.

Mara mines had the lowest $k_1$ (0.074), indicating that there are more losses in this sub catchment than in Nyangores and Amala which had a $k_1$ of 0.157 and 0.179 respectively. This is supported by the fact that Mara mines consist of large relatively flat semi-arid regions with savannah vegetation typically associated with high ET. Results of the soil analysis also showed that the soils in this region are pre-dominantly clay loam and sandy clay. The implication of this is that the region is poorly drained hence more water is available for ET. Similarly, the small $k_1$ for Nyangores, which according to literature has more forest cover, proves that it has higher ET losses than Amala. The reasoning is that the forest cover promotes infiltration hence more water is available for ET. Similarly, Amala with less forest cover and steep slopes quickly drains most of the rainfall as quick runoff with little left for ET. An analysis of the long term simulated runoff components shows that Nyangores and Amala which make about 12% of the total area of Mara mines contribute about 54% and 32% of the total simulated runoff in the sub-catchment respectively. Comparison of the groundwater runoff components shows that Nyangores generates the highest volumes followed by Amala and Mara mines in that order. These findings are supported by results from an analysis of historical data and also findings from previous studies which indicate that Nyangores has relatively higher base flows compared to Amala (Dessu et al., 2014; Mango et al., 2011; Mati et al., 2008). This clearly demonstrates that the dry season runoff of the MR is largely sustained by groundwater storage of the two upstream sub-catchments. During validation and calibration, the model is observed to perform relatively well in Nyangores than in the other sub-catchments and poorly in Amala. It is also observed to simulate dry season runoff better than the wet season runoff. These results support the arguments of Meier et al. (2011) that this model is not suitable for catchments with low storage and quick peak runoff. However, for Mara mines with the lowest storage the model is observed to perform better than in Amala. This is indicates that the peak flows could be having more influence on the model performance than the catchment storage capacity. It can also be argued that the uncertainties noted in the runoff data of Amala by this research  and previous research (Dessu & Mellesse, 2012) could have contributed to the overall poor performance in this sub-catchment. There was a feeling that the model results would have been better if the quantity and quality of the in-situ runoff measurements for all the three sub-catchment was better.

These results are a further proof that satellite-observed soil moisture has great potential in hydrological modelling at catchment scale (Ceballos et al., 2005; Khan et al., 2012; Meier et al., 2011).

# 7. FINAL REMARKS

## 7.1. Conclusions

The overall objective of this research was to develop a runoff simulation model based on satellite observed soil moisture and rainfall as a forcing data in the MRB. Based on the research questions the research was answering in its bid to achieve this objective, we make the following conclusions:

- There is relationship between satellite observed soil moisture and rainfall, and measured runoff of the Mara mines, Nyangores and Amala. The relationship between soil moisture versus rainfall is logarithmic while that of runoff versus soil moisture is exponential. The runoff simulation model was developed on the basis of these relationships.
- The model found was found to be sensitive only model parameter $k_1$ is. This research has found that this parameter is affected by catchment characteristics like, for example, land cover, infiltration/groundwater storage capacity, which affects ET. Mara mines assumed to have the highest ET has the lowest $k_1$. Amala assumed to have the least ET has the highest $k_1$.
- The performance of the model is found to be affected by catchment's infiltration capacity and quality of in-situ measured runoff. The relatively good performance of the model in Nyangores is attributed to the catchment's apparent high infiltration capacity compared with the other two sub-catchments.
- The model is found to be applicable in monitoring dry season runoff even in catchments with low storage capacities like Mara mines. However it has a weakness in simulating wet season flows in such catchments.

## 7.2. Recommendations

Drawing from the findings and lessons from this research, we make the following recommendations.

- To reduce the uncertainty brought by in-situ measured runoff data, investigations on the possible calibration of the data with satellite based ET data and rainfall need to be done including the assessment of the reliability of the satellite products.
- Investigation should be done on the validation of the simulated groundwater storage components with in-situ borehole piezometric measurements and terrestrial water storage data from for example Gravity Recovery and Climate Experiment (GRACE) and Global Land Data Assimilation System (GLDAS) model.
- To improve simulation of the peak flows, there is a need to do more investigation on improving the routing process of the model. It would also be interesting to see the effect of having separate loss factor for surface and ground storage components to the overall performance of the model.

# LIST OF REFERENCES

Beven, K. (1996). A Discussion of Distributed Hydrological Modelling. In M. Abbott & J. Refsgaard (Eds.), *Distributed Hydrological Modelling* (Vol. 22, pp. 255-278): Springer Netherlands.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., . . . Bittelli, M. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sensing of Environment, 115*(12), 3390-3408. doi: http://dx.doi.org/10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., Moramarco, T., Wagner, W., & Hasenauer, S. (2010). ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy. *Remote Sensing of Environment, 114*(11), 2745-2755. doi: http://dx.doi.org/10.1016/j.rse.2010.06.009

Brocca, L., Morbidelli, R., Melone, F., & Moramarco, T. (2007). Soil moisture spatial variability in experimental areas of central Italy. *Journal of Hydrology, 333*(2), 356-373.

Ceballos, A., Scipal, K., Wagner, W., & Martínez-Fernández, J. (2005). Validation of ERS scatterometer-derived soil moisture data in the central part of the Duero Basin, Spain. *Hydrological Processes, 19*(8), 1549-1566. doi: 10.1002/hyp.5585

Dessu, S. B., Melesse, A. M., Bhat, M. G., & McClain, M. E. (2014). Assessment of water resources availability and demand in the Mara River Basin. *CATENA, 115*(0), 104-114. doi: http://dx.doi.org/10.1016/j.catena.2013.11.017

Dessu, S. B., & Mellesse, A. M. (2012). Modelling the rainfall–runoff process of the Mara River basin using the Soil and Water Assessment Tool. *Hydrological Processes, 26*(26), 4038-4049. doi: doi: 10.1002/hyp.9205

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., . . . Oevelen, P. v. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences, 15*(5), 1675-1698.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Hegyiová, A., Sanchis-Dufau, A., . . . Drusch, M. (2013). Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network. *Vadose Zone Journal*.

Engman, E. T., & Chauhan, N. (1995). Status of microwave soil moisture measurements with remote sensing. *Remote Sensing of Environment, 51*(1), 189-198.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., . . . Van Zyl, J. (2010). The Soil Moisture Active Passive (SMAP) Mission. *Proceedings of the IEEE, 98*(5), 704-716. doi: 10.1109/jproc.2010.2043918

FAO, IIASA, ISRIC, ISSCAS, & JRC. (2012). Harmonized World Soil Database (version 1.2). from FAO, Rome, Italy and IIASA, Laxenburg, Austria

Gereta, E., Mwangomo, E., & Wolanski, E. (2009). Ecohydrology as a tool for the survival of the threatened Serengeti ecosystem. *Ecohydrology & Hydrobiology, 9*(1), 115-124. doi: http://dx.doi.org/10.2478/v10104-009-0035-7

Juston, J., Jansson, P.-E., & Gustafsson, D. (2013). Rating curve uncertainty and change detection in discharge time series: case study with 44-year historic data from the Nyangores River, Kenya. *Hydrological Processes*, n/a-n/a. doi: 10.1002/hyp.9786

Khan, S. I., Yang, H., Vergara, H. J., Gourley, J. J., Brakenridge, G. R., De Groeve, T., . . . Bin, Y. (2012). Microwave Satellite Data for Hydrologic Modeling in Ungauged Basins. *Geoscience and Remote Sensing Letters, IEEE, 9*(4), 663-667. doi: 10.1109/lgrs.2011.2177807

Krhoda, G. O. (2005). The Hydrology of Mara River. *WWF–EASTERN AFRICA REGIONAL PROGRAMME OFFICE (EARPO) MARA RIVER CATCHMENT BASIN INITIATIVE*.

Liu, Z., Ostrenga, D., Teng, W., & Kempler, S. (2012). Tropical Rainfall Measuring Mission (TRMM) Precipitation Data and Services for Research and Applications. *Bulletin of the American Meteorological Society, 93*(9), 1317-1325. doi: 10.1175/bams-d-11-00152.1

Mango, L. M., Melesse, A. M., McClain, M. E., Gann, D., & Setegn, S. G. (2011). Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: Results of a modeling study to support better resource management. *Hydrology and Earth System Sciences, 15*(7), 2245-2258. doi: 10.5194/hess-15-2245-2011

Mati, B. M., Mutie, S., Gadain, H., Home, P., & Mtalo, F. (2008). Impacts of land-use/cover changes on the hydrology of the transboundary Mara River, Kenya/Tanzania. *Lakes & Reservoirs: Research & Management, 13*(2), 169-177. doi: 10.1111/j.1440-1770.2008.00367.x

Meier, P., Fromelt, A., & Kinzelbach, W. (2011). Hydrological real-time modelling in the Zambezi river basin using satellite-based soil moisture and rainfall data. *Hydrology and Earth System Sciences, 15*(3), 999-1008. doi: 10.5194/hess-15-999-2011

NASA. (2011). Explanation of Data Products Retrieved 28 August 2013, from http://pmm.nasa.gov/data-access/data-products

Parrens, M., Zakharova, E., Lafont, S., Calvet, J. C., Kerr, Y., Wagner, W., & Wigneron, J. P. (2012). Comparing soil moisture retrievals from SMOS and ASCAT over France. *Hydrol. Earth Syst. Sci., 16*(2), 423-440. doi: 10.5194/hess-16-423-2012

Prakash, S., & Gairola, R. M. (2013). Validation of TRMM-3B42 precipitation product over the tropical Indian Ocean using rain gauge data from the RAMA buoy array. *Theoretical and Applied Climatology*, 1-10. doi: 10.1007/s00704-013-0903-3

Prasetia, R., As-syakur, A. R., & Osawa, T. (2013). Validation of TRMM Precipitation Radar satellite data over Indonesian region. *Theoretical and Applied Climatology, 112*(3-4), 575-587. doi: 10.1007/s00704-012-0756-1

Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology, 198*(1–4), 69-97. doi: http://dx.doi.org/10.1016/S0022-1694(96)03329-X

Refsgaard, J. C., & Knudsen, J. (1996). Operational Validation and Intercomparison of Different Types of Hydrological Models. *Water Resources Research, 32*(7), 2189-2202. doi: 10.1029/96wr00896

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., . . . Namkhai, A. (2000). The Global Soil Moisture Data Bank. *Bulletin of the American Meteorological Society, 81*(6), 1281-1299. doi: 10.1175/1520-0477(2000)081<1281:tgsmdb>2.3.co;2

Scipal, K., Scheffler, C., & Wagner, W. (2005). Soil moisture-runoff relation at the catchment scale as observed with coarse resolution microwave remote sensing. *Hydrology and Earth System Sciences, 9*(3), 173-183.

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes, 17*(15), 3163-3170. doi: 10.1002/hyp.5155

USDA. (1999). *Soil taxonomy : a basic system of soil classification for making and interpreting soil surveys* (Second edition ed. Vol. 436). Washington, D.C.: United States Department of Agriculture, Natural Resources Conservation Service.

USGS. (2012, 24 January 2012). Shuttle Radar Topography Mission (SRTM) Retrieved 29 August, 2013, from https://lta.cr.usgs.gov

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., . . . Rubel, F. (2013). The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications. *Meteorologische Zeitschrift, 22*(1), 5-33. doi: 10.1127/0941-2948/2013/0399

Wagner, W., Lemoine, G., & Rott, H. (1999). A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment, 70*(2), 191-207. doi: http://dx.doi.org/10.1016/S0034-4257(99)00036-X

Wagner, W., Lemoine, G., & Rott, H. (1999b). A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data. *Remote Sensing of Environment, 70*(2), 191-207. doi: http://dx.doi.org/10.1016/S0034-4257(99)00036-X

Wolff, D. B., Marks, D. A., Amitai, E., Silberstein, D. S., Fisher, B. L., Tokay, A., . . . Pippitt, J. L. (2005). Ground Validation for the Tropical Rainfall Measuring Mission (TRMM). *Journal of Atmospheric and Oceanic Technology, 22*(4), 365-380. doi: 10.1175/jtech1700.1

# APPENDICES

Appendix A: IDL script to read subset and convert SWI data from hdf5 to tiff

```
;***********************************************************
PRO read_SWI
;-----------------------------------------------------------
; Image coordinates
;-----------------------------------------------------------
ilat = 89.95
ilon = -179.95
res  = 0.1
;-----------------------------------------------------------
; Define study area and File output and input paths
;-----------------------------------------------------------
ulat = -0.125
llat = -1.875
llon = 33.625
rlon = 35.875
x_st  = uint((llon-ilon)/res)
y_st  = uint((ilat-ulat)/res)
x_en  = uint((rlon-ilon)/res)
y_en  = uint((ilat-llat)/res)
s_xsize = abs(x_en-x_st)
s_ysize = abs(y_en-y_st)
print, x_en,x_st
print, y_en,y_st
;-----------------------------------------------------------
i_path    = 'F:\G2_SWI_hdf\'
o_path    = 'F:\G2_SWI_tif\'
i_files   = file_search(i_path, 'g2_BIOPAR_SWI_*')
b_files   = file_basename(i_files, 'g2_BIOPAR_SWI_*')
n_f        = n_elements(i_files)
print, n_f
;-----------------------------------------------------------
; Read variables
;-----------------------------------------------------------
n_v       = 1
var       = strarr(n_v)
var(0)    = 'SWI_001'
for f = 0, n_f-1 do begin
print, f
  file  = i_files(f)
  h5_open
  fid=h5f_open(file)
  finfo=h5_parse(file)
  ntags=n_tags(finfo)
;-----------------------------------------------------------
; Determine Study area
;-----------------------------------------------------------
  for v = 0, n_v-1 do begin
    v_name  = var(v)
    for itag=0,ntags-1 do begin
        if (size(finfo.(itag),/type) eq 8) then begin
          ; if variable is a structure
        if (finfo.(itag)._type eq 'GROUP') then begin
          ; and then if it is a group, get all its members
          n=h5g_get_nmembers(fid,finfo.(itag)._name)
                 for i=0,n-1 do begin
              name=h5g_get_member_name(fid,finfo.(itag)._name,i)
              data1_id=h5d_open(fid,finfo.(itag)._NAME+'/'+name)
              data1=h5d_read(data1_id)
                if name EQ v_name  then begin
                if v EQ 0 then begin
                  dims  = size(data1)
                  xs    = dims(1)
                  ys    = dims(2)
                  all_data = fltarr(xs,ys,n_v)
                  all_data(*,*,v) = data1(*,*)
                 endif else begin
                  all_data(*,*,v) = data1(*,*)
                 endelse
```

```
            endif
         endfor ; i
      endif
    endif
  endfor
endfor ; v
h5_close ; Close all identifiers
dims = size(all_data)
i_xsize = dims(1)
i_ysize = dims(2)
output  = fltarr(s_xsize, s_ysize)
;-------------------------------------------------------------
; location study area
;-------------------------------------------------------------
  for x = x_st, x_en-1 do begin
    for y = y_st, y_en-1 do begin
      p_x = x - x_st
      p_y = y - y_st
      if all_data(x,y) LT 255 then begin
        output(p_x,p_y) = (all_data(x,y)/255) * 100
      endif else begin
        output(p_x,p_y) = -99.
      endelse
    endfor
  endfor
o_file =o_path+ b_files(f) + '.tif'
write_tiff, o_file, output, /FlOAT, PLANARCONFIG=2
endfor ; f
;--------------------------------------------------------------
End
;--------------------------------------------------------------
```

Appendix B: IDL script for creating a map-list (same modified was modified for TRMM data)

```
;****************************************************************
PRO create_stack
;----------------------------------------------------------------
close, 1
;----------------------------------------------------------------
i_path = 'F:\G2_SWI_tif\'
;----------------------------------------------------------------
o_file = 'F:\G2_Stack_tiff\SWI_stack.tif'
in_files    = file_search(i_path,'*.tif')
n_f         = n_elements(in_files)
;----------------------------------------------------------------
data  = read_tiff(in_files(0), INTERLEAVE = 2, GEOTIFF = GEO)
;----------------------------------------------------------------
dims        = size(data)
xsize       = dims(1)
ysize       = dims(2)
output      = fltarr(xsize, ysize, n_f)
;----------------------------------------------------------------
  for f = 0, n_f-1 do begin
   data = read_tiff(in_files(f), INTERLEAVE = 2, GEOTIFF = GEO)
  output (*,*,f) = (data(*,*))
   endfor
;----------------------------------------------------------------
  print, n_f, ' ', o_file
  write_tiff, o_file, output, /float, PLANARCONFIG= 2
;----------------------------------------------------------------
End
;----------------------------------------------------------------
```

Appendix C: IDL script to read subset and convert TRMM data from NetCDF to tiff

```
;****************************************************************
PRO read_TRMM
;----------------------------------------------------------------
;*Define, path and file name*
;----------------------------------------------------------------
out_path = 'D:\sat_prec\TRMM\subset\
in_path = 'D:\sat_prec\TRMM\wget-1.11.4b\'
```

```
;------------------------------------------------------------
; Define the study area
;------------------------------------------------------------
ulat = -0.125
llat = -1.875
;------------------------------------------------------------
llon = 33.625
rlon = 35.875
;------------------------------------------------------------
ysize = 400
xsize = 1440
res   = 0.25
;------------------------------------------------------------
m_files = FILE_SEARCH (in_path, '3B42_daily.*')
b_files = file_basename (m_files, '*.7.nc')
nr_files = n_elements (m_files)
;------------------------------------------------------------
; Count and open the files
;------------------------------------------------------------
nr_files =n_elements (m_files)
for t=5626,nr_files-1 do begin;
  print, t,'   ' , m_files(t), nr_files
  ncid = NCDF_OPEN(m_files(t))
;------------------------------------------------------------
; Read NetCDF data
;------------------------------------------------------------
  NCDF_VARGET, ncid, 3, rain      ; Read pixel values
  NCDF_VARGET, ncid, 2, lat       ; Read Latitude
  NCDF_VARGET, ncid, 1, lon       ; Read Longitude
;------------------------------------------------------------
  t_da    = fltarr(xsize, ysize)
;------------------------------------------------------------
; transform data
;------------------------------------------------------------
  for j = 0, ysize-1 do begin
    p_j = (ysize-1) - j
    for i = 0, xsize-1 do begin
      t_da[i,p_j] = rain[i,j]
    endfor
  endfor
  ;------------------------------------------------------------
  ; find the study area needed
  ;------------------------------------------------------------
  nl = size(lat)
  ns = size(lon)
   for l = 0, nl(1)-1 do begin
    p_l = (nl(1)-1) - l
    if lat(p_l) EQ llat then y_en = l
    if lat(p_l) EQ ulat then y_st = l
  endfor
  for s = 0, ns(1)-1 do begin
    if lon(s) EQ llon then x_st = s
    if lon(s) EQ rlon then x_en = s
  endfor
 ;------------------------------------------------------------
 ; create memory for output
 ;------------------------------------------------------------
  x_size = x_en - x_st
  y_size = y_en - y_st

  output = fltarr(x_size,y_size)
 ;------------------------------------------------------------
 ; write tiff file
 ;------------------------------------------------------------
   for i = x_st, x_en-1 do begin
    for j = y_st, y_en-1 do begin
        p_i = i - x_st
        p_j = j - y_st
        if (t_da[i,j] GE 0.) and (t_da[i,j] LE 1000.) then begin
        output[p_i,p_j] = t_da[i,j]   ;   Read rainfall [m3/m3]
        endif else begin
           output[p_i,p_j] = t_da[i,j] ;   Read rainfall [m3/m3]
        endelse
    endfor
  endfor
```

```
o_file =out_path+ b_files(t) + '.tif'
write_tiff, o_file, output, /FlOAT, PLANARCONFIG=2
endfor ;t
END
;-----------------------------------------------------------------
```

Appendix D: IDL script to create an ASCII file from tiff map-list (same modified was modified for TRMM data)

```
;****************************************************************
PRO create_stack_ascii
;----------------------------------------------------------------
close, 1
;----------------------------------------------------------------
i_path = 'F:\G2_SWI_tif\'
o_file = 'F:\G2_SWI_tif\SWI_SCAT.txt'
;----------------------------------------------------------------
in_files    = file_search(i_path,'*.tif')
n_f         = n_elements(in_files)
print, n_f
stop
;----------------------------------------------------------------
data  = read_tiff(in_files(0), INTERLEAVE = 2, GEOTIFF = GEO)
;----------------------------------------------------------------
dims        = size(data)
xsize       = dims(1)
ysize       = dims(2)
output      = fltarr(xsize* ysize, n_f)
;----------------------------------------------------------------
  for f = 0, n_f-1 do begin
    data = read_tiff(in_files(f), INTERLEAVE = 2, GEOTIFF = GEO)
  count = 0
  for x = 0, xsize-1 do begin
    for y = 0, ysize-1 do begin
        output(count,f) = (data(x,y))
        count = count + 1
    endfor
  endfor
  endfor
 ;----------------------------------------------------------------
  openw, 1, o_file
  printf, 1, output, FORMAT = '(414F15.4)'
  close, 1
;-----------------------------------------------------------------
End
;-----------------------------------------------------------------
```