

**ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES**

ADDIS ABABA INSTITUTE OF TECHNOLOGY

**VALIDATION OF GLOBAL PRECIPITATION DATA SETS TO RUNOFF
GENERATION USING SWAT MODELING SYSTEM**

**A Thesis Submitted to the School of Graduate Studies of Addis Ababa University in
Partial Fulfillment of the Degree of Master of Science in Civil & Environmental
Engineering**

(Major in Hydraulic Engineering)

By Fenet Tadese Neme

Advisors: Dr. Dereje Hailu and Dr. Semu A. Moges

Addis Ababa

Ethiopia

December 2015



VALIDATION OF GLOBAL PRECIPITATION DATA SETS TO RUNOFF GENERATION USING SWAT MODELING SYSTEM

**A Thesis Submitted to the School of Graduate Studies of Addis Ababa University in
Partial Fulfillment of the Degree of Master of Science in Civil & Environmental
Engineering**

(Major in Hydraulic Engineering)

By Fenet Tadese Neme

Approved By Board of Examiners

	Signature	Date
1. Dr. Dereje Hailu (Advisor)	_____	_____
2. Dr. Semu Ayalew Moges (Advisor)	_____	_____
3. Dr. Agizew Nigusie (Internal examiner)	_____	_____
4. Dr. Yilma Seleshi (External examiner)	_____	_____
5. Dr. Esayas G/Youhannes (Chairman)	_____	_____

Acknowledgments

Above all I thank the almighty GOD for His guidance and grace upon me during all my works and in all my life.

I would also like to express my sincere thanks to my supervisors Dr. Dereje Hailu (Interim Director, Ethiopian Institute of water Resources) and Dr. Semu A. Moges for their valuable guidance, encouragement, suggestions and constructive comments from inception to the completion of this research work. I am also very thankful to Mr. Dejene & Mr. Haileyesus for their immediate response and assistance in providing satellite rainfall data.

I would also like to thank the Ethiopian National Meteorology Agency and Ministry of Water Irrigation and Energy for their help by providing me Meteorological and hydrological data respectively to conduct this study without which this research was not possible.

Furthermore, I would also want to give special thanks to Earth2Observe project which is funded by European Union at School of Civil and Environmental Engineering, AAiT, AAU, for sponsoring my study.

Last but not least, I am very grateful to my lecturers, families and friends who in one way or another supported me during the whole length of this research.

Abstract

The main objective of this study is to evaluate the performance of globally gridded high – resolution satellite rainfall products (TMPA 3B42V7, Corrected coarse CMORPH, and Adj-PERSIANN) under sparse ground based data and complex topography of Angar Watershed (3627.6 km²) in Upper Nile basin Ethiopia through semi-distributed hydrological model SWAT for monthly streamflow simulation.

The model is calibrated for the period of 2003 – 2007 and validated for the period of 2008 – 2010. Comparisons of the simulations to the observed streamflow at the outlet of Angar Watershed form the basis for the conclusions of this study. The Nash-Sutcliffe Efficiency (NSE), Coefficient of Determination (R^2), Root Mean Square Error (RMSE), linear correlation coefficient (R) and BIAS indices were used to benchmark the model performance.

Our results reveal that the performance of the satellite rainfall products as input to SWAT for monthly stream flow simulation strongly depends on the product type. Simulation from all rainfall inputs captured the trend of observed hydrograph. Simulation based on the new version of TMPA 3B42V7 and CMORPH showed consistent and modest skills in their simulations but slightly overestimates the large flood peaks in some years (2005 & 2006). On the other hand, the infrared – based algorithm PERSIANN simulation showed poor skills in reproducing the hydrograph at the outlet of the watershed. One of the reasons may be due to PERSIANN product is relatively less correlated to observed rainfall and other satellite products.

Overall, the results indicate that although some uncertainties exist in these gridded datasets (3B42V7 & CMORPH), the application of these gridded data prove useful for hydrological studies in the absence of station data. It strongly indicates the growing potential of the predictive accuracy of the satellite rainfall products in reproducing hydrological features.

Key words: Satellite rainfall products, SWAT model, Angar watershed

Abbreviations

ANN	-	Artificial Neural Network
ARS	-	Agriculture Research Service
CMORPH	-	Climate Prediction Center Morphing Method
CN	-	Curve Number
ENS	-	Nash-Sutcliffe Efficiency
HBV	-	Hydrologiska Byråns Vattenbalans-avdelning
HEC-HMS	-	Hydraulic Engineering Centre-Hydrologic Modeling System
HRU	-	Hydrological Response Units
IR	-	Infrared
Km ²	-	Square Kilometer
MW	-	Microwave
PERSIANN	-	Precipitation Estimation from Remotely Sensed Information Using Neural Networks
PR	-	Precipitation Radar
R ²	-	Coefficient of Determination
SCS	-	Soil Conservation Service
SOM	-	Self-Organizing Map
SWAT	-	Soil and Water Analysis Tool
TPMA	-	TRMM Multisatellite Precipitation Analysis
TRMM	-	Tropical Rainfall Measuring Mission
USDA	-	United States Department of Agriculture
WMO	-	World Meteorological Organization
WXGEN	-	Weather Generator

Table of Contents

Acknowledgments.....	ii
Abstract.....	iv
Abbreviations.....	v
Table of Contents.....	vi
List of Figures.....	ix
List of Tables.....	xi
CHAPTER ONE	1
1. Introduction	1
1.1. Background	1
1.2. Statement of the problem	3
1.3. Objectives.....	4
1.3.1. General objective.....	4
1.3.2. Specific objectives.....	4
1.4. Research questions	4
1.5. Organization of the thesis.....	5
CHAPTER TWO	6
2. Literature Review	6
2.1. Hydrological cycle and rainfall – runoff processes in a watershed.....	6
2.2. Rainfall – runoff modeling.....	7
2.3. Stream flow analysis	7
2.4. Description of satellite rainfall products	8
2.4.1. TMPA 3B42	8
2.4.2. CMORPH	9
2.4.3. PERSIANN.....	9
2.5. Application of high – resolution satellite rainfall products in hydrological modeling .	10
2.6. SWAT model.....	12

2.6.1. SWAT model application worldwide	13
2.6.2. SWAT model application in Ethiopia	14
CHAPTER THREE	16
3. Materials and Methods	16
3.1. Description of the study area	16
3.1.1. Location	16
3.1.2. Climate	17
3.1.3 Topography and geology	19
3.1.4. Soil types and land uses	20
3.2. Hydrological model selection criteria	21
3.2.1 Description of SWAT model	21
3.2.1.1. Surface runoff	23
3.2.1.2. Potential evapotranspiration	24
3.2.1.3. Ground water flow	25
3.2.1.4. Flow routing phase	25
3.3. Methodology	27
3.4. Model inputs data collection and analysis	28
3.4.1. Digital elevation model	28
3.4.2. Weather data	29
3.4.3. Satellite rainfall products	33
3.4.4. Stream flow data	33
3.4.5. Soil data	34
3.4.6. Land use and land cover	36
3.5. Model Setup	37
3.5.1. Watershed delineation	37
3.5.2. Hydrologic response units analysis	38
3.5.3. Weather generator	40

3.5.4. Sensitivity analysis	41
3.5.5. Model calibration.....	42
3.5.6. Model validation.....	44
3.6. Model performance evaluation.....	45
CHAPTER FOUR.....	48
4. Results and Discussion.....	48
4.1. Satellite rainfall products versus rain gauges	48
4.2. Stream flow modeling	54
4.2.1. Sensitivity analysis	54
4.2.2. Calibration and validation of stream flow simulation	54
4.3. Satellite rainfall simulation of stream flow	56
4.4. Discussion	61
CHAPTER FIVE	63
5. Conclusions and Recommendations	63
5.1. Conclusion.....	63
5.2. Recommendations	64
6. References	65
APPENDICES	68
Appendix I List of tables.....	68
Appendix II. List of figures	71
Appendix III. Symbols and description of weather generator parameters (WGEN) used by the SWAT model	73
Appendix IV. Soils parameter and legend used in SWAT model	74
Appendix V. Soils parameters values used in SWAT model	75
Appendix VI. Water balance parameters values of Angar watershed	76

List of figures

Figure 3.1 Location map of the area (Source: Own analysis).....	16
Figure 3.2 Mean monthly rainfall distribution of principal meteorological station for the period of 1992 - 2012.....	18
Figure 3.3 Mean annual rainfall from 1992- 2012.....	18
Figure 3.4 Slope classification of the watershed	19
Figure 3.5 Frame work of the study.....	27
Figure 3.6 Digital elevation model of Angar watershed.....	29
Figure 3.7 Meteorological stations around Angar watershed	30
Figure 3.8 Non-dimensional plot of selected stations around Angar watershed	31
Figure 3.9 Angar watershed with ten satellite rainfall grids ($0.25^0 \times 0.25^0$).....	34
Figure 3.10 Map of the soil types of Angar watershed.....	35
Figure 3.11 Land use/land cover map of the study area	37
Figure 3.12 Sub watersheds map of the Angar watershed.....	38
Figure 4.1 Annual (a) and percentage of seasonal (b) rainfall (i.e. JJAS: June, July, August, and September) from various rainfall products for the period of 2003 – 2007.	49
Figure 4.2 Monthly (a) and mean monthly (b) rainfall over Angar Watershed obtained from rain gauge data and various satellites rainfall products for the period of 2003 – 2007.	50
Figure 4.3 Intercomparison of daily rainfall from satellite rainfall products (i.e. 3B42V7, CMORPH, PERSIANN) and rain gauge for Angar watershed for the period of January 2003 – December 2007	52
Figure 4.4 Intercomparison of monthly rainfall from satellite rainfall products (i.e. 3B42V7, CMORPH, PERSIANN) and rain gauge for Angar watershed for the period of January 2003 – December 2007	53
Figure 4.5 Comparison of Arc SWAT simulated stream flow (based on rainfall input data from various sources in the legend) and observed monthly stream flow during the calibration period.	56
Fig. 4.6 Statistical comparison of simulations of monthly streamflow based on various rainfall inputs indicated in the legend.....	57
Figure 4.7 Intercomparison of simulated monthly stream flow (based on 3B42V7, CMORPH, PERSIANN and rain gauge rainfall input data) and observed monthly stream flow, the period of Jan, 2003 - Dec, 2007.....	59

Figure 4.8 Comparison of mean monthly rainfall estimates of various rain fall products versus simulations monthly stream flow (based on 3B42V7, CMORPH, PERSIANN and rain gauge rainfall input data), the period of Jan, 2003 - Dec, 2007.....60

List of tables

Table 2.1 Studies on performance of satellite rainfall products through hydrological modeling	10
Table 3.1 Meteorological station names, locations and variables	30
Table 3.2 Major soil distribution within Angar watershed	35
Table 3.3 Land use/cover classification of Angar watershed as per SWAT model	36
Table 3.4 SWAT parameters sensitivity class	42
Table 4.1 Default & final calibrated flow parameter values of the watershed	55

CHAPTER ONE

1. Introduction

1.1. Background

Weather measured directly by rain gauges or indirectly by remote sensing techniques, all precipitation estimates contain uncertainty. While rain gauges provide a direct measurement of precipitation reaching the ground, they may contain significant bias arising from coarse spatial resolution (yielding underestimation especially during events with low spatial coherency, i.e., convective showers), location, wind, and mechanical errors among others (Groisman and Legates 1994 cited by Yilmaz et al.).

Precipitation is the most important atmospheric input to land surface hydrology models, and therefore accurate precipitation inputs are essential for reliable hydrologic prediction. However, in many remote parts of the world and particularly in developing countries, ground-based precipitation measurements are either sparse or nonexistent, mainly because of the high cost of establishing and maintaining infrastructure. For rivers that cross international boundaries, inconsistencies in instrumentation and administrative limitations to data access further hamper the effective use of hydrology models in support of reliable flood and drought diagnosis and prediction (Ymeti, 2007). Both of these limitations are true for the Nile basin.

Satellite-based precipitation estimates with high spatial and temporal resolution and large areal coverage provide a potential alternative source of forcing data for hydrological models in regions where conventional in situ precipitation measurements are not readily available.

The growing availability of high resolution (and near real-time) satellite-based rainfall estimates has significant potential in applications such as hydrological analysis for engineering design, assimilation of precipitation information into forecast models, flood forecasting and water resources management in general.

These applications would even have far – reaching effects for several developing countries whose ground based – rain gauges are sparse and no radar technology for measuring representative rainfall magnitude. However, there are errors associated to satellite based rainfall estimates which prompt several scientific questions. How accurate are satellite based

rainfall products? Can high resolution satellite rainfall products be used for hydrological applications? Lack of knowledge on the accuracy of those satellite products is a challenge to the hydrological community especially under complex terrain and less gauged regions (Solomon, 2013).

During the last two decades, satellite-based instruments have been designed to collect observations mainly at thermal infrared (IR) and microwave (MW) wavelengths that can be used to estimate rainfall rates. Observations in the IR band are available in passive modes from (near) polar orbiting (revisit times of 1–2 days) and geostationary orbits (revisit times of 15–30 min), while observations in the passive and active MW band are only available from the (near) polar-orbiting satellites. A number of algorithms have been developed to estimate rainfall rates by combining information from the more accurate (but infrequent) MW with the more frequent (but less accurate) IR to take advantage of the complementary strengths (Bitew and Gebremichael, 2011).

The TMPA method (Huffman et al., 2007) uses MW data to calibrate the IR-derived estimates and creates estimates that contain MW-derived rainfall estimates when and where MW data are available and the calibrated IR estimates where MW data are not available. The TMPA products are available in two versions: real-time version (3B42RT) and post-real-time research version (3B42). The main difference between the two versions is the use of monthly rain gauge data for bias adjustment in the post-real-time research product. The 3B42 products are released 10–15 days after the end of each month, while the 3B42RT are released about 9 h after overpass. The CMORPH method (Joyce et al., 2004) obtains the rainfall estimates from MW data but uses a tracking approach in which IR data are used only to derive a cloud motion field that is subsequently used to propagate raining pixels.

The PERSIANN method (Sorooshian et al., 2000) uses a neural network approach to derive relationships between IR and MW data which are applied to the IR data to generate rainfall estimates. The resolutions of these (often dubbed as „high-resolution“) products are 0.25_ and 3 hourly, although finer resolutions are also available for CMORPH and PERSIANN. Besides these widely known products, there are also other high-resolution products, such as, Hydro-estimator (Scofield and Kuligowski, 2003), NRL-blended (Turk and Miller, 2005), PMIR (Kidd and Muller, 2009), and GSMaP (Ushio and Kachi, 2009).

It is well known that the satellite rainfall values are just estimates that are subject to a variety of error sources (gaps in revisit times, poor direct relationship between remotely sensed signals and rainfall rate, atmospheric effects that modify the radiation field) and require a thorough validation. The validation efforts can be grouped into two categories. The first is the direct comparison of the satellite rainfall estimates to the rain gauge networks and ground-based radar estimates (Dinku et al., 2007, 2008; Hirpa et al., 2010; Bitew and Gebremichael, 2010).

The second is the evaluation of satellite rainfall estimates based on their predictive ability of streamflow rate in a hydrological modeling framework. This has two advantages. One, since the evaluation is performed at the watershed scale, it is not subject to the scale discrepancy problem that arises when using rain gauge data for validation. Two, the satellite rainfall estimates are evaluated with respect to a specific application, as a driving input variable in a hydrologic model (Bitew and Gebremichael, 2011).

The purpose of this study is to assess the capability and limitation of satellite rainfall products as input into a hydrological model for run off simulation in a mountainous Upper Blue Nile Basin in Ethiopia, for three different satellite rainfall products in Angar watershed. This study will be limited to the following specific cases: the Soil and Water Assessment Tool (SWAT) hydrological model, Angar watershed (3627.6 km²), and three satellite precipitation products (3B42-V7, corrected coarse CMORPH and Adj-PERSIANN).

1.2. Statement of the problem

Ground based water resources monitoring systems are often difficult to maintain consistently in developing countries. The decline in the number of stations, data quality and changes in the data holding policy has made water resources data less reliable for use in operational purposes.

Obtaining representative meteorological data for watershed-scale hydrological modeling can be difficult and time consuming. Land-based weather stations do not always adequately represent the weather occurring over a watershed, because they can be far from the watershed of interest and can have gaps in their data series, or recent data are not available. An additional complication is that rain gauge data are effectively point measurements, which

may represent precipitation poorly across a watershed, particularly if there are large hydro-climatic gradients.

Moreover, weather records are seldom complete, which requires substituting other measurements or incorporating some sort of “estimated” weather conditions. Thus, there is a need to consider new methods to estimate weather conditions for watershed-scale modeling.

1.3. Objectives

1.3.1. General objective

The main objective of this study is to evaluate the performance of globally gridded high – resolution satellite rainfall products (TMPA 3B42V7, Corrected coarse CMORPH, and Adj-PERSIANN) under sparse ground based data and complex topography of Angar watershed (3627.6 km²) in Upper Nile basin Ethiopia through semi-distributed hydrological model SWAT for monthly streamflow simulation.

1.3.2. Specific objectives

To attain the above main objective the following specific objectives were derived.

- To evaluate streamflow simulation capability of high resolution satellite rainfall products over the watershed.
- To compare satellite rainfall products with ground based rain gauge rainfall both averaged over the watershed.
- Assessing the performance of semi-distributed hydrological model (SWAT) using ground truth and satellite data as input.

1.4. Research questions

To address the above objectives, the following research questions were designed.

- Can high – resolution satellite rainfall products accurately estimate rainfall compared to ground based rain gauge rainfall observation over Angar watershed?
- Can measurements from sparse rain gauges represent areal rainfall estimates of Angar watershed? Can they be used to evaluate the accuracy of high – resolution satellite rainfall products?
- How good are a high – resolution satellite rainfall products for prediction of streamflow?

1.5. Organization of the Thesis

The paper is organized into five chapters: chapter one is an introduction part where the background, statement of the problem, objectives of the study and research questions are discussed. In chapter two, review of related literatures where hydrological cycle and rainfall-runoff process in a watershed, rainfall-runoff modeling, stream flow analysis, description of satellite rainfall products, application of high-resolution satellite rainfall products in hydrological modeling and application of SWAT model worldwide and in Ethiopia are reviewed.

In Materials and Methods chapter, Description of the study area, hydrological model selection criteria, model inputs data collection and analysis, model setup and model performance evaluation are elaborated. The fourth chapter describes with the result and discussion which are satellite rainfall products versus rain gauges, stream flow modeling, satellite rainfall simulation of stream flow and overall discussion are made. The stream flow modeling includes sensitivity analysis, calibration and validation of stream flow simulation, and the performance evaluation of the model. Finally, in section five, conclusion and recommendations of the study are provided.

CHAPTER TWO

2. Literature Review

2.1. Hydrological cycle and rainfall – runoff processes in a watershed

Precipitation is a primary factor in hydrology and is the main input of water to land surface. Knowledge of the spatial and temporal distribution of rainfall in hydrological modeling is very important. Rainfall partly intercepted by the vegetation canopy before it reaches the ground surface. The quantity of rain captured by trees, plants, depends not only on their sort, stage of development, and density of vegetation canopy, but also on the intensity and duration of precipitation. For rainfall – runoff, interception is considered as a loss.

The remaining rain that reaches the ground either evaporates, infiltrates, or remains at the land surface as overland flow. Infiltrated rainfall may percolate to recharge the saturated domain, but also may flow to a channel through interflow. Due to evaporation and transpiration at the land surface the water availability in the subsurface reduces. On the other hand, due to an intensive rainfall event groundwater level may rise and water starts to move from saturated zone toward unsaturated zones as capillary rise. The dominant flow processes that contributes to catchment are difficult to understand because hydrological domains (overland, unsaturated, saturated and channel) are coupled. In reality several flow processes can be distinguished, Horton overland flow, this occurs when the intensity of the rainfall is greater than the infiltration capacity of the soil. Due to this water is stored at the land surface. According to (Dunne, 1983) Horton overland flow is common in arid and semiarid zones when rainfall events are very intensive and the vegetation cover is poor. When the soil is saturated due to rise of groundwater level to the land surface the saturation overland flow is generated. Saturated overland flow is typical for humid zones. Finally, channel flow occurs when the water reaches the catchment drainage system.

Flow processes do not occur only at the land surface, but also occurs in the subsurface. Indeed, due to infiltration water enters the subsurface as unsaturated flow in the form of matrix or macro pore flow. When the saturated hydraulic conductivity of a given layer is lower compared to the layer above, perched subsurface flow is generated. In the saturated zone groundwater flow can be rapid or delayed. Groundwater flow is very important component for the base flow.

2.2. Rainfall – runoff modeling

Understanding the rainfall – runoff relation has been a subject of hydrological research for a long time (Ragan 1968, Betson and Marius 1969). These studies have concluded that runoff occurs due to a complex interaction between surface, unsaturated, and saturated flow. In the same time researches in the field of hydrology focused on streamflow simulation processes using hydrological model. A model is a representation of the real world into a conceptual world. It is obvious that modeling the flow processes that occur in the real world will be simplified in modeling. However, these models should be able to capture the dominant processes at different space and time scale in a catchment. Models can be predictive (to obtain a specific answer to a specific question) or investigative (to further understanding hydrological processes) (O’Connell, 1991). Another classification of hydrological models is given in the report of World Meteorological Organization (WMO, 1990), which divides the models in two classes: deterministic model (models, which value of all parameters are uniquely defined) and stochastic models (models having probability distribution in parameter space). Within the deterministic models, three types of models are distinguished based on spatial discretization. In rainfall-runoff modeling, models are characterized as lumped, semi-distributed (White box) models account for spatial variability of topography, geology, soil type and landuse within a catchment. These models use grid layers with elements. Within this sub-catchment the characteristics are lumped. Also another classification is known: event models that represent a single runoff event occurring over a period of time from one hour to several days. In general event models tend to be lumped. Continuous models that determine flow during storm and inter-storm period. Considering the spatial scale, these models can range from lumped to distributed models. Based on the objectives of this study Soil and Water Assessment Tool (SWAT) model was used to simulate stream flow for both in situ dataset and satellite rainfall products.

2.3. Stream flow analysis

Watershed response to the event is characterized by measuring the discharge (volume rate of flow) at the outlet. This is described by a hydrograph. It is clear that hydrograph is a spatial and temporally integrated response of catchments which is determined by spatially and temporal variation of input (rainfall), and travel time of each drop of water from where it strikes on the stream network to outlet. After begins of the rainfall event the flow rate to basin increase rapidly and at same moment it reaches the peak discharge. Each hydrograph has a different shape because it depends on the catchment input (rainfall) and the catchment

characteristics. It is important to know climatic factors (Precipitation, evaporation, transpiration), geometric factors (size, shape, elevation and stream density), geology, soil type, and land use as well as channel factors (size, shape of the channel, cross-section, slope and length, roughness, and number of tributaries). Catchment response to an event is characterized by the discharges (volume rate of flow) at the outlet that is described by a hydrograph. Sometime after the start of the event the flow rate begins to increase till the well defined peak discharge. At this moment the hydrograph rises, which is characterized by the rising limb. After that the hydrograph declines as a result of ending of direct runoff, which is described by the recession limb. At the recession period, the discharge is related primarily to the base flow. From this description, one can identify the hydrograph components, which are overland flow, interflow, and base flow.

2.4. Description of satellite rainfall products

The different observational data that are used in this study are described in this section.

2.4.1. TMPA 3B42

TRMM Multisatellite Precipitation Analysis (TMPA) provides a calibration-based sequential scheme for combining precipitation estimates from multiple satellites as well as gauge analyses (where feasible) (Huffman et al., 2007).

The TMPA 3B42 algorithm covers the globe between the 50°S and 50°N latitudes and is published with a 0.25°×0.25° spatial resolution encompassing periods of three hours. With a data series from 1998 to present, the algorithm has undergone several modifications. In late 2012 TMPA 3B42 version 7a was released.

In the derivation of the research grade TMPA data, contributions from several passive MW sensors (TMI, SSM/I, AMSR-E and AMSU-B), rainfall estimates from IR sensors, combined estimates from TRMM's Microwave Imager (TMI) and PR (TRMM Combined Instrument, TCI), and interpolated ground rainfall measurements are used.

Towards the final product, the best MW estimates from the different sensors, calibrated based on TCI and TMI observations, are merged resulting in a high quality MW precipitation measurement. The next step is to calibrate IR brightness temperatures using the aforementioned high quality measurement. Calibrated MW and IR data are then merged,

being MW's taken "as is" where available, promoting local accuracy at the expense of a more homogeneous dataset. The final step is to scale the MW-IR estimates resorting to rain gauge data, which is done based on monthly aggregations (Huffman et al., 2007).

2.4.2. CMORPH

The Climate Prediction Center Morphing Method (CMORPH) presents yet another approach at rainfall estimation. Being a morphing technique, it does not simply "merge" data from several sources, but uses motion vectors from IR observations to propagate the relatively high quality precipitation derived from MW data (Joyce et al., 2004).

The basic idea underlying the morphing concept is that, as IR precipitation estimates are affected by a number of problems, using only estimates derived from passive MW sensors is advantageous. Unlike other algorithms, which employ IR precipitation in areas where no MW coverage is available, CMORPH uses IR indirectly, solely as a guide to propagate the MW features between valid observations.

The CMORPH dataset produces estimates every 30 minutes in the $\pm 60^\circ$ latitude band, with a spatial resolution of $0.0727^\circ \times 0.0727^\circ$. MW rainfall from the TMI, SSM/I, AMSU-B and AMSR-E sensors is used in its derivation.

The main steps of the algorithm are: passive MW rainfall ordering and mapping in a fine-resolution grid; surface snow and ice scanning, as it affects passive MW sensors; normalization of the AMSU-B estimates; derivation of half-hourly cloud system advection vectors from IR data and application of bias corrections; forward and backward propagation of MW rainfall estimates using the derived cloud system; and morphing the rainfall fields by a weighted combination of the forward and backward propagated fields.

2.4.3. PERSIANN

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) is designed to be capable of extracting and combining information from data of various types including, for example, IR and microwave satellite imagery, rain gauge and ground based radar data, and ground-surface topographic information (Hsu et al., 1997).

Various contributions describing PERSIANN enhancements have been published and retrieving exact details about the characteristics of the published dataset was not possible. In general, the PERSIANN algorithm starts by the application of a SOM (Self-Organizing Map) ANN in order to classify inputs relevant to rainfall estimation (such as topography, IR-derived and MW-derived precipitations in the vicinity of the location of interest, radar data, variability of these fields, etc.). The activated node of the SOM layer is then linearly transformed in order to produce estimates, being that the coefficients of this transformation are adaptively adjusted in order to calibrate the model (Hsu et al., 1997, Sorooshian et al., 2000).

2.5. Application of high – resolution satellite rainfall products in hydrological modeling

A number of studies have been carried out to assess the capability of satellite rainfall products through streamflow simulation using hydrological modeling. We summarized some of the studies that evaluated the performance of satellite rainfall products for hydrological simulation as follow: These studies revealed the capability of satellite rainfall products in the application of hydrological modeling.

Table 2.1 Studies on performance of satellite rainfall products through hydrological modeling.

Satellite rainfall products evaluated	Hydrological model used	Watershed characteristics	Main results	References
CMORPH, TMPA 3B42RT, TMPA 3B42 and PERSIANN)	SWAT (Arnold et al., 1998)	Gilgel Abay at 1656 km ² Koga at 299 km ² , Ethiopian highlands.	TMPA 3B42RT and CMORPH – based model simulations were performed better compared to TMPA 3B42 and PERSIANN. Also simulation based on	M. Bitew and M. Gebremichael, (2011).

			3B42RT and CMORHP improved when watershed area increases while 3B42 and PERSIANN based deteriorates.	
Satellite-based rainfall estimates (SBRE) (Xie and Arkin 1997)	Geospatial Stream Flow Model (GeoSFM) (Artan et al., 2001)	2646 km ² , Nyando Kenya 6044 km ² , Se Done, China 19013 km ² , nam Ou, China 22150 km ² , Gash, Ethiopia – Eritrea – Sudan	SBRE – based model simulation performed better when model was calibrated with satellite data than with sparse rain gauge data.	Artan et al., (2009)
TMPA 3B42 (Huffman et al., 2007)	SWAT (Arnold et al., 1998)	7720 km ² , middle Nueces, Texas 8950 km ² , midlle Rio Grande, Mexico – united States	While simulation with ground – based radar data were superior, TMPA 3B42 – based data were acceptable and comparable to those based on rain gauge data.	Tobin and Bennett (2009)
CMORPH, TRMM3B42 and PERSIANN)	Hillslope River Routing (HRR) model (Beighley et al., (2009)	3.7 Million km ² Congo Basin, the world’s second largest	The three satellite product vary significantly in	R.E.Beighley et al., (2011)

		river basin.	terms of magnitude but produces a reasonable streamflow hydrograph, overall this study revealed that, TRMM (3B42) performed best in spatial and temporal distributions and magnitudes while CMORPH and PERSIANN tend to overestimate.	
--	--	--------------	---	--

2.6. SWAT model

SWAT, developed by the United States Department of Agriculture (USDA) – Agricultural Research Service (ARS) in early 1990s (Arnold et al., 1998), is a continuous, semi-distributed hydrologic model that runs on a daily time step. Hydrologic response units (HRUs), defined by combinations of land cover and soil combinations, are the computational elements of SWAT. The daily water budget in each HRU is computed based on daily precipitation, runoff, evapotranspiration, percolation, and return flow from the subsurface and groundwater flow. Runoff volume in each HRU is computed using the Soil Conservation Service (SCS) curve number method (SCS, 1986).

The interface of SWAT model is compatible with ArcGIS that can integrate numerous available geospatial data to accurately represent the characteristics of the watershed. In SWAT model, the impacts of spatial heterogeneity in topography, land use, soil and other

watershed characteristics on hydrology are described in subdivisions. There are two scale levels of subdivisions; the first is that the watershed is divided into a number of sub-watersheds based upon drainage areas of the attributes, and the other one is that each sub-watershed is further divided into a number of Hydrologic Response Units (HRUs) based on land use and land cover, soil and slope characteristics.

The SWAT model simulates eight major components: hydrology, weather, sedimentation, soil temperature, crop growth, nutrients, pesticides, and agricultural management (Neitsch, et al., 2005). Major hydrologic processes that can be simulated by this model include evapotranspiration, surface runoff, infiltration, percolation, shallow aquifer and deep aquifer flow, and channel routing (Arnold et al., 1998). Stream flow is determined by its components (surface runoff and ground water flow from shallow aquifer).

2.6.1. SWAT model application worldwide

The SWAT model has good reputation for best use in agricultural watersheds and its uses have been successfully calibrated and validated in many areas of the USA and other continents (Ndomba, 2002; Tripathi et al., 2003). The studies indicated that the SWAT model is capable in simulating hydrological process and erosion/sediment yield from complex and data poor watersheds with reasonable model performance statistical values. Jayakrishnan et al. (2005) modeled the hydrology of the 3050 km² Sondu River basin in Kenya using land use, soil and elevation data with limited spatial resolution (1 - 10 km²). The objective was to assess impacts of land use changes as a result of changes to intensive dairy farming. The simulation Nash-Sutcliffe efficiency (NSE) coefficient of < 0.1 was attributed to inadequate rainfall and other model input data. The use of one rain gauge station situated at the upper end of the catchment was not representative of the basin.

In modeling the hydrology of the Mitano River basin in Uganda, Kingston and Taylor (2010) used the gridded 0.5° CRUTS3.0 database as the climatic input. Although there was a good agreement between observed and simulated monthly means and flow duration curves, the model performance after calibration was poor, resulting in an NSE of -0.09. According to the author, the poor performance in the hydrological modeling was attributed to “model-observation divergences with the calibration period that are simply too large to be resolved by an auto-calibration routine”.

SWAT has gained international acceptance as a robust interdisciplinary watershed modeling tool as evidenced by international SWAT conferences, hundreds of SWAT- related papers presented at numerous other scientific meetings, and large number of articles published in peer-reviewed journals (Gassman, 2007).

2.6.2. SWAT model application in Ethiopia

The SWAT model application was calibrated and validated in some parts of Ethiopia, frequently in Blue Nile basin. Setegn et al. (2009a) used SWAT to model the hydrological water balance of the Lake Tana basin in Ethiopia with the objective of testing the performance of the SWAT model for stream flow prediction. These authors calibrated and validated on four tributaries of Lake Tana using SUFI-2, GLUE and ParaSol algorithms. This paper reported that the SWAT model was more sensitive to HRU definition thresholds than to sub-basin discretization. Further, the paper reported that more than 60% of the observed river discharge falls within the 95% confidence bounds.

Gessese (2008) used the SWAT model performed to predict the Legedadi reservoir sedimentation. According to this study, the SWAT model performed well in Predicting sediment yield to the Legedadi reservoir. The study further put that the model proved to be worthwhile in capturing the process of stream flow and sediment transport of the watershed of the Legedadi reservoir. Mekonnen et al. (2009) developed a generic rainfall-runoff model better suited to Ethiopian catchments. They used a spectrum analysis method to extract the relationships between different temporal scales of available daily rainfall and runoff series that reflect the temporal and spatial scales of 25 discharges in two watersheds in Ethiopia. The paper reported that frequencies in rainfall and stream discharge longer than 50 days had a sufficient coherence to warrant model calibration.

In addition to the above, Tibebe and Bewket (2011) assessed surface runoff generation and soil erosion rates for a small watershed in the Awash River basin of Ethiopia using the SWAT model. Comparing monthly predicted runoff against the measured values, the study demonstrates that distribution of observed and simulated runoff was quite uniform throughout the simulation period. The study presents a high correlation value of 0.831. It further reports a NSE of 0.789 to demonstrate that the model was able to generate monthly runoff close to the observed.

The literature reviewed and presented above showed that SWAT is capable of simulating hydrological and soil erosion process with reasonable accuracy and can be applied to large and complex watersheds.

CHAPTER THREE

3. Materials and Methods

3.1. Description of the study area

3.1.1. Location

The Angar watershed is one of the tributary river basins of upper Blue Nile Basin. It is located in south parts of the Blue Nile basin within $9.0^{\circ} - 10.0^{\circ}\text{N}$ and $36^{\circ}30' - 37^{\circ}10'\text{E}$. The topography of the watershed is mainly mountainous or hilly with elevation ranging from 1286 m to 3057 m.a.s.l. The total area of the watershed is estimated to be 3627.6 Km^2 .

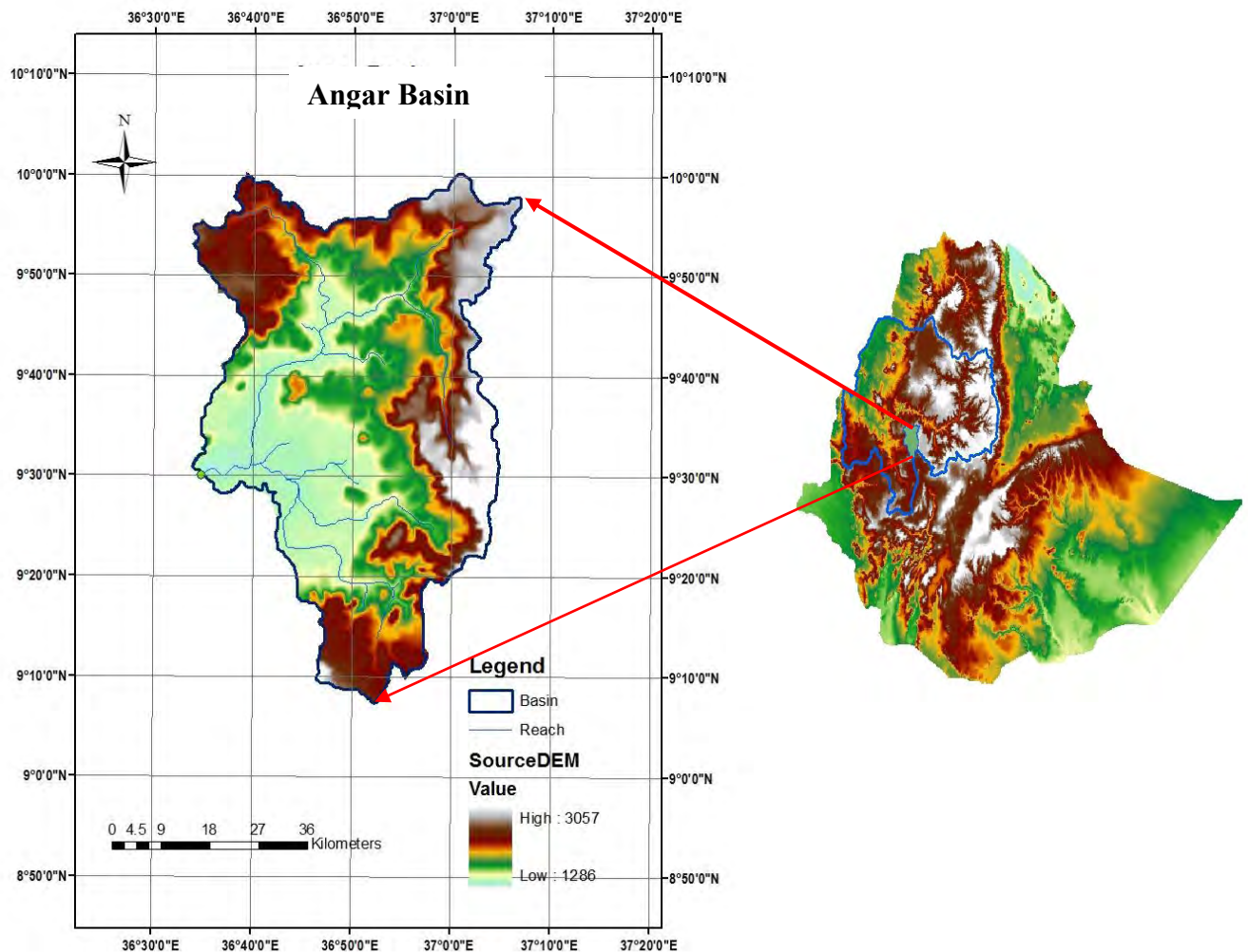


Figure 3.1 Location map of the area (Source: Own analysis)

3.1.2. Climate

The climate of Ethiopia is mainly controlled by the seasonal migration of the Inter-tropical convergence Zone (ITCZ) and associated atmospheric circulation as well as by the complex topography of the country. It has a diversified climate ranging from semi-arid desert type in the low lands to humid and temperate type in the southwest.

There are three recognized seasons in upper Blue Nile basin. The main rainy season (Kiremt) lasts generally from June – September during which south-west winds bring rains from the Atlantic Ocean. About 70-90 per cent of total rainfall occurs during this season which is also typified by minimum levels of sunshine, low variation in daily temperatures and high relative humidity. A dry season (Bega) lasts from October to January during which clear skies are associated with maximum sunshine, high daily temperature variation, and low relative humidity. Finally, the minor rainy season (Belg) lasts from mid-February to mid-May during which south-east winds bring the small rains from the Indian Ocean and temperatures are at their highest; however, this is only significant in the extreme east of the basin. The timing and duration of these seasons varies considerably by location. Spatial characteristics combine with these seasonal characteristics to define the pattern of the climate in different areas.

The climate of the Angar basin is dominated by two factors: its near-equatorial location and an altitude ranging from 1286 meters above sea level to 3057 meters. There is high spatial and temporal variation of rainfall in the study area. The main rainfall season which accounts around 70-90% of the annual rainfall occurs from June to September, while small rains also occur during December to March.

The monthly rainfall distributions of the study area indicate that July and August are the wettest months of the year in all the selected stations. The mean monthly rainfall of Nekemte, which is first class stations for the period of 1992 – 2012, is shown in Figure 3.2. The mean annual rainfall recorded at different stations is 1729 mm and annual average daily minimum and maximum temperature ranges between 12⁰c and 25⁰c.

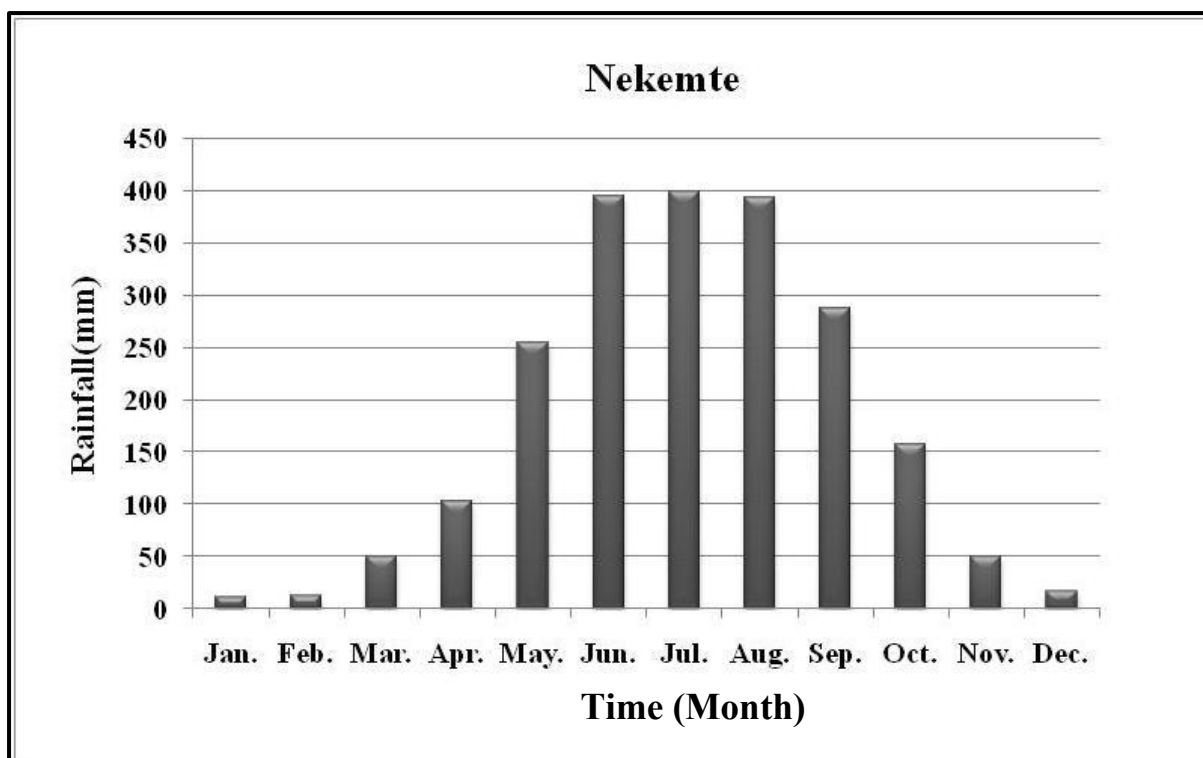


Figure 3.2 Mean monthly rainfall distribution of principal meteorological station for the period of 1992 - 2012

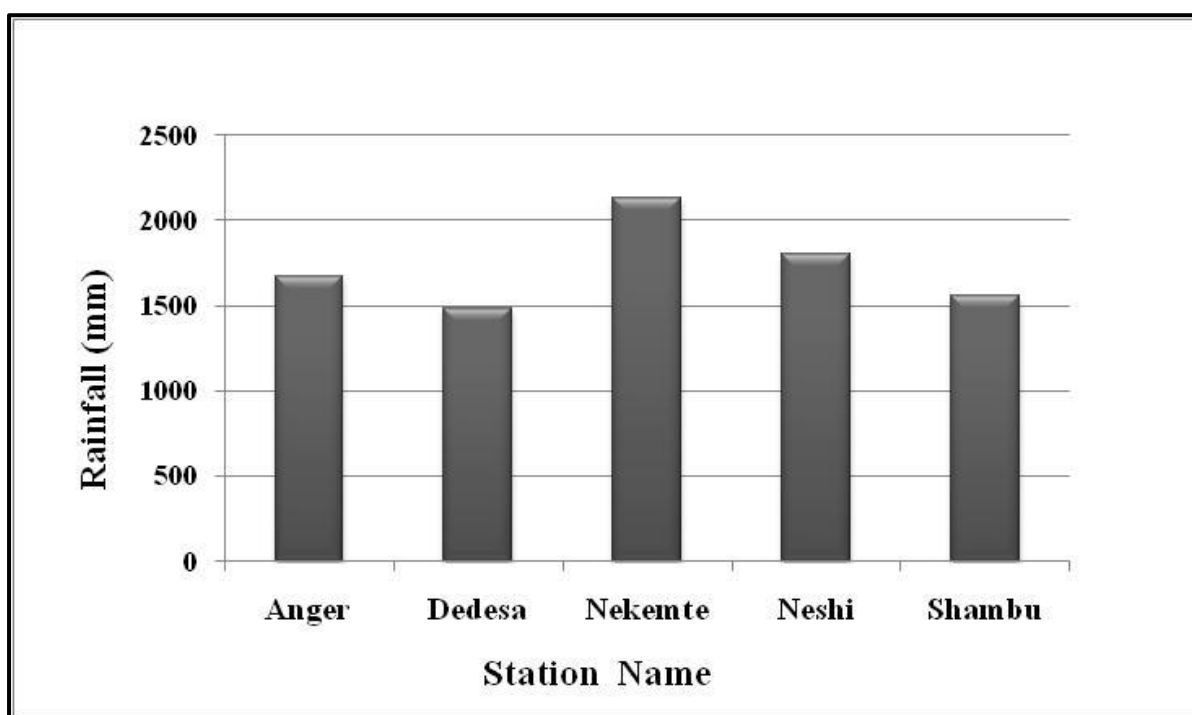


Figure 3.3 Mean annual rainfall from 1992- 2012

3.1.3 Topography and geology

Angar watershed is dominated by high elevation, which ranges from 1286 – 3057 meter above sea level and with mean elevation of 1774 m.a.s.l. The upstream part is characterized by a mountain area with very steep slope while downstream part is flat area.

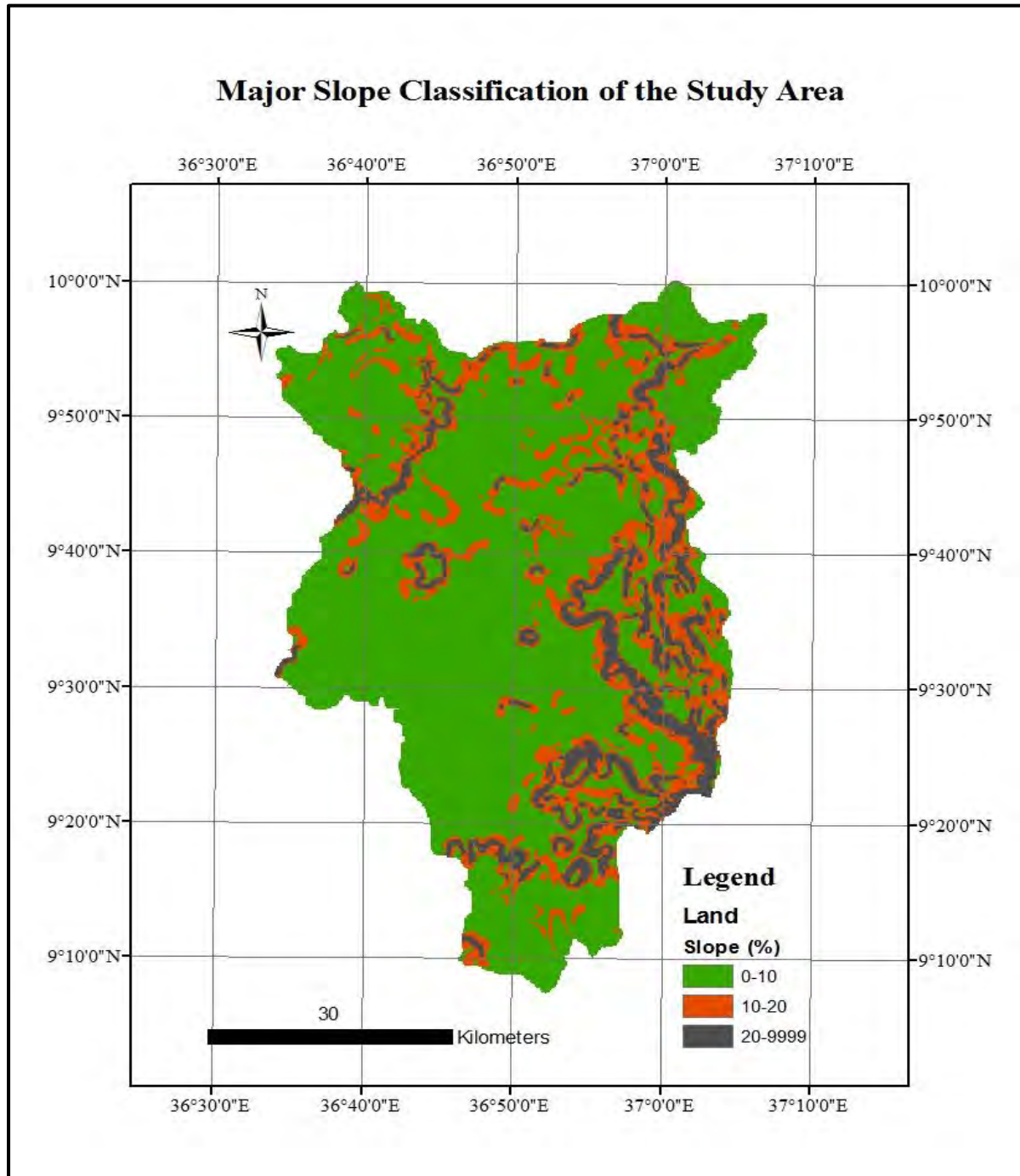


Figure 3.4 Slope classification of the watershed

3.1.4. Soil types and land uses

The Angar watershed is mainly formed from Clay and Clay-loam soil type, but the riverbed has a loam and sandy-loam type of soil. The infiltration capacity of the soil depends, among others, on the porosity of the soil, which determines its storage capacity and affects the resistance of the water to flow into deep layers. Since the soil infiltration capacity depends on the soil texture, the highest infiltration rates are observed in sandy soil. This shows that, surface runoff is higher in heavy clay and loamy which has low infiltration rate. Soils that have predominantly available in the study area are generally Nitisols, Alisols, Acrisols, Leptosols and Vertisols.

Vertisols are deep to very deep, moderately well to poorly drained, very dark grey to dark yellowish brown in the topsoil, and clay textured throughout. The soils have large surface cracks in the dry season. Run-off formation from Vertisols is high and hence it is susceptible to erosion. Vertisols are found in relatively small areas in the watershed area. The texture of Leptosols varies from sandy clay loam to clay and has excessive drainage characteristics.

Nitisols occupied about 29% of the watershed area. They are reddish brown to red clay soils. These soils are deep and have very good potential agriculture. The Nitisols of the area are intensely cultivated for annual crops. The major red clay soils (Alisols) occur mainly on flat to rolling upland plain and flat to undulating land features. These are deep, well drained, permeable and medium textured soils.

The landuse is dominated mainly by forest, range-grasses and in some part, by the mixture of both agriculture and pastoral.

3.2. Hydrological model selection criteria

There are various criteria which can be used for choosing the right hydrological model for a specific problem. These criteria are always project dependent, since every project has its own specific requirements and needs. Further, some criteria are also user-dependent (and therefore subjective). Among the various project-dependant selection criteria, there are four common, fundamental ones that must be always answered (Cunderlik, 2003):

- Required model outputs important to the project and therefore to be estimated by the model (Does the model predict the variables required by the project such as long-term sequence of flow?),
- Hydrologic processes that need to be modeled to estimate the desired outputs adequately (Is the model capable of simulating single-event or continuous processes?),
- Availability of input data (Can all the inputs required by the model be provided within the time and cost constraints of the project?),
- Price (Does the investment appear to be worthwhile for the objectives of the project?).

Reasons for selecting SWAT model

The reasons behind for selecting SWAT model for this study are;

- The model simulates the major hydrological process in the watersheds
- It is less demanding on input data, and
- It is readily and freely available.

A major limitation to large area hydrologic modeling of SWAT is the spatial detail required to correctly simulate environmental processes. For example, it is difficult to capture the spatial variability associated with precipitation within a watershed. Another limitation is data files can be difficult to manipulate and can contain several missing records. The model simulations can only be as accurate as the input data. The third limitation is that, the SWAT model does not simulate detailed event-based flood.

3.2.1 Description of SWAT model

Soil and Water Assessment Tool (SWAT) was applied in Angar Watershed to assess the streamflow simulation capability of high resolution satellite rainfall products. The SWAT

model is embodied in ArcGIS that can integrate various readily available geospatial data to accurately represent the characteristics of the watershed.

The SWAT watershed model is one of the most recent models developed by the USDA-ARS to predict the impacts of land management practices on water, sediment and agricultural chemicals yields in watersheds with varying soils, land use and management practices over long periods of time (Neitsh, et al., 2005).

The model is a physical based, semi-distributed, continuous time, and operating on daily time step (Neitsch, et al., 2005). As a physical based model, SWAT uses Hydrological Response Units (HRUs) to describe spatial heterogeneity in terms of land use, soil types and slope with in a watershed. In order to simulate hydrological processes in a watershed, SWAT divides the watershed in to sub watersheds based upon drainage areas of the tributaries. The sub watersheds are further divided in to smaller spatial modeling units known as HRUs, depending on land use and land cover, soil and slope characteristics.

One of the main advantages of SWAT is that it can be used to model watersheds with less monitoring data. For simulation, SWAT needs digital elevation model; land use and land cover map, soil data and climate data of the study area. These data are used as an input for the analysis of hydrological simulation of surface runoff and groundwater recharge.

SWAT splits hydrological simulations of a watershed in to two major phases: the land phase and the routing phase. The land phase of the hydrological cycle controls the amount of water, sediment, nutrient, and pesticide loadings to the main channel in each sub watershed. While the routing phase considers the movement of water, sediment and agricultural chemicals through the channel network to the watershed outlet.

The land phase of the hydrologic cycle is modeled in SWAT based on the water balance equation (Neitsch, et al., 2005):

$$SW_t = SW_o + \sum_{i=1}^t (R_{day} - Q_{surf} - E_a - W_{seep} - Q_{gw}) \quad (3.1)$$

Where, SW_t is the final soil water content (mm)

SW_o is the initial water content (mm)

t is the time (days)

R_{day} is the amount of precipitation on day i (mm)

Q_{surf} is the amount of surface runoff on day i (mm)

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

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

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

The model has eight major components: hydrology, weather, sedimentation, soil, temperature, crop growth, nutrients, pesticides, and agricultural management (Neitsch, et al., 2005). However, brief description of some of the SWAT computation procedures which are considered in this study are presented under the following subsections. For complete model description, one may refer to SWAT theoretical documentation (Neitsch, et al., 2005).

3.2.1.1. Surface runoff

Surface runoff refers to the portion of rain water that is not lost to interception, infiltration, and evapotranspiration (Solomon, 2005). Surface runoff occurs whenever the rate of precipitation exceeds the rate of infiltration. SWAT offers two methods for estimating the surface runoff: the Soil Conservation Service (SCS) curve number method (USDA-SCS, 1972) or the Green and Ampt infiltration method (Green and Ampt, 1911). The Green and Ampt method needs sub-daily time step rainfall which made it difficult to be used for this study due to unavailability of sub-daily rainfall data. Therefore, the SCS curve number method was adopted for this study.

The general equation for the SCS curve number method is expressed by equation

$$Q_{surf} = \frac{(R_{day} - I_a)^2}{(R_{day} - I_a + S)} \quad (3.2)$$

Where, Q_{surf} is the accumulated runoff or rainfall excess (mm),

R_{day} is the rainfall depth for the day (mm water),

I_a is initial abstraction which includes surface storage, interception and infiltration prior to runoff (mm water),

S is retention parameter (mm water).

The retention parameter varies spatially due to changes with land surface features such as soils, land use, slope and management practices. This parameter can also be affected temporally due to changes in soil water content. It is mathematically expressed as:

$$S = 25.4 * \left(\frac{1000}{CN} - 10 \right) \quad (3.3)$$

Where, CN is the curve number for the day and its value is the function of land use practice, soil permeability and soil hydrologic group.

The initial abstraction, I_a , is commonly approximated as $0.2S$ and equation (3.2) becomes

$$Q_{surf} = \frac{(R_{day} - 0.2S)^2}{(R_{day} + 0.8S)} \quad (3.4)$$

For the definition of hydrological groups, the model uses the U.S. Natural Resource Conservation Service (NRCS) classification. The classification defines a hydrological group as a group of soils having similar runoff potential under similar storm and land cover conditions. Thus, soils are classified in to four hydrologic groups (A, B, C, and D) based on infiltration which represent high, moderate, slow, and very slow infiltration rates, respectively.

3.2.1.2. Potential evapotranspiration

Potential Evapotranspiration is a collective term that includes evaporation from the plant (transpiration) and evaporation from the water bodies and soil. Evaporation is the primary mechanism by which water is removed from a watershed. An accurate estimation of evapotranspiration is critical in the assessment of water resources and the impact of land use change on these resources.

There are many methods that are developed to estimate potential evapotranspiration (PET). SWAT Provides three options for PET calculation: Penman-Monteith (Monteith, 1965), Priestley-Taylor (Priestley and Taylor, 1972), and Hargreaves (Hargreaves et al., 1985) methods. The methods have various data needs of climate variables. Penman-Monteith method requires solar radiation, air temperature, relative humidity and wind speed; Priestley-

Taylor method requires solar radiation, air temperature and relative humidity; whereas Hargreaves method requires air temperature only and therefore, selected for this study.

3.2.1.3. Ground water flow

To simulate the ground water, SWAT partitions groundwater into two aquifer systems: a shallow, unconfined aquifer which contributes return flow to streams within the watershed and a deep, confined aquifer which contributes return flow to streams outside the watershed (Arnold et al., 1993). In SWAT the water balance for a shallow aquifer is calculated with equation (3.5)

$$aq_{sh,i} = aq_{sh,i-1} + W_{rchrg} - Q_{gw} - W_{revap} - W_{deep} - W_{pump,sh} \quad (3.5)$$

Where, $aq_{sh,i}$ is the amount of water stored in the shallow aquifer on day i(mm),

$aq_{sh, i-1}$ is the amount of water stored in the shallow aquifer on day i-1(mm)

W_{rchrg} is the amount of recharge entering the aquifer on day i (mm)

Q_{gw} is the ground water flow, or base flow, or return flow, into the main channel on day i (mm)

W_{revap} is the amount of water moving in to the soil zone in response to water deficiencies on day i (mm)

W_{deep} is the amount of water percolating from the shallow aquifer in to the deep aquifer on day i (mm), and

$W_{pump,sh}$ is the amount of water removed from the shallow aquifer by pumping on day I (mm)

3.2.1.4. Flow routing phase

The second component of the simulation of the hydrology of a watershed is the routing phase of the hydrologic cycle. It consists of the movement of water, sediment and other constituents (e.g nutrients pesticides) in the stream network. The change in channel dimensions with time due to down cutting and widening is also included. Similar to the case for the overland flow, the rate and velocity of flow is calculated by using the manning's equation. The main channels or reaches are assumed to have a trapezoidal shape by the model.

Two options are available to route the flow in the channel network: the variable storage and Muskingum methods. The variable storage method uses a simple continuity equation in routing the storage volume, whereas the Muskingum routing method models the storage volume in a channel length as a combination of wedge and prism storages. In the latter

method, when a flood wave advances into a reach segment, inflow exceeds outflow and a wedge of storage is produced. As the flood wave recedes or retreat, outflow exceeds inflow in the reach segment and a negative wedge is produced. In addition to the wedge storage, the reach segment contains a prism of storage formed by a volume of constant cross-section along the reach length. While calculating the water balance in the channel flow, the transmission and evaporation are also well considered by the model.

The variable storage method was used for this study. The method was developed by (Williams, 1969). The equation of the variable storage routing is given by:

$$\Delta V_{Stored} = V_{in} - V_{out} \quad (3.6)$$

Where, $\Delta V - \text{stored}$ is the change in volume of storage during the time step ($\text{m}^3 \text{ water?}$)

V_{in} – is the volume of inflow during the time step ($\text{m}^3 \text{ water}$), and

V_{out} – is the volume of outflow during the time step ($\text{m}^3 \text{ water}$).

The above equation can also be rewritten in detail as follows:

$$V_{stored,2} - V_{stored,1} = \Delta t \left(\frac{q_{in,1} + q_{in,2}}{2} \right) - \Delta t \left(\frac{q_{out,1} + q_{out,2}}{2} \right) \quad (3.7)$$

Where: Δt is the length of the time step (s), $q_{in,1}$ is the inflow rate at the beginning of the time step (m^3/s), $q_{in,2}$ is the inflow rate at the end of the time step (m^3/s), $q_{out,1}$ is the outflow rate at the beginning of the time step (m^3/s), $q_{out,2}$ is the outflow rate at the end of the time step (m^3/s), $V_{stored,1}$ is the storage volume at the beginning of the time step (m^3) and $V_{stored,2}$ is the storage volume at the end of the time step (m^3)

3.3. Methodology

The general methodology for the study can be described by the following flow chart;

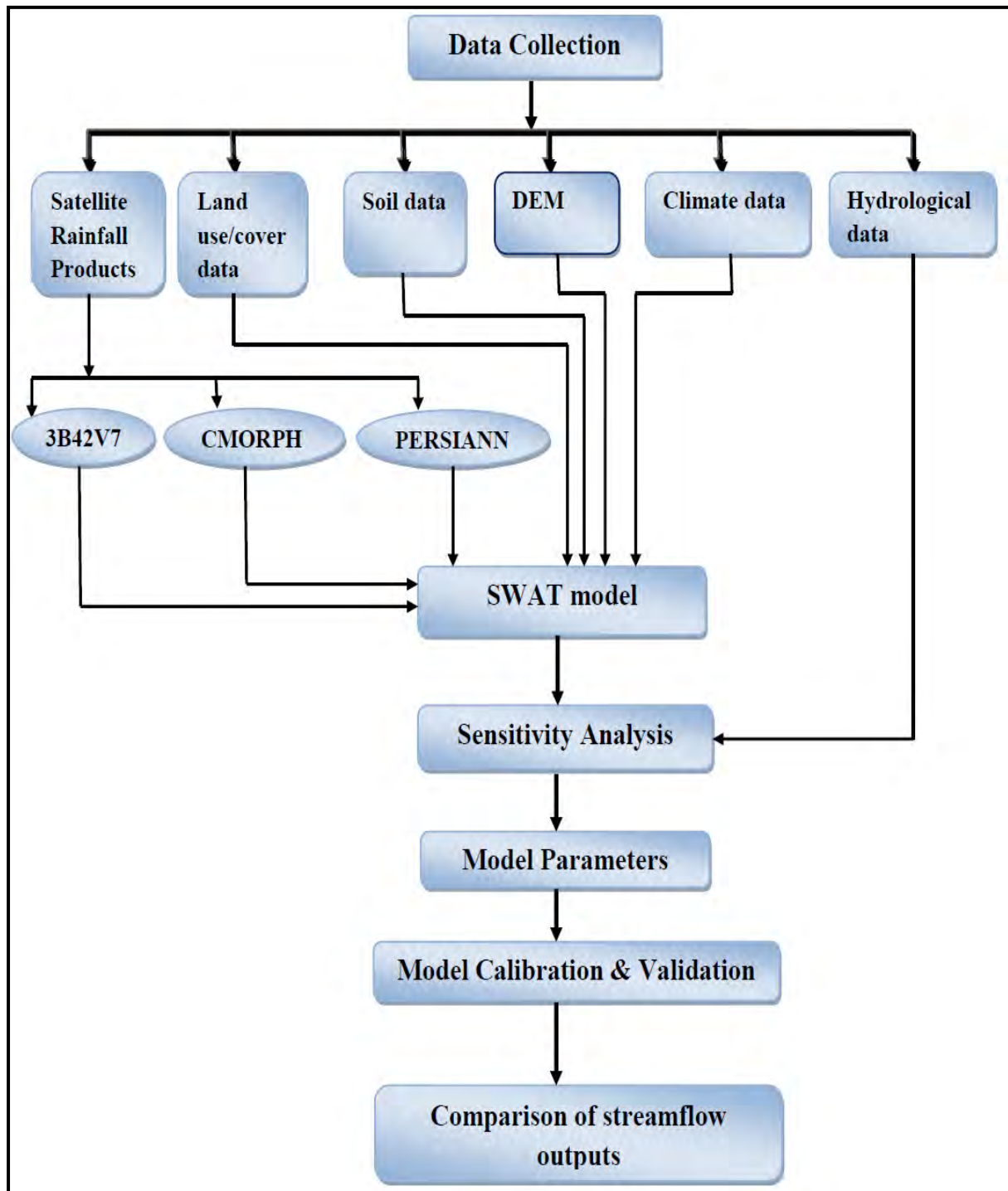


Figure 3.5 Frame work of the study

3.4. Model inputs data collection and analysis

The necessary data that was collected and used for this study can be classified into spatial and time series data. Spatial data used are DEM, land use/cover and soil map of the study area and collected from Ministry of Water, Irrigation and Energy (MoWIE). The time series data used are Meteorological and hydrological data and these data were collected from Ethiopian National Meteorological Agency and MoWIE respectively.

Models and software's used for simulation of stream flow in the study area was Arc GIS 9.3 which used for input preparation of SWAT model, to extend the Arc SWAT model and to prepare the Thiessen polygon of the Meteorological stations in the watershed and extension of SWAT model that is Arc SWAT 2009.

3.4.1. Digital elevation model

The digital elevation model (DEM) is any digital representation of a topographic surface and it is specifically made available in the form of raster or regular grid of spot heights. Digital Elevation Model (DEM) data is required to calculate the flow accumulation, stream networks, and watershed delineation using SWAT watershed delineator tools. The DEM obtained for this study was obtained from Ministry of Water, Irrigation and Energy and it has a resolution of 90m x 90m. Figure 3.6 shows elevation distribution of the Angar Watershed.

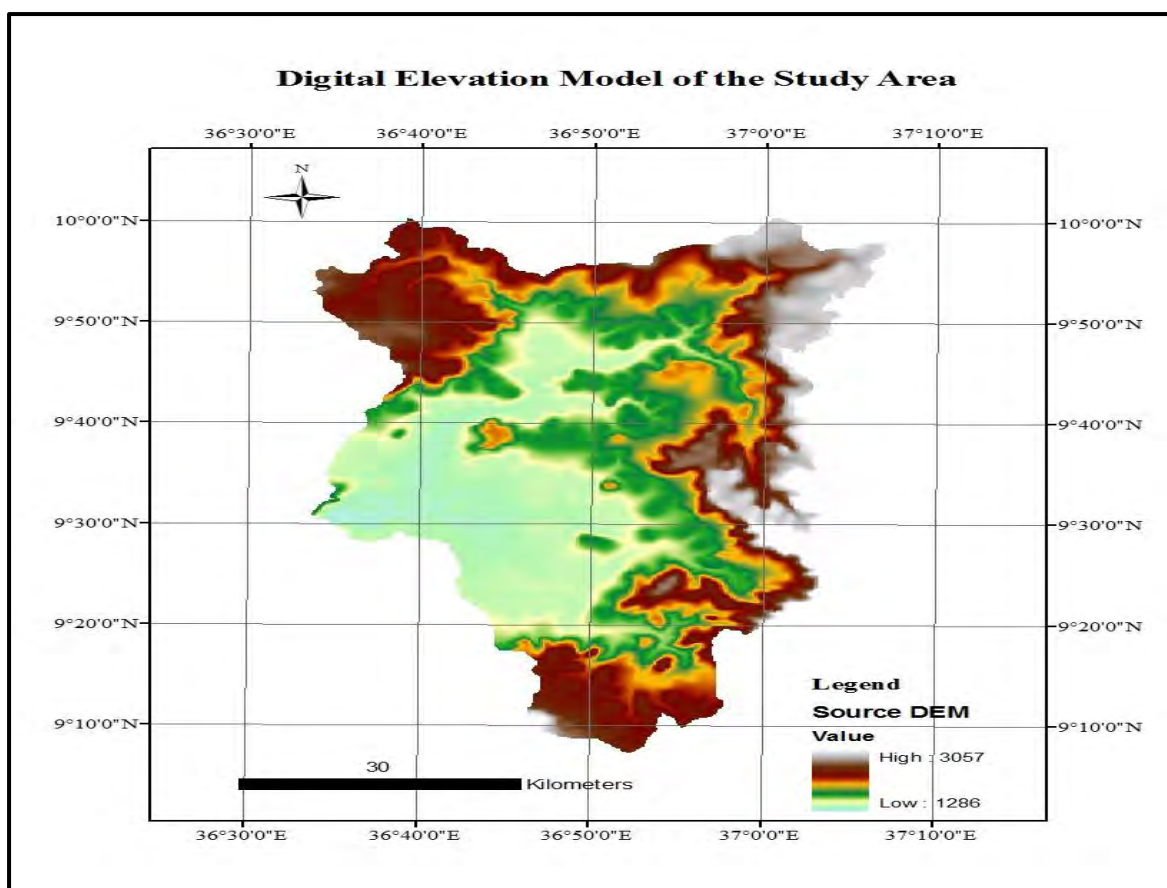


Figure 3.6 Digital elevation model of Angar watershed

3.4.2. Weather data

Weather data are among the main demanding input data for the SWAT simulation. The weather input data required for SWAT simulation includes daily data of Precipitation, maximum and minimum temperature, relative humidity, wind speed and solar radiation. These were obtained from the Ethiopian National Meteorological Agency. The weather data used were represented from five stations around Angar watershed, such as Anger, Dedessa, Neshi, Nekemte and Shambu. Only Nekemte station is the first class that contains records on all climatic variables, whereas the other stations are second class stations. The climatic data used for this study covers 21 years from January 1992 to December 2012. However, missing values were identified in some of the climatic variables. These values were assigned with no data code (-99) which then filled by the weather generator embodied in the SWAT model.

Finally, the weather data were prepared in DBF format with lookup tables as required by the model.

Table 3.1 Meteorological station names, locations and variables

No	Station Name	Latitude (deg)	Longitude (deg)	Rainfall	Max Temp	Min Temp	RH	WN	SHR
					p	p	D	D	
1	Nekemte	9.08	36.46	X	X	X	X	X	X
2	Angar	9.27	36.33	X	X	X	X	X	
3	Shambu	9.57	37.12	X	X	X	X		X
4	Dedessa	9.38	36.1	X	X	X	X		
5	Neshi	9.72	37.27	X	X	X			

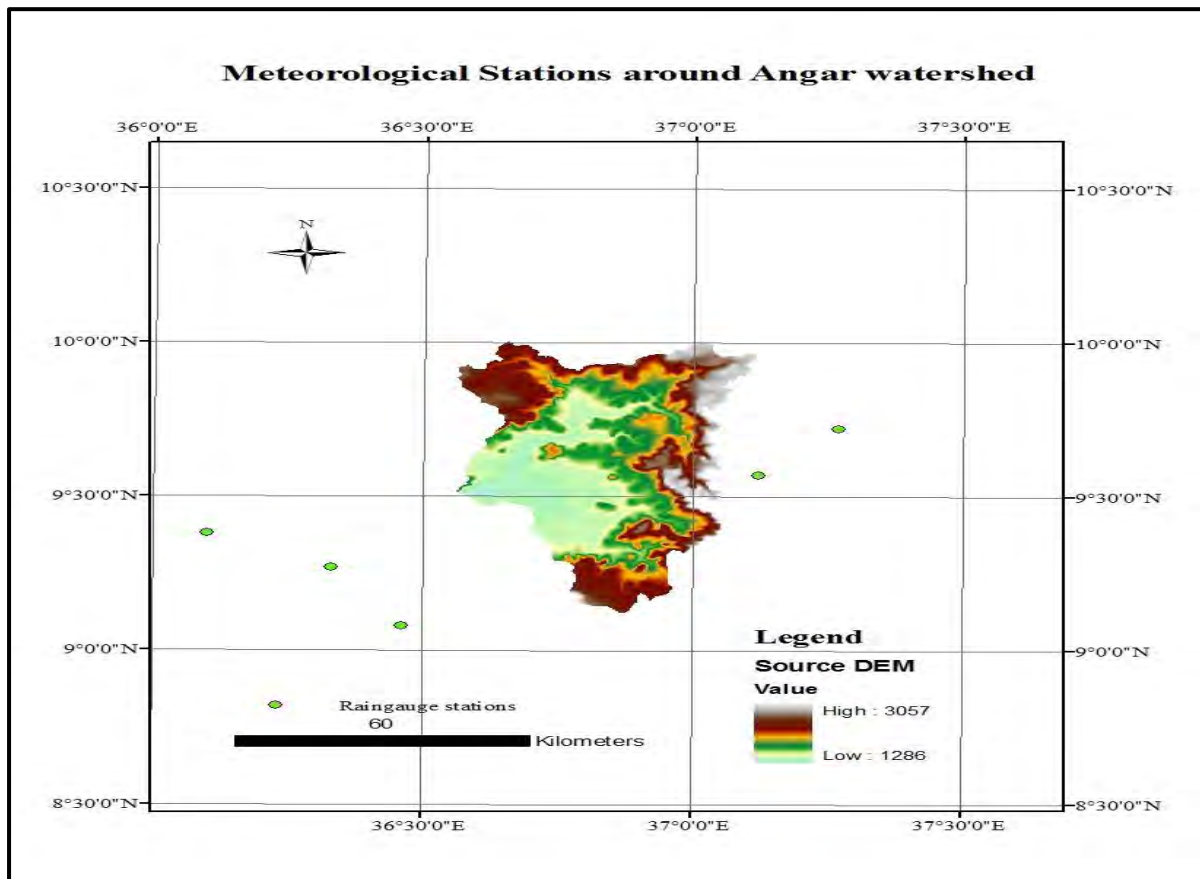


Figure 3.7 Meteorological stations around Angar watershed

Checking homogeneity of selected rainfall station

One of the methods to check homogeneity of the selected stations is the non-dimensional rainfall records and plotted to compare the stations with each other. Non-dimensional values of the monthly precipitation of each station can be computed by:

$$P_i = \frac{P_{i,av}}{P_{av}} * 100 \quad (3.8)$$

Where P_i is non-dimensional value of precipitation for the month in station i . $P_{i,av}$ over years averaged monthly precipitation for the station i and P_{av} is over year's averaged yearly precipitation of the station i .

The stations that used for this study have a rainfall pattern of uni-modal with high rainfall season in June to September.

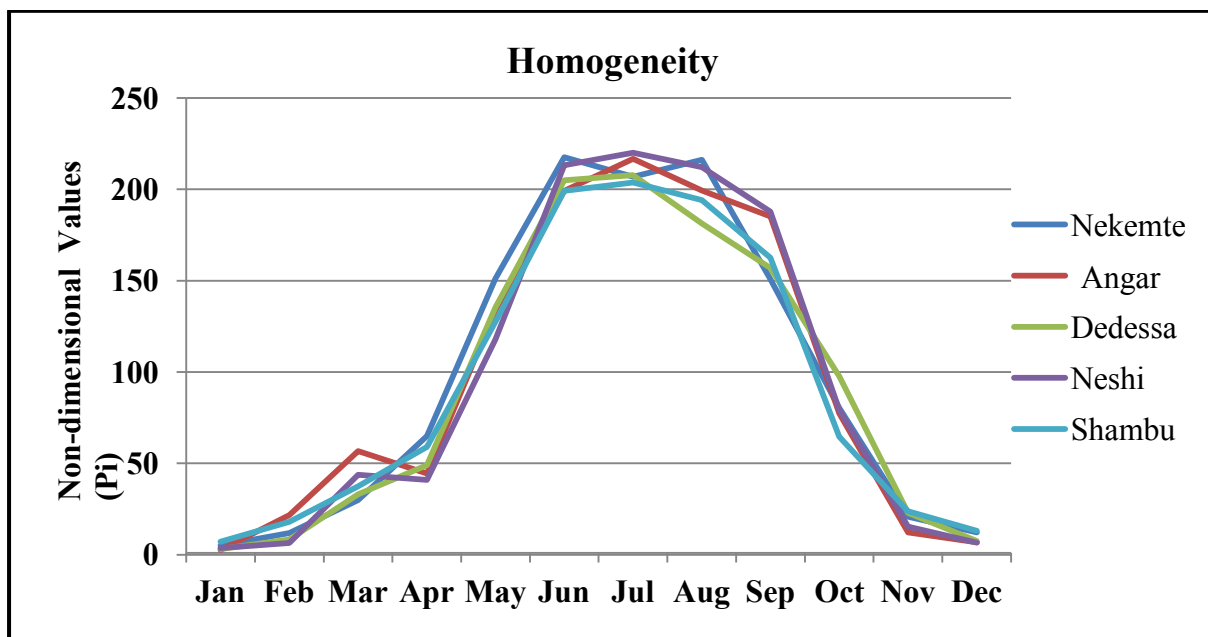


Figure 3.8 Non-dimensional plot of selected stations around Angar watershed

Checking consistency and adjustment of rainfall stations

A consistent record is the one where the characteristics of the record have not changed with time. Adjusting for gauge consistency involves the estimation of an effect rather than a missing value. The consistency of rainfall records on selected stations commonly checked by double mass curve analysis. Double mass curve is a graphical method for identifying and adjusting inconsistency in a station record by comparing its time trend with those of adjacent

stations. If the conditions relevant to the recording of a rain gauge station have undergone a significant change during the period of record, inconsistency would arise in the rainfall data of that station. This inconsistency can be differentiated from the time the significant change took place. If significant change in the regime of the curve is observed, it should be corrected by using equation (3.9). The stations used in this study have not undergone a significant change during the base line period (1992 to 2012) of the study.

$$P_{cx} = P_x * \frac{M_c}{M_a} \quad (3.9)$$

Where: P_{cx} is corrected precipitation at any time period, P_x is original recorded precipitation at time period, M_c is corrected slope of the double mass curve and M_a is original slope of the double mass curve.

Filling missing rainfall data

Measured precipitation data are important to many problems in hydrologic analysis and design. But, due to failure of the observer to make the necessary visit to the gauge, vandalism of recording gauges or instrument failure (by mechanical or electrical malfunctioning) may result in missing data. It is very essential to fill in the missing data when a hydrological basin analysis needs a continuous data. There are a number of deterministic and stochastic weighting methods used for infilling missing rainfall data values, some of them are: weighted average method, Isohyetal method, station year method and weighted distance method. There are several other types of methods, including inverse distance, simple arithmetic averaging and normal ratio methods. These are mainly related to spatial interpolations techniques. The rest are model based methods used to fill missing data. WXGEN is a stochastic model (Sharpley and Williams 1990). It was developed for the contiguous US, is used in SWAT to fill missing data using monthly statistics which is calculated from existing daily data. For this study missing values was estimated from other stations around the missed record station by considering the assumptions of at least three as close to and evenly spaced around the station with the missing record station as possible. Simple Arithmetic mean method was used where the mean monthly rainfall of all the index stations is within 10% of the station under consideration (station x) and calculated the missing data by equation (3.10) below. Whereas the mean monthly rainfall of one or more of the adjacent (index) stations differs from that of station x by more than 10% then the normal ratio method was used (equation 3.11). The data

obtained from five weather stations namely Anger, Didessa, Shambu, Neshi and Nekemt were used to perform the interpolation for filling the gaps of the missing data.

$$P_x = \frac{1}{N} (P_A + P_B + P_C \dots + P_N) \quad (3.10)$$

$$P_x = \frac{1}{N} \left(\frac{N_x}{N_A} P_A + \frac{N_x}{N_B} P_B + \frac{N_x}{N_C} P_C \dots + \frac{N_x}{N_N} P_N \right) \quad (3.11)$$

Where P_x is the precipitation for the station with missed record, $P_A, P_B, P_C \dots P_N$ are the corresponding precipitation at the index stations and $N_A, N_B, N_C \dots N_N$ and N_x are the long term mean monthly precipitation at the index stations and at station X under consideration respectively.

3.4.3. Satellite rainfall products

In this study three high resolution satellite-based rainfall estimations (SREs) were evaluated, which estimates precipitation at the same spatial (0.25° by 0.25°) and temporal (3-hour) resolution. The SREs considered are (i) the Tropical Rainfall Measuring Mission (TRMM) the TMPA method post-real-time research version 7 product (3B42V7), (ii) The Climate Prediction Center Morphing Method (CMORPH) and The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN). Figure 3.9 shows ten satellite rainfall grids in Angar watershed.

3.4.4. Stream flow data

In rainfall – runoff modeling, the stream flow data (available throughout the entire year) was used for calibration of the results obtained from model simulation. The daily stream flow data (2001 - 2011) is relatively long period and were collected from the Ministry of Water, Irrigation and Energy of Ethiopia for the Angar Watershed.

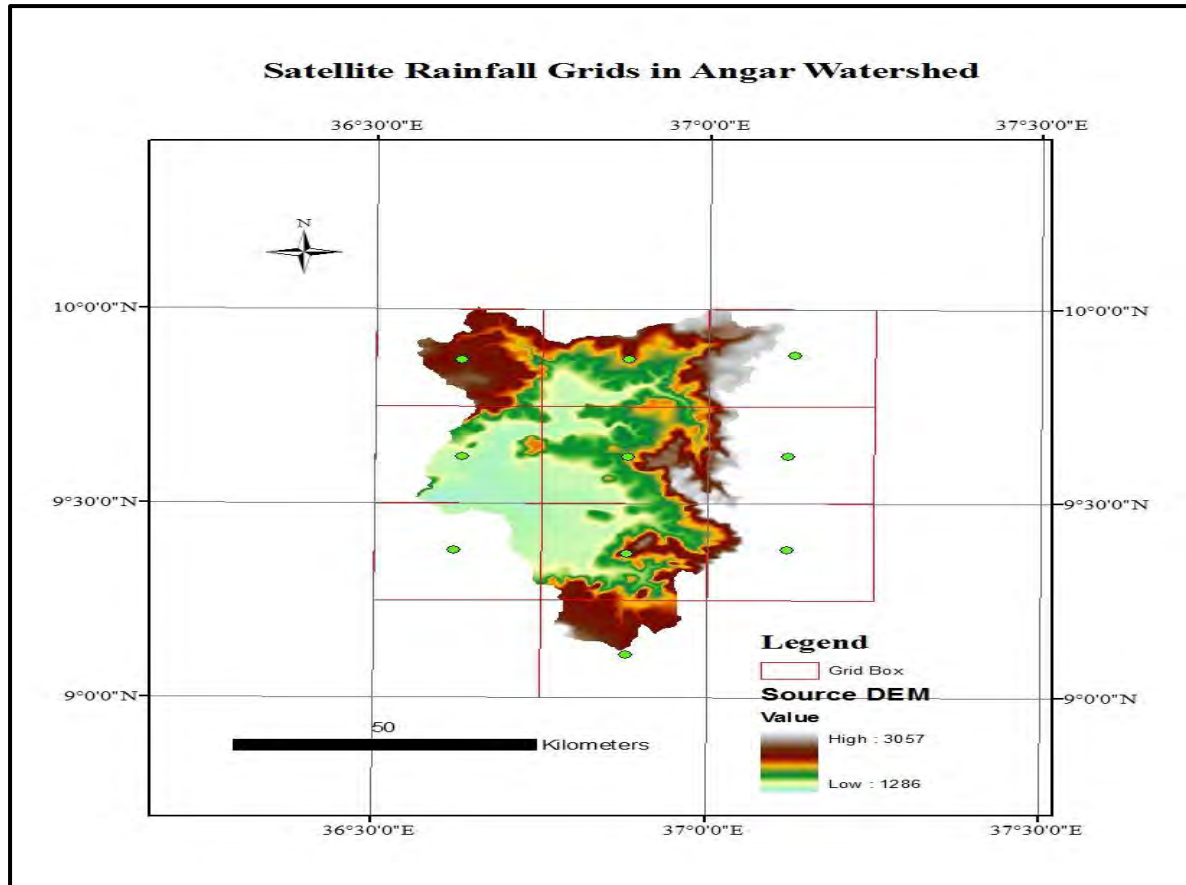


Figure 3.9 Angar watershed with ten satellite rainfall grids ($0.25^{\circ} \times 0.25^{\circ}$)

3.4.5. Soil data

Soil data is one of the major input data for the SWAT model. The soil textural and physiochemical properties required by the SWAT model include Soil texture, available water content, hydraulic conductivity, bulk density and organic carbon content for each soil type. These data were obtained from FAO (1998, 2002, and 2005) and the Ministry of Water, Irrigation and Energy of Ethiopia (MoWIE).

The shape file which describes the distribution of soil in the study area was obtained from the base line maps available at MoWIE. It was observed that Rhodic Nitisols, Haplic Nitisols, Haplic Alisols, Haplic Acrisols, Eutric Vertisols, Eutric Leptosols and Dystric Leptosols are the most dominant soils in the watershed.

To integrate the soil map with SWAT model, a user soil database which contains textural and chemical properties of soils was prepared for each soil layers and added to the SWAT user soil databases using the data management append tool in ArcGIS. The area coverage of the soil types are presented in Table 3.2

Table 3.2 Major soil distribution within Angar watershed

Land use	Area (ha)	Percentage of the area covered
1 Haplic Alisols	137544	37.92
2 Rhodic Nitisols	105684	29.13
3 Dystric Leptosols	74124	20.43
4 Haplic Acrisols	37344	10.29
5 Eutric Leptosols	4596	1.27
6 Haplic Nitisols	2568	0.71
7 Eutric Vertisols	896	0.25

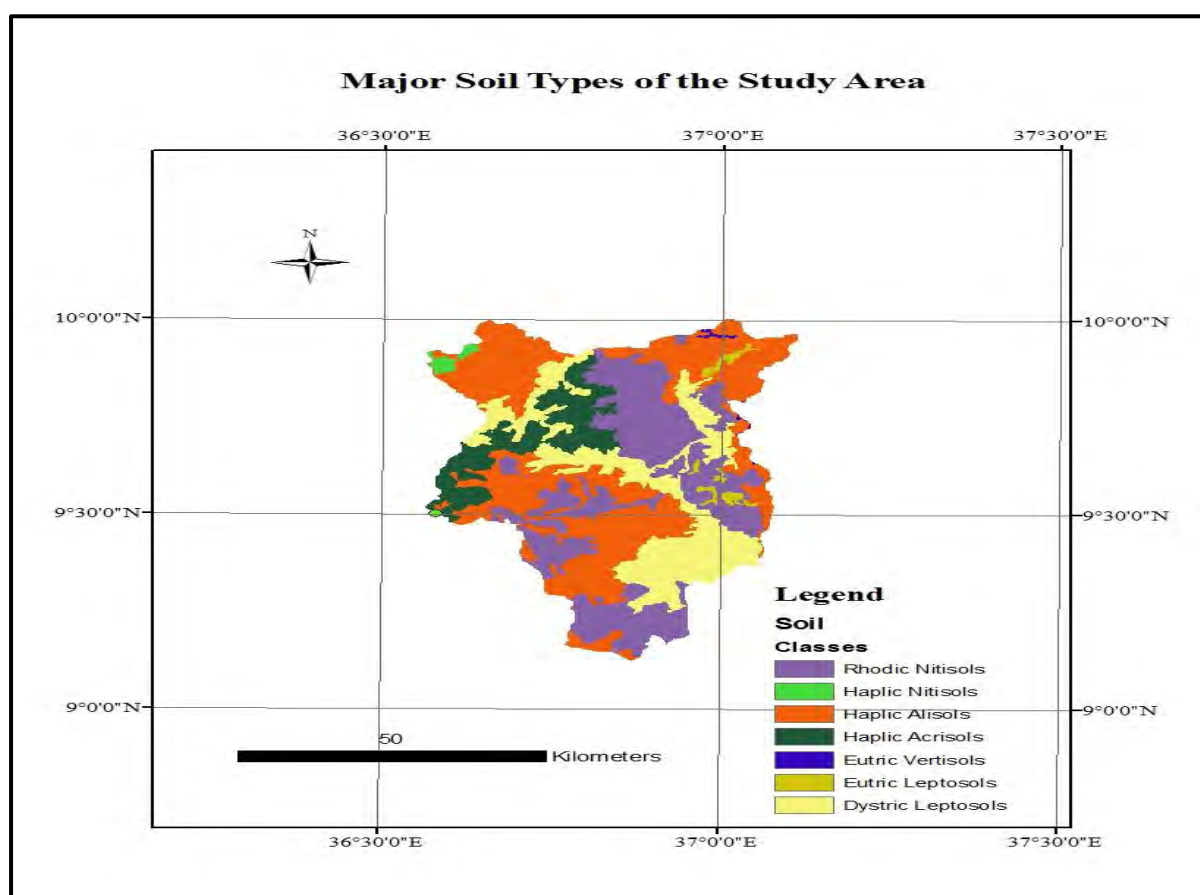


Figure 3.10 Map of the soil types of Angar watershed

3.4.6. Land use and land cover

Spatial distribution and specific land use parameters were required for modeling. SWAT has predefined land uses identified by four letter codes and it uses these codes to link land use maps to SWAT land use databases in the GIS interface.

Hence, while preparing the lookup table, the land use types were made compatible with the input needs of the model. Hence the classified land use map and its attribute were adjusted to the SWAT model requirement format and database. Forest is the dominant land use in the Angar River watershed.

Table 3.3 Land use/cover classification of Angar watershed as per SWAT model

Land use	Area (ha)	Percentage of the area Covered
1 Forest-Mixed (FRST)	129900	35.81
2 Forest-Evergreen (FRSE)	95112	26.22
3 Range-Grasses (RNGE)	54572	15.04
4 Agricultural Land-Generic (AGRL)	37004	10.2
5 Winter Pasture (WPAS)	24388	6.72
6 Pasture (PAST)	11660	3.21
7 Agricultural Land-Row Crops (AGRR)	10012	2.76
8 Agricultural Land-Close-grown (AGRC)	108	0.03

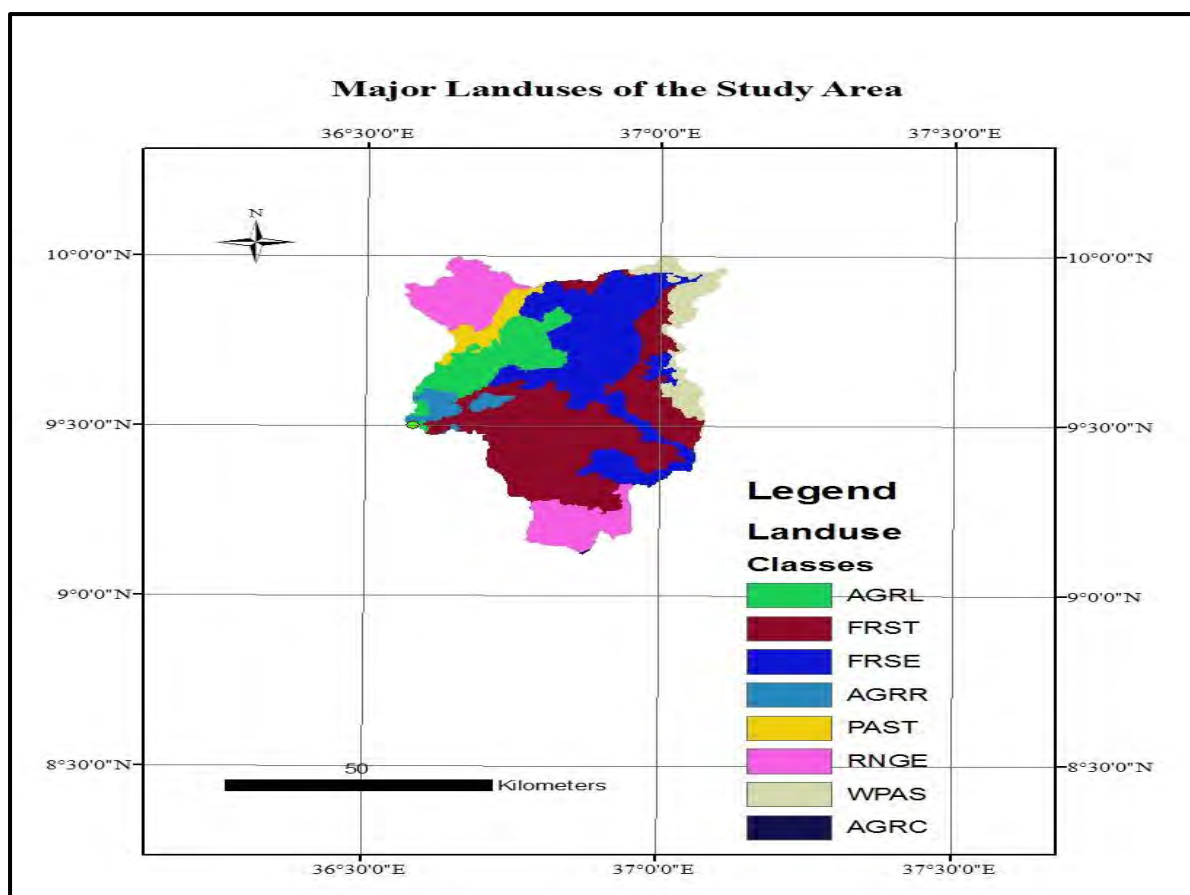


Figure 3.11 Land use/Land cover map of the study area

3.5. Model setup

3.5.1. Watershed delineation

The watershed and sub watershed delineation was performed using 90 m resolution DEM data using Arc SWAT model watershed delineation function. First, the SWAT project set up was created. The watershed delineation process consists of five major steps, DEM setup, stream definition, outlet and inlet definition, watershed outlets selection and definition and calculation of sub basin parameters. Once, the DEM set up was completed and the location of outlet was specified on the DEM, the model automatically calculates the flow direction and flow accumulation. Subsequently, stream networks, sub watersheds and topographic parameters were calculated using the respective tools.

The stream definition and the size of sub basins were carefully determined by selecting threshold area or minimum drainage area required to form the origin of the streams. Using a threshold value suggested by the Arc SWAT interface (7255.0 hectares), the Angar Watershed was delineated in to 23 sub watersheds having an estimated total area of 3627.6

Km². But the total area of the watershed as obtained from the Ministry of Water, Irrigation and Energy (MoWIE) was estimated to be 3742 Km². There is a slight deviation between the delineated and that obtained from the MoWIE database. The difference in the total area between the delineated and the database may be due to the difference in the DEM resolution or the watershed delineator model used.

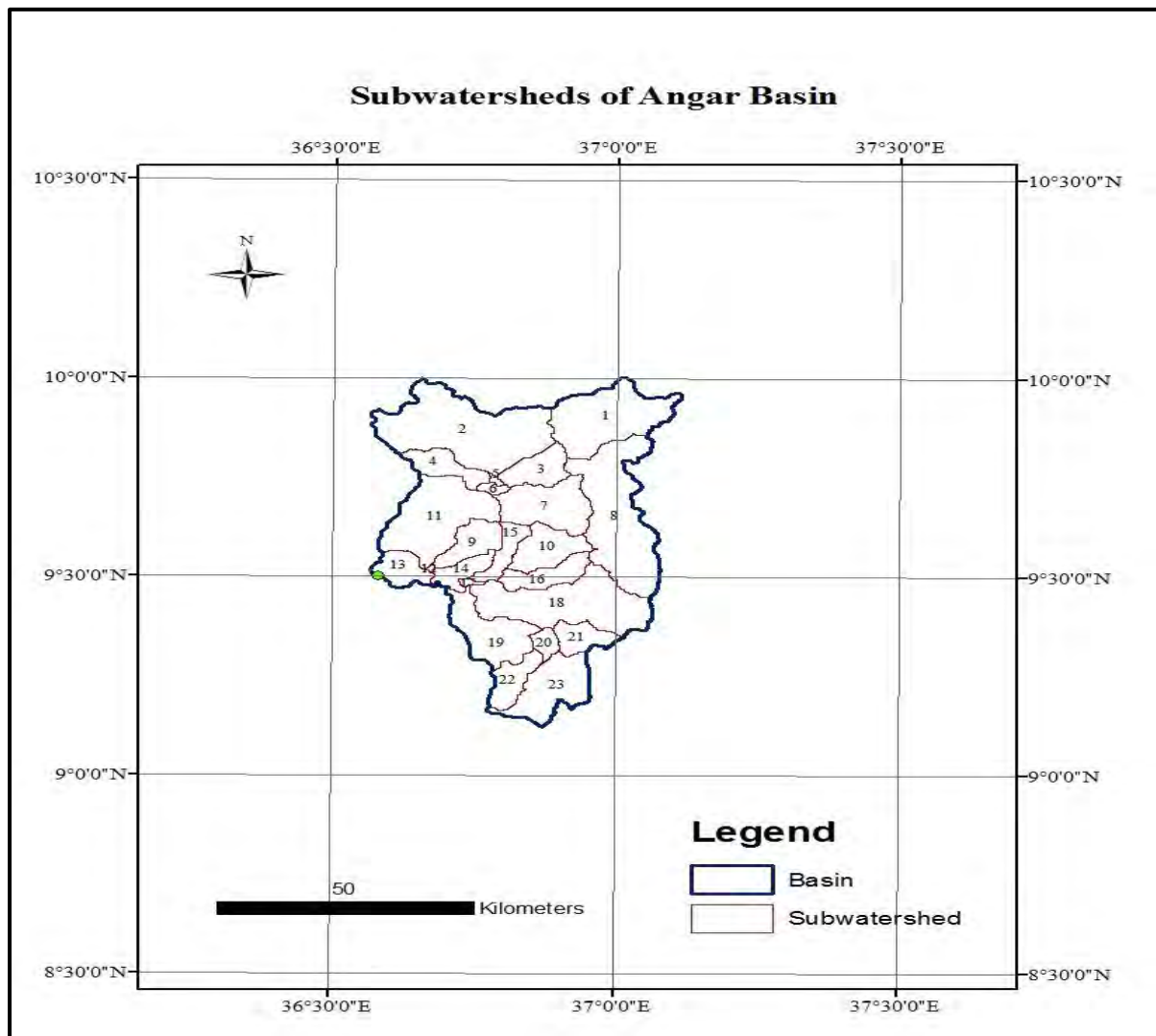


Figure 3.12 Sub watersheds map of the Angar watershed

During the watershed delineation process, the topographic parameters (elevation, slope) of the watershed and its sub watershed were also generated from the DEM data. Accordingly the elevation of the watershed ranges from 1286 m to 3057 above mean sea level.

3.5.2. Hydrologic response units analysis

Hydrologic response units (HRUs) are lumped land areas within the sub-basin that comprised of unique land cover, soil and management combinations. HRUs enable the model to reflect

differences in evapotranspiration and other hydrologic conditions for different land covers and soils. The runoff is estimated separately for each HRU and routed to obtain the total runoff for the watershed. This increases the accuracy in flow prediction and provides a much better physical description of the water balance. The land use and soil data in a projected shape file format were loaded into the SWAT interface to determine the area and hydrologic parameters of each land-soil category simulated within each sub-watershed. The land cover classes were defined using the look-up table. A look-up table that identifies the 4-letter SWAT code for the different categories of land cover/land use was prepared so as to relate the grid values to SWAT land cover/land use classes. After the land use SWAT code is assigned to all map categories, calculation of the area covered by each land use and reclassification were done. As for the land use, the soil layer in the map was linked to the user soil database information by loading the soil look-up table and reclassification applied. The DEM data used during the watershed delineation was also used for slope classification. After the reclassification of the land use, soil overlay operation was performed.

The second step in the HRU analysis was the HRU definition. The HRU distribution in this study was determined by assigning multiple HRU to each sub-watershed. In multiple HRU definition, a threshold level was used to eliminate minor land uses, soils or slope classes in each sub-basin. Land uses, or soils which cover less than the threshold level are eliminated. After the elimination process, the area of the remaining land use, or soil was reapportioned so that 100% of the land area in the sub-basin is modeled. The threshold levels set is a function of the project goal and amount of detail required. The SWAT user's manual suggests that a 20% slope threshold is adequate for most modeling application. However, Setegn *et al*, 2008, suggested that HRU definition with multiple options that account for 10% land use, 20% soil and 10% slope threshold combination gives a better estimation of runoff and sediment components. Therefore, for this study, HRU definition with multiple options that accounts for 10% land use, 20% soil and 10% slope threshold combination was used.

The third step in HRU definition is selection of slope classification option (single or multiple). For this study multiple slope option (an option for considering different slope classes for HRU definition) was selected and the slope class was classified to three and the range was 0-10%, 10-20% and above 20%.

Lastly, by define the HRUs within a sub-basin completes the HRU setup. For this study the option of multiple HRU was selected and 10%, 20% and 10% were the threshold area of land

use, soil and slope in each HRU from the sub-basin values respectively. These threshold values indicate that land uses which form at least 10% of the sub watershed area and soils which form at least 20% of the area within each of the selected land uses will be considered in HRU.

The reason for taking these thresholds values was in order to keep the HRUs to a reasonable and manageable number and also considering computer processing time required. However, application of these thresholds eliminates the land uses and soils that covered relatively small areas in the sub-basins it creates a total of 105 HRUs for 23 sub-basins.

3.5.3. Weather generator

In developing countries, there is a lack of full and realistic long period of climatic data. Therefore, the weather generator solves this problem by generating data from the observed one (Danuso, 2002). The weather generator model included in SWAT is used to fill missing values in measured records and also to simulate the data if simulation option is selected. The Model requires the daily values of all climatic variables from measured data or generated from values using monthly average data over a number of years. The WXGEN was provided with all the necessary statistical information from the meteorological records of the watershed to fill the missing portion properly. After loading this WXGEN parameter and location table, the daily meteorological data (daily precipitation, daily minimum and maximum air temperature, relative humidity, sunshine hour and wind speed) with the missing data filled with a missing data identifier of -99 and including the corresponding location table prepared according to the SWAT format were loaded to the model. The climatic stations which were used in the study are called *Anger, Shambu, Neshi, Dedessa and Nekemte*. *Nekemte* station selected to be principal station for the weather generator.

The SWAT Model contains weather generator model called WXGEN (Shapley and Williams, 1990). It is used in SWAT model to generate climatic data or to fill missing data using monthly statistics which is calculated from existing daily data. From the values of weather generator parameters, the weather generator first separately generates precipitation for the day. Maximum temperature, minimum temperature, solar radiation and relative humidity are then generated. Lastly, the wind speed is generated independently.

To generate the data, weather parameters were developed by using the weather parameter calculator pcpSTAT.exe and dew point temperature calculator Dewo2. The pcpSTAT program calculates the monthly daily average and standard deviation as well as probability of

wet and dry days, skew coefficient, and average number of precipitation days in the month by reading of the daily values of the variables. Average daily dew point temperature was calculated using the dew point calculator (Dew 02) from daily maximum temperature, daily minimum temperature and average relative humidity.

3.5.4. Sensitivity analysis

There are several parameters which affect a complex hydrological modeling. Most of the values of these parameters are not exactly known. This can be for many reasons. Spatial variability, measurement error, incompleteness in description of both the elements and processes present in the system are some of the reasons (Holvoet et al., 2004). Therefore, optimizing internal parameters of a model is an important task in order to achieve a well representative hydrological model. This kind of task is called model calibration which is usually supported by sensitivity analysis. Sensitivity analysis helps to determine the sensitivity of parameters by comparing the output variance due to input variability. It also facilitates selecting important and influential parameters for a model calibration by indicating the parameters that shows higher sensitivity to the output due to the input variability. Therefore, the number of parameters that can be involved for calibration will be less in number and influential. It also evaluates the model capacity and helps to understand the behavior of the system being modeled.

Sensitivity analysis was performed to determine the influence a set of parameters had on predicting total flow. The analysis was carried out to identify the SWAT's hydrologic sensitive parameters by comparing their relative sensitiveness. It was performed on Twenty-six different SWAT parameters. An observed stream flow data of 11 years from 01 January 2001 to 31 December 2011, of little Angar gauge station was used. Using ArcSWAT interface, the position of the sub basin that contains the gauge station was identified. Then the model parameters used in the sensitivity analysis of stream flow were selected and the method algorithm for analysis was defined. In the project the Latin Hypercube One factor At a Time (LH-OAT) sensitivity analysis method was used. It is combination of the One-factor-At-a-Time (OAT) design for simulation and Latin Hypercube (LH) sampling. It basically has the same concept as that of Monte Carlo simulation except the sampling method which is used by LH-OAT is stratified sampling rather than random sampling (McKay et al., 1979; McKay, 1988).

By applying default lower and upper boundary parameters values, the parameters were tested for sensitivity analysis for the simulation of the streamflow. „Average criteria“ options have been selected for „sensitivity analysis output“. Finally the sensitivity analyses were run for the Little Angar gauge station. In the analysis, the sensitive parameters of the stream flow of the basin were identified. Up on the completion of sensitivity analysis, the mean relative sensitivity (MRS) values of the parameters were used to rank the parameters, and their category of classification. Out of the twenty-six stream flow parameters the nine most sensitive ones were chosen for calibration processes. The category of sensitivity was defined based on the (Lenhart et al., 2002) classification presented below (Table 3.4).

Table 3.4 SWAT parameters sensitivity class

Class	MRS	Sensitivity category
I	$0.00 \leq \text{MRS} < 0.05$	Small to negligible
II	$0.05 \leq \text{MRS} < 0.20$	Medium
III	$0.2 \leq \text{MRS} < 1$	High
IV	$\text{MRS} > 1$	Very high

Based on the above classification, parameters producing MRS values of medium, high and very high were selected for calibration process.

3.5.5. Model calibration

Model inputs and values of parameters are associated with a number of uncertainties. Therefore model calibration is an important task to improve the result of model simulation. It is a process in which parameter adjustment are made in order to simulate as closely as possible the hydrological behavior of the watershed. A proper model calibration is necessary to consider a good fit between simulated and observed watershed runoff volume (water balance), the shape of the hydrograph, the peak flow, and the base flow. All these objectives are considered during model calibration because a single objective function cannot establish a reasonable match between simulated and observed data. Most calibrations are supported by sensitivity analysis which avoids performing calibration on non-effective parameters.

Calibration can be performed in two ways: either manually or automated. In ArcSWAT2009 Manual Calibration Helper used for making adjustment to parameters across a user defined group of HRUs or sub basins. Auto calibration and Uncertainty of ArcSWAT 2009 is used for automated calibration. It has two dialogue boxes namely Auto-Calibration Input and

Auto-Calibration Output. The earlier allows performing the automatic model calibration by selecting a simulated model and a subbasin which a discharge outlet located at. The latter provides option to refine to the out parameters for an analysis.

The parameters which had been selected by sensitivity analysis were optimized by both auto and manual calibration. In auto calibration, an Objective Function such as Sum of the Squared of the Residuals (SSQ) works at matching a simulated series to a measured time series.

The auto-calibration was performed incorporating with Parameter Solution (Parasol) method which is an optimization and statistical method for the assessment of parameter uncertainty (Van Griensven and Meixner, 2006). It was performed based on a multiobjective function which incorporates the algorithms of the Shuffled Complex Evolution (Duan et al., 1992). The method uses best-fit statistics to adjust parameter values. In other word the method use algorithms perform to optimize an objective function by systematically searching the parameter space according to a fixed set of rule (Van Griensven et al., 2007).

The streamflow observed data of eleven years, from 01 January 2001 to 31 December 2011, derived from Little Angar gauging station was used for auto calibration. Before performing the Auto calibration process the data was divided into two parts: 01 January 2001 to 31 December 2007 for calibration and 01 January 2008 to 31 December 2011 for validation.

While dealing SWAT auto calibration process, the location of the subbasin which the discharge outlet located at was indicated. The observed streamflow data was added. The nine most sensitive parameters, which were identified through sensitivity analysis process, were selected for auto calibration. The PARASOL was taken as a method for the calibration process. Next, auto calibration output evaluations were defined. „Flow“ was set as an output parameter to evaluate. The sum of squared residuals was used as the objective function of the process. The objective function weight was set and the model output evaluation criterion was also defined. Finally, the calibration process was run.

The auto calibration process has produced optimized parameters and their output values in the form of text files which some of them were used for rerun the simulated model and further manual calibration process to much up the trend between the observed and the simulated flow parameter. The Best Parameters and their values, which are resulted from auto

calibration process, were reused to rerun the simulated model to fit the parameters of simulated flow to the parameters of observed streamflow.

The simulated model was rerun using Best Parameters and the values which were obtained from auto calibration. Then the output result was imported to SWAT database and stored in the output file which contains the simulated flow and other water quality parameters in stream. The simulated flow of the outlet subbasin was sorted out from entire simulated flow of Angar Basin. Next, it was compared to the measured streamflow of the watershed. The comparison simulated and observed streamflow was done using time-series plot of the measured and simulated monthly flows diagram created in excel 2007.

Since manual calibration gives a better result on fitting the parameters of simulated and observed flow, it was utilized following to the auto calibration. Among the nine parameters which resulted from sensitivity analysis method Base flow alpha factor (Alpha_Bf), Curve number (CN) and available water capacity of the soil layer (Sol_Awc) were found the most influential parameters and were used for further manual calibration. The rest of six parameters have also been used for manual calibration but in lesser degree compared to above three parameters.

Following the manual calibration process, the model was run using the best parameter values and the simulations were compared with observed stream flow data using Nash and Sutcliffe coefficient of efficiency (NSE), coefficient of determination (R^2), Root Mean Square Error (RMSE) and BIAS.

3.5.6. Model validation

Validation was also done to compare the model outputs with an independent data set without making further adjustment of the parameter values. Model validation is comparison of the model outputs with an independent data set without making further adjustment which may be adjusted during calibration process.

Model performance in calibration and validation periods may not be similar. Recent studies revealed that there are a number of difficulties of climate model validation. That is because of the complexity of the nature of climate and time dependent uncertainties of modeling dataset. Another reason is the hydrologic condition in the calibration period may not be the same as the hydrologic condition during the validation period (Beven, 2006; Liu and Gupta, 2007; Zhang et al., 2009a).

The measured data of average monthly stream flow data of 3 years from January 2008 to December 2010 were used for the model validation process. In this process, the two model performance values were also checked here to make sure that the simulated values are still within the accurate limits.

3.6. Model performance evaluation

To evaluate the model simulation outputs in relative to the observed data, model performance evaluation is necessary. There are various methods to evaluate the model performance during the calibration and validation periods. For this study, five methods were used: Coefficient of determination (R^2), Nash and Sutcliffe simulation efficiency (ENS), Root Mean Square Error (RMSE), linear correlation coefficient (R) and BIAS.

The determination coefficient (R^2) describes the proportion the variance in measured data by the model. It is the magnitude linear relationship between the observed and the simulated values. R^2 ranges from 0 (which indicates the model is poor) to 1 (which indicates the model is good), with higher values indicating less error variance, and typical values greater than 0.6 are considered acceptable (Santhi et al., 2001). The R^2 is calculated using the following equation:

$$R^2 = \left[\frac{\sum (x_i - x_{av})(y_i - y_{av})}{\sqrt{\sum (x_i - x_{av})^2} \sqrt{\sum (y_i - y_{av})^2}} \right]^2 \quad (3.12)$$

Where, x_i – simulated value (m^3/s)

x_{av} – average simulated value (m^3/s)

y_i – observed value (m^3/s) and

y_{av} – average observed value (m^3/s)

The Nash- Sutcliffe simulation efficiency (ENS) indicates that how well the plots of observed versus simulated data fits the 1:1 line. ENS is computed using the following equation:

$$ENS = 1 - \frac{\sum (x_i - y_i)^2}{\sum (y_i - y_{av})^2} \quad (3.13)$$

Where, x_i – simulated value

y_i – observed value and

y_{av} – average observed value

The value of ENS ranges from negative infinity to 1 (best) i.e, $(-\infty, 1]$. ENS value ≤ 0 indicates the mean observed value is better predictor than the simulated value, which indicates unacceptable performance. While ENS values greater than 0.5, the simulated value is better predictor than mean measured value and generally viewed as acceptable performance (Santhi *et al.*, 2001).

The Root Mean Square Error (**RMSE**) (also called the root mean square deviation, RMSD) is a frequently used measure of the difference between values predicted by a model and the values actually observed from the environment that is being modeled. These individual differences are also called residuals, and the RMSE serves to aggregate them into a single measure of predictive power. The RMSE of a model prediction with respect to the estimated variable X_{model} is defined as the square root of the mean squared error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{model,i})^2}{n}} \quad (3.14)$$

Where x_{obs} is observed values and x_{model} is modelled values at time/place i .

Correlation – often measured as a correlation coefficient – indicates the strength and direction of a linear relationship between two variables (for example model output and observed values) and is obtained by dividing the covariance of the two variables by the product of their standard deviations. If we have a series n observations and n model values, then the correlation coefficient can be used to estimate the correlation between model and observations.

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3.15)$$

The correlation is +1 in the case of a perfect increasing linear relationship, and -1 in case of a decreasing linear relationship, and the values in between indicates the degree of linear relationship between for example model and observations. A correlation coefficient of 0 means the there is no linear relationship between the variables.

$$Bias = \frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n y_i} \quad (3.16)$$

Where, X_i – simulated value and

Y_i – observed value

CHAPTER FOUR

4. Results and Discussion

4.1. Satellite rainfall products versus rain gauges

To better understand the impacts of rainfall inputs on the output of model, we begin by comparing watershed – averaged rainfall magnitudes derived from satellite rainfall products (i.e. 3B42V7, CMORPH, and PERSIANN) and rain gauge that have been used in Arc SWAT modeling for Angar Watershed. The comparison performed for the study period of 2003-2007, corresponding to ArcSWAT calibration period at Angar Watershed. The direct intercomparison of rainfall products was based on Thiessen Polygon interpolated average rainfall over the watershed and the intercomparison of the satellite rainfall products is meaningful as the products represent average rainfall extracted from each grid that lies in the watershed.

Figure 4.1 clearly shows that the interannual variation of rain observed by rain gauges was generally captured by all satellite rainfall products. The PERSIANN product shows small underestimation of annual rainfall for the whole study period except 2005, but shows better performance in capturing interannual variation of rainfall compared to other. The new version TMPA 3B42V7 product shows small overestimation for the whole study period when compared with CMORPH product. But, when we compare the mean monthly rainfall (Figure 4.2) observed at rain gauge (1580.11 mm) and satellite products (i.e. 1618.39 mm for 3B42V7, 1783.30 mm for CMORPH & 1517.70 mm for PERSIANN), the new version of TMPA 3B42V7 shows relatively very near to rainfall observed at rain gauge. According to Fig.4.3, the correlation values for the time series of daily rainfall observed between rain gauge and satellite products ($R = 0.71$ for 3B42V7 Vs Rain gauge, $R = 0.69$ for Rain gauge Vs CMORPH and $R = 0.64$ for Rain gauge Vs PERSIANN) shows that the new version of TMPA 3B42V7 and CMORPH show relatively better correlation with rainfall observed at the rain gauge.

Analyzing the seasonal rainfall variation was very important since this study area is characterized by only one rainy season (mono-modal) which lasts from June to end of September (JJAS). The seasonal rainfall comparison result revealed that on average more than 70 % of annual rainfall occurs in the summer rainy season (JJAS) for the same study period as observed from rain gauge and satellite products. According to Fig.4.2, the average

percentage of seasonal rainfall variation observed at rain gauge (75.33%) was approximately similar to new research version TMPA 3B42V7 products (75.15%), while CMORPH & PERSIANN underestimates the average percentage of seasonal rainfall; however it's large of PERSIANN.

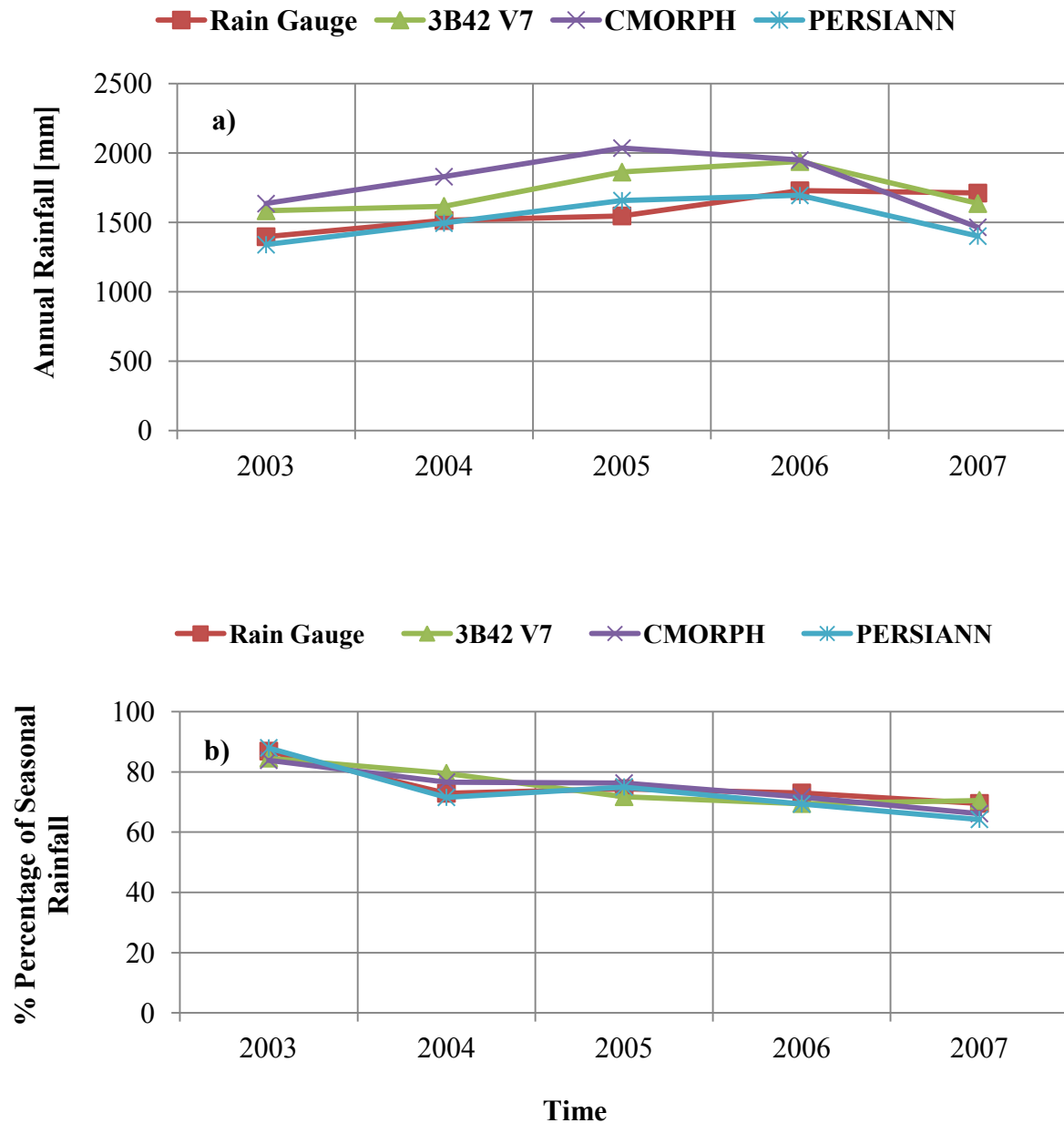


Figure 4.1 Annual (a) and percentage of seasonal (b) rainfall (i.e. JJAS: June, July, August, and September) from various rainfall products for the period of 2003 – 2007.

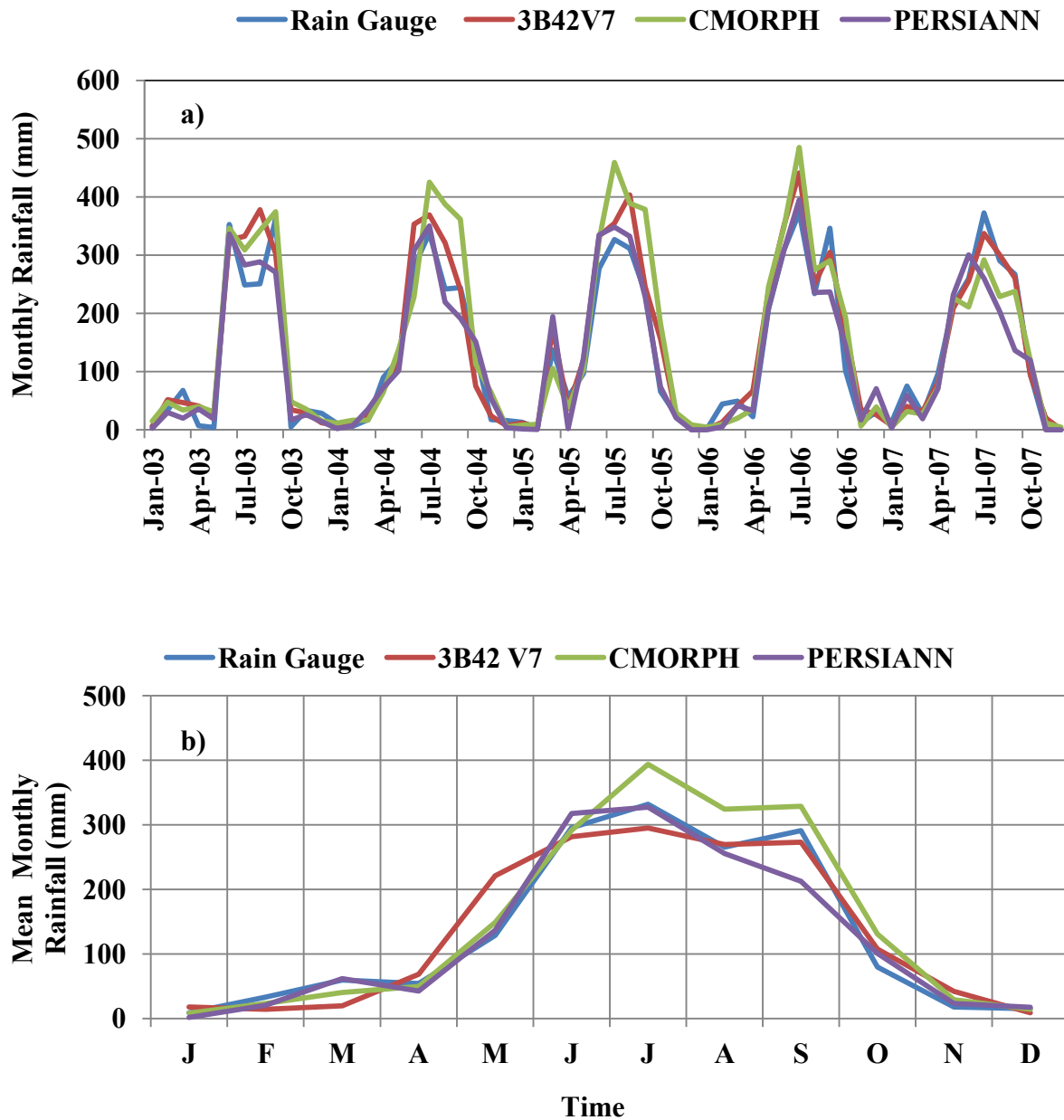
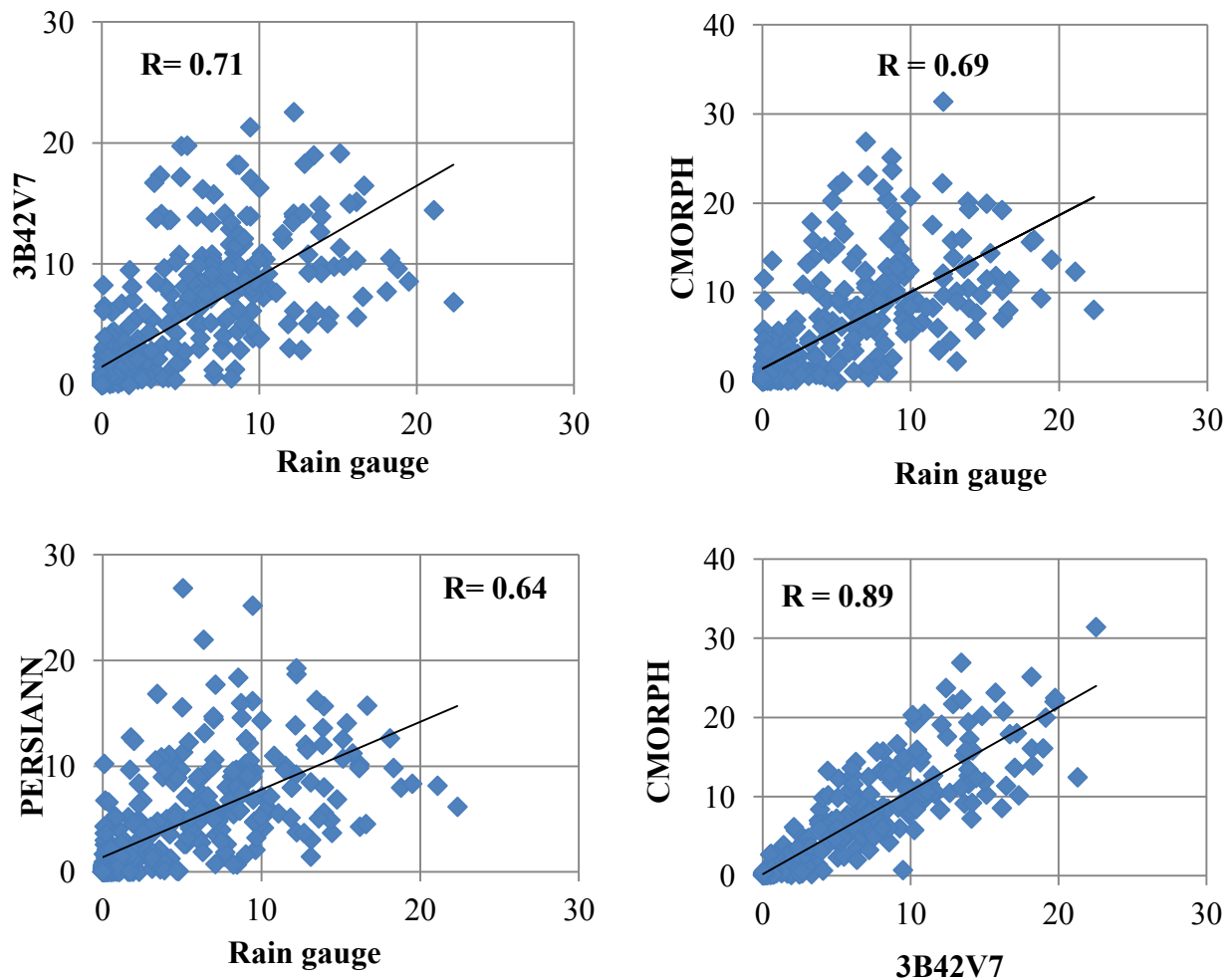


Figure 4.2 Monthly (a) and mean monthly (b) rainfall over Angar watershed obtained from rain gauge data and various satellites rainfall products for the period of 2003 – 2007.

The mean monthly peak rainfall occurs in July month for all rainfall products. Compared to rain gauges values, CMORPH satellite rainfall products over estimated the mean monthly rainfall by 38%, 3B42V7 and PERSIANN under estimated the mean monthly rainfall by 16% and 18% respectively.

Figure 4.3 Presents intercomparison of daily rainfall estimates. The comparison statistics (R) are given in each plot. Correlations between satellite rainfall values and rain gauge values were good (R Values are ranging from 0.64 to 0.89). Relatively Poor correlation was observed between rain gauge and PERSIANN ($R = 0.64$). Correlations between pairs of daily satellite rainfall values were high (R values are ranging from 0.84 to 0.89), with the highest correlation observed between 3B42V7 and PERSIANN ($R = 0.89$) and between 3B42V7 and CMORPH ($R = 0.89$).



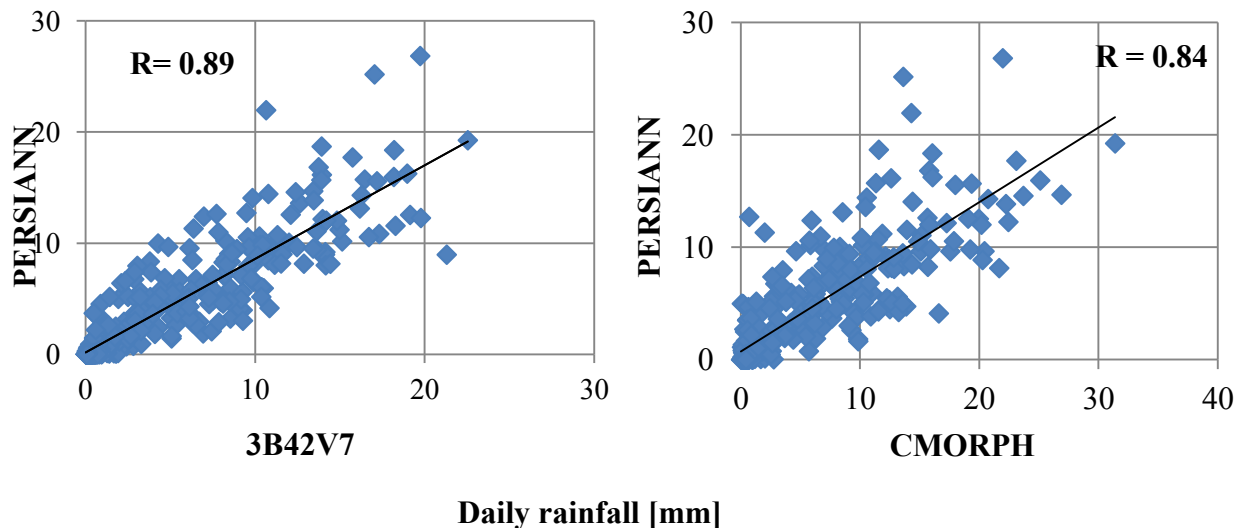
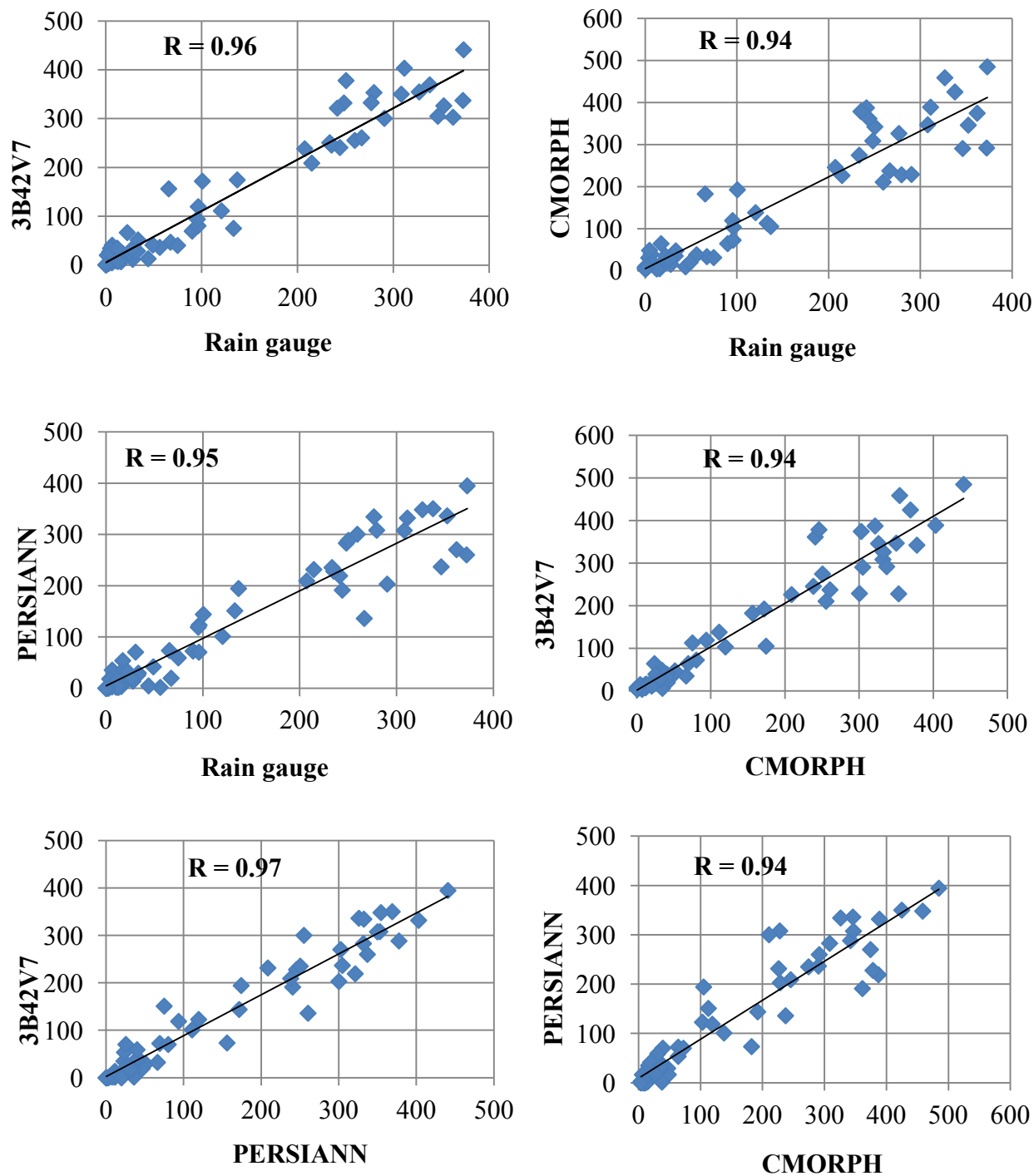


Figure 4.3 Intercomparison of daily rainfall from satellite rainfall products (i.e. 3B42V7, CMORPH, PERSIANN) and rain gauge for Angar watershed for the period of January 2003 – December 2007

Figure 4.4 Presents intercomparison of monthly rainfall estimates. At monthly time scale, the correlation between rain gauges and satellite rainfall values were almost similar with correlation between satellite rainfall values. Correlations between monthly satellite rainfall values and rain gauge values were high (R values ranging from 0.94 to 0.96), which is almost similar with the correlation between satellite monthly rainfall values (R values ranging from 0.94 to 0.97). The highest correlation was observed between 3B42V7 and PERSIANN ($R = 0.97$) and the second highest correlation was observed between rain gauge and 3B42V7 ($R = 0.96$) where as the lowest between Rain gauge and CMORPH ($R = 0.94$).

From the above comparison, especially the daily basis is very important in our case because Arc SWAT uses daily observed rainfall data. Therefore, both 3B42V7 and CMORPH rainfall estimates gave relatively good result and this result also consistent with findings in Koga and Gilgel Abay watersheds of Ethiopia (Bitew and Gebremichael, 2011).

There have been subsequent studies that conducted in Blue Nile River basin to evaluate satellite rainfall products in the estimation of rainfall, Bitew and Gebremichael., 2011, Romilly and Gebremichael (2011), Solomon (2013), and Gebremichael (2010); based on their studies and the result of this study, it is concluded that 3B42V7 and CMORPH are much closer to the actual rainfall fields in Ethiopian highlands compared to PERSIANN.



Monthly rainfall [mm]

Figure 4.4 Intercomparison of monthly rainfall from satellite rainfall products (i.e. 3B42V7, CMORPH, PERSIANN) and rain gauge for Angar watershed for the period of January 2003 – December 2007

4.2. Stream flow modeling

4.2.1. Sensitivity analysis

Sensitivity analysis was performed on flow parameters of SWAT on monthly time steps with observed data of the Little Angar River gauge station. For this analysis, 26 parameters were considered and only 9 Parameters were identified to have significant influence in controlling the stream flow in the watershed. The result of the sensitivity analysis indicated that these 9 flow parameters were sensitive to the SWAT model i.e. the hydrological process of the study watershed mainly depends on the action of these parameters. Alpha factor (ALPHA_BF), Curve number (CN2), soil available water capacity (SOL_AWC), maximum canopy index (CANMX) and hydraulic conductivity of main channel alluvium (CH_K2) were identified to be highly sensitive parameters and retained rank 1 to 5, respectively. The other parameters such as soil evaporation compensation factor (ESCO), threshold depth of water in the shallow aquifer required for return flow (GWQMN), evaporation to occur (REVAPMN) and ground water evaporation coefficient (GW_REVAP) were identified as slightly important parameters that were retained rank 6 to 9, respectively. The remaining parameters (17 parameters) were not considered during calibration process as the model simulation result was not sensitive to these parameters in the watershed.

These parameters are related to ground water, runoff and soil process and thus influence the stream flow in the watershed. The result of the analysis was found that ALPHA_BF is the most important factor influencing stream flow in the Angar watershed. The ALPHA_BF is a direct index of ground water flow response to changes in recharges. In addition, (M.M. Bitew and M. Gebremichael, 2011) through modeling of Koga and Gilgel Abbay watershed in Upper Blue Nile basin found CH_K2, ALPHA_BF and CN2 to retain rank 1 to 3, respectively. Therefore, these may be an additional support to the result of the sensitivity analysis.

4.2.2. Calibration and validation of stream flow simulation

The simulation of the model with the default value of parameters in the Angar watershed showed relatively weak matching between the simulated and observed stream flow hydrographs. Hence, calibration was done for sensitive flow parameters of SWAT with observed average monthly stream flow data. First, auto calibration was run using sensitive flow parameters that were identified during sensitivity analysis. Then, model parameters were calibrated manually. In this procedure, the values of the parameters were varied iteratively

within the allowable ranges until the simulated flow as close as possible to observed stream flow. Table 4.1 presents the result of calibrated flow parameters for various rainfall inputs.

Table 4.1 Default & final calibrated flow parameter values of the watershed

Name	Parameter description	Range	Initial (Default) values	Calibrated value
ALPHA_BF	Base flow alpha factor (days)	0-1	0.048	0.00085
CN2	Curve Number (%)	($\pm$)25	36-74	(+)15
SOL_AWC	Available water capacity of the soil layer (mm/mm)	($\pm$)25	0.08	(+)25
CANMX	Maximum canopy Index	0-10	0	3
CH_K2	Hydraulic conductivity of main channel alluvium (mm/hr)	0-150	0	130
ESCO	Soil evaporation compensation factor	0-1	0	0.01
GWQMN	Threshold depth of water in the shallow aquifer required for return flow (mm)	0-5000	0	1000
REVAPMN	Evaporation to occur (mm)	0-500	1	2
GW_REVAP	Ground water evaporation coefficient	0.02-0.2	0.02	0.16

Therefore, we calibrated the model parameters for the period of seven-years from 2001 to 2007. However, as the first two years were considered for model warm up period, calibration was performed for 5 years from 2003 to 2007.

Generally, it's observed that there was an overall agreement between simulated and observed hydrographs, even if the simulated flows (especially, the satellite rainfall products) overestimated the observed peak flow for some year.

4.3. Satellite rainfall simulation of stream flow

According to Fig.4.5, the comparison of model simulation using rainfall input from various rainfall input (i.e. rain gauge, 3B42V7, CMORPH and PERSIANN) with observed monthly stream flow for Angar watershed is presented.

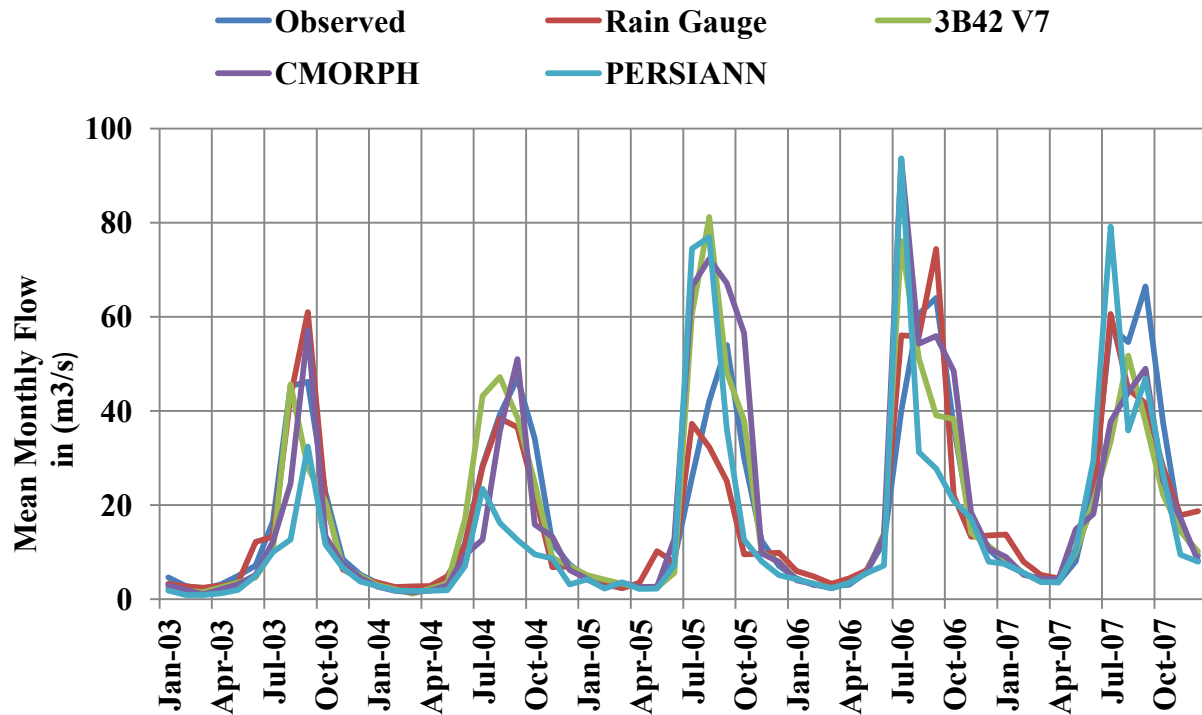


Figure 4.5 Comparison of Arc SWAT simulated stream flow (based on rainfall input data from various sources in the legend) and observed monthly stream flow during the calibration period.

Fig. 4.5 shows the statistical comparisons of simulations of Arc SWAT model from various rainfall inputs. Simulations of stream flow from rain gauge, 3B42V7, and CMORPH inputs showed comparatively good performance, while simulation based PERSIANN inputs resulted in small R^2 and NSE value. Simulation based on rain gauge has relatively high NSE (0.83) and R^2 (0.83) values. Simulation based on 3B42V7 and CMORPH has nearly similar NSE and R^2 value (0.66 & 0.7 for 3B42V7 and 0.59 & 0.7 for CMORPH) respectively. There were very small values of NSE (0.32) and R^2 (0.49) in PERSIANN based simulation of stream flow. RMSE result for the rain gauge and satellite rainfall products are shown in Table 8.

Fig 4.6 shows the statistical comparisons of simulations of Arc SWAT model from various rainfall inputs. Simulations of streamflow from rain gauge, 3B42V7 and CMORPH inputs

had showed comparatively similar performance, while simulations based PERSIANN inputs resulted in large negative bias and small NSE value. Simulation based on rain gauge has relatively high NSE (0.83) and R^2 (0.83) values. Simulation based on 3B42V7 and CMORPH have nearly similar NSE and R^2 value (0.66 & 0.7 for 3B42V7 and 0.59 & 0.7 for CMORPH) respectively. Small value of biases were observed in the simulations based on rain gauge (% BIAS = - 6 %), 3B42V7 (% BIAS = -2 %), and CMORPH (% BIAS = 4 %). There were very small values of NSE (0.32), R^2 (0.49) and relatively large negative bias (% BIAS = -18 %) in PERSIANN based simulation of streamflow.

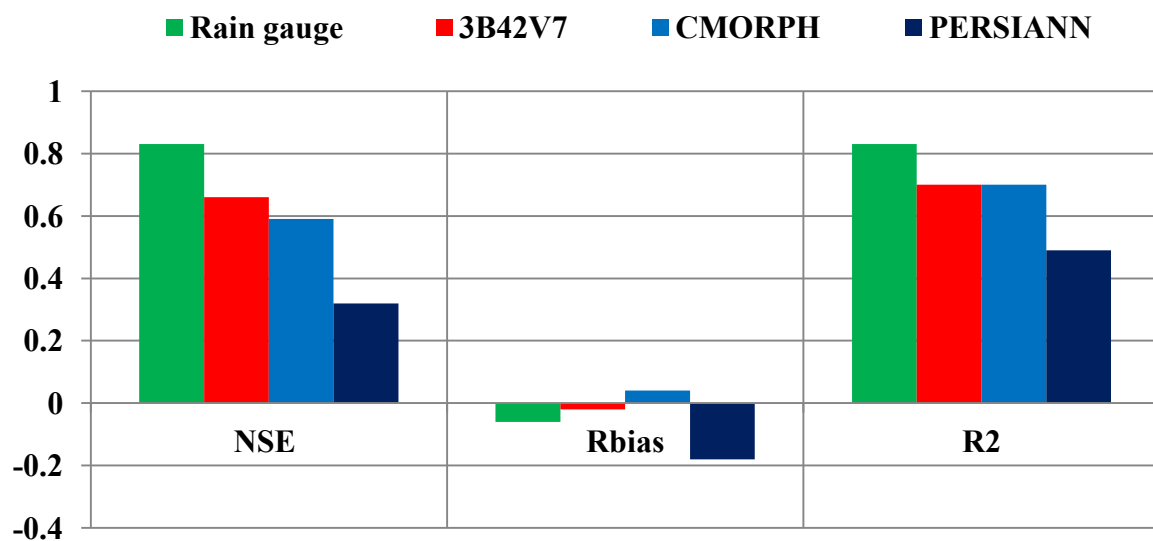
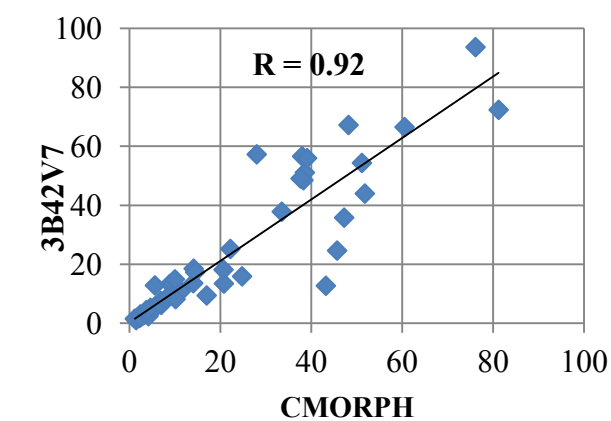
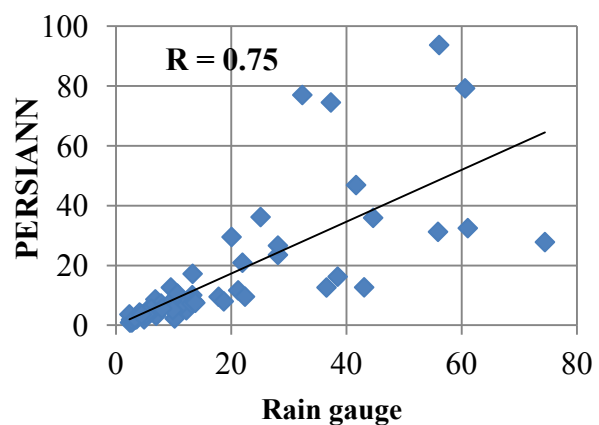
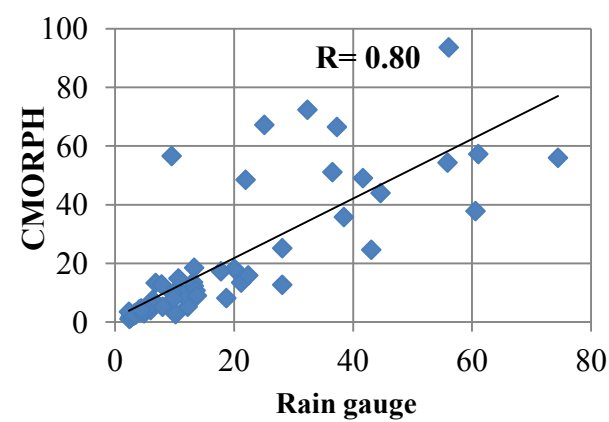
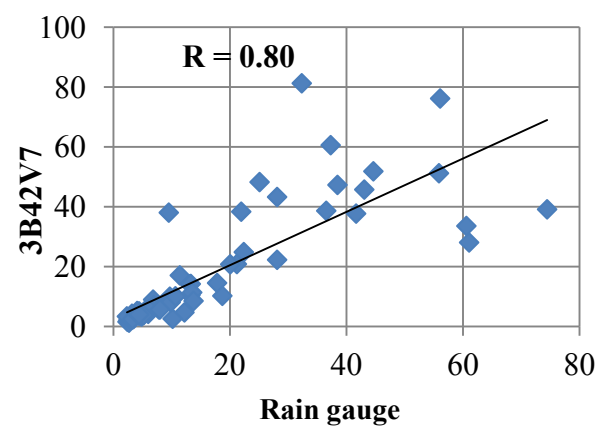
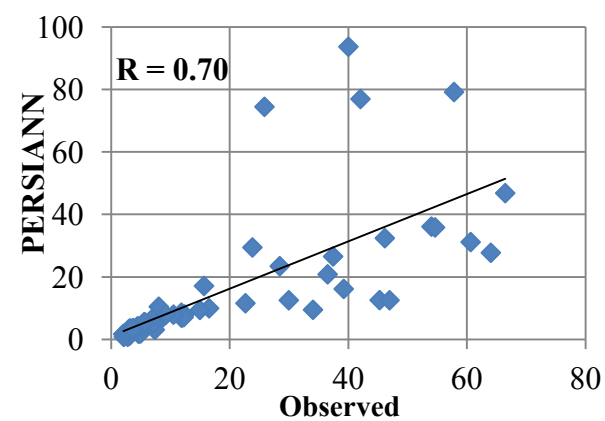
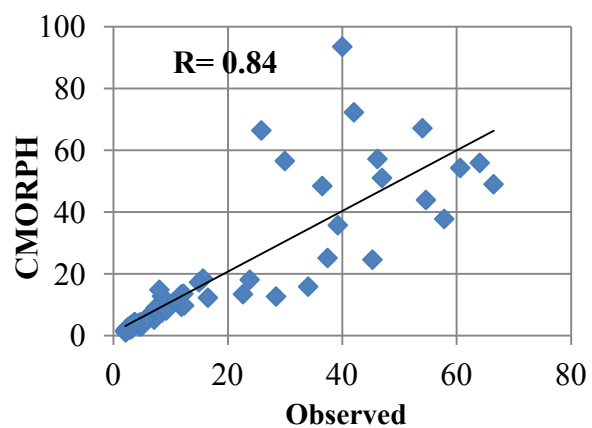
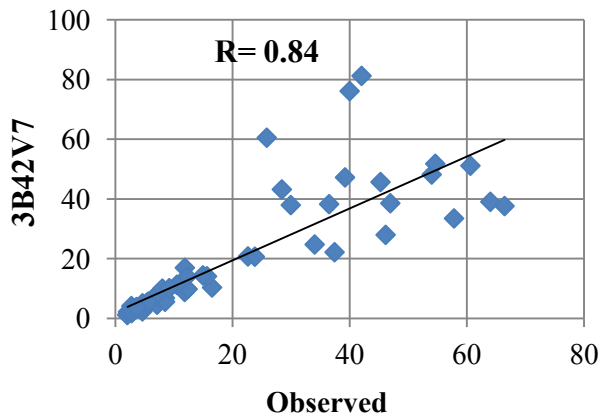
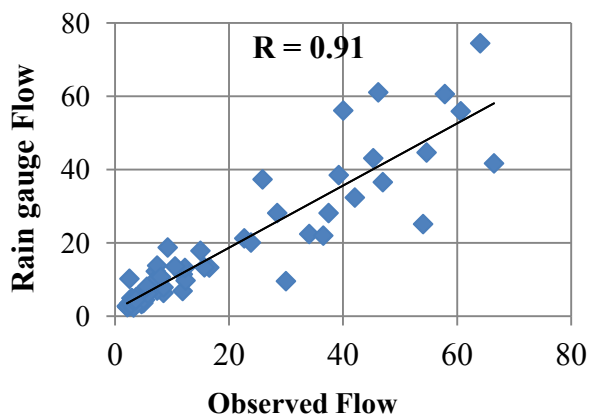


Fig. 4.6 Statistical comparison of simulations of monthly streamflow based on various rainfall inputs indicated in the legend.

Fig. 4.7 presents the intercomparison of model simulation based on various rainfall inputs using linear correlation coefficient (R). The result revealed that high correlation ($R = 0.92$) was observed between simulations based on 3B42V7 and CMORPH rainfall inputs whereas the correlation between Observed and PERSIANN based simulation were the lowest ($R = 0.70$). The highest correlation value was observed between model simulations from satellite rainfall inputs. The next highest correlation was observed between observed and simulations based on rain gauge rainfall inputs ($R = 0.91$). In general, good correlation value observed between simulation based on rain gauge rainfall inputs and simulations based on satellite rainfall inputs.



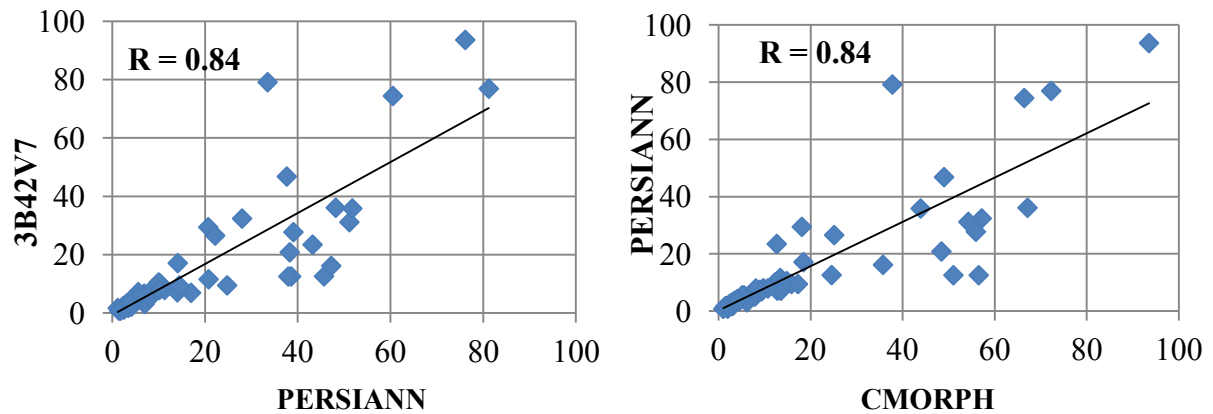


Figure 4.7 Intercomparison of simulated monthly stream flow (based on 3B42V7, CMORPH, PERSIANN and rain gauge rainfall input data) and observed monthly stream flow, the period of Jan, 2003 - Dec, 2007

Fig. 4.8 presents the comparison of mean monthly rainfall estimates of various rainfall products versus simulations of mean monthly streamflow from the rainfall products. Mean monthly rainfall estimates of 3B42V7 shows relatively small positive bias (% BIAS = 2%) compared with CMORPH which shows large positive bias (% BIAS = 13%) and PERSIANN shows negative bias (% BIAS = - 4%).

Though CMORPH shows relatively large rainfall variation with rain gauge, its streamflow simulation shows small variation with observed streamflow data. In other words, the error in CMORPH rainfall product did not propagate to streamflow. But, for the rest of the satellite products (3B42V7 & PERSIANN) the error in their rainfall product propagated to streamflow.

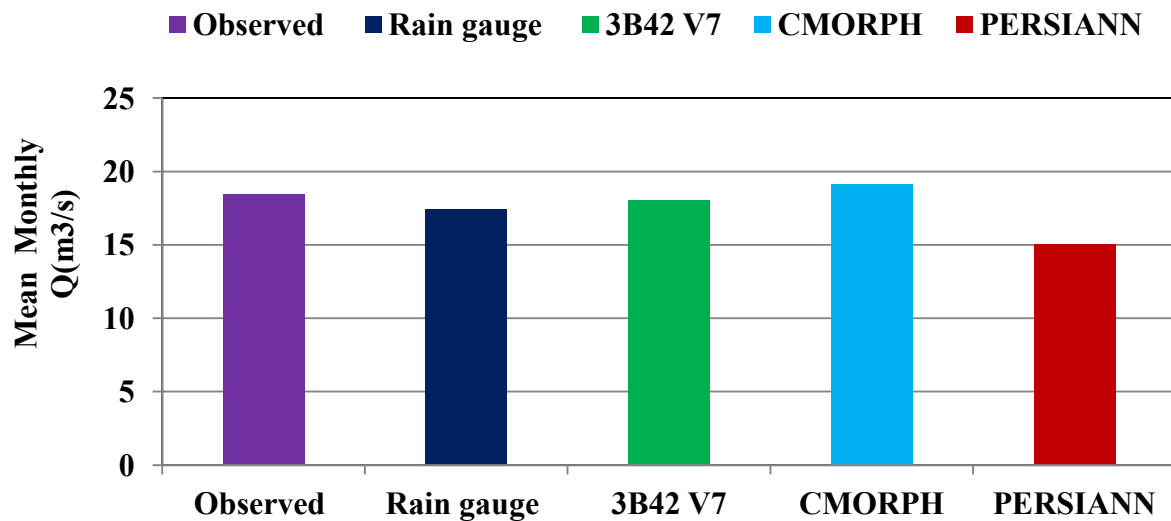
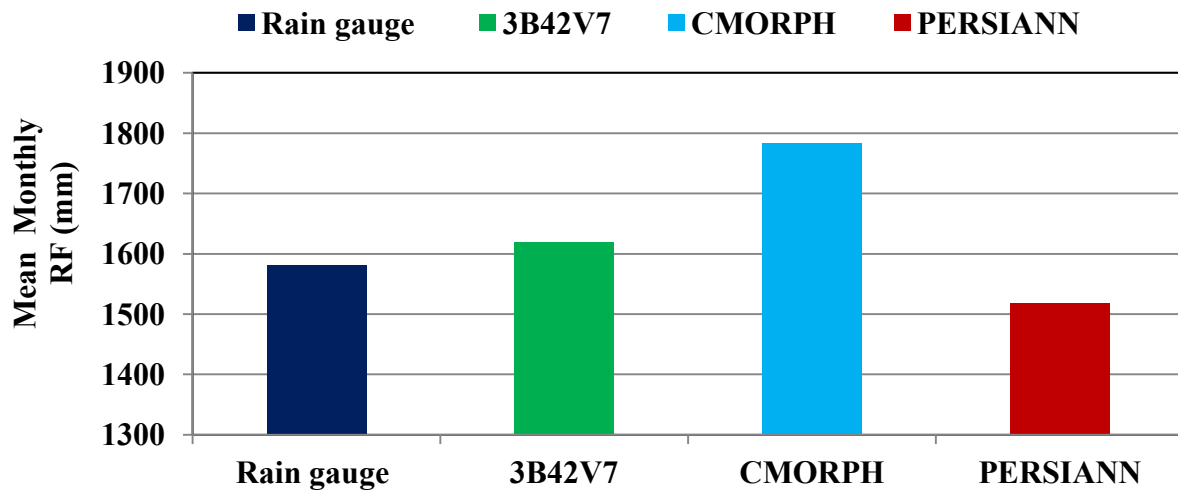


Figure 4.8 Comparison of mean monthly rainfall estimates of various rain fall products versus simulations monthly stream flow (based on 3B42V7, CMORPH, PERSIANN and rain gauge rainfall input data), the period of Jan, 2003 - Dec, 2007

4.4. Discussion

The study of the three satellite rainfall products (3B42V7, CMORPH and PERSIANN) indicated that the performance of the satellite rainfall products are progressively improving and proving its usefulness in hydrological and water resources studies. In this study, the 3B42V7 and CMORPH rainfall estimates gave relatively good performance compared to observed rainfall. This result is consistent with findings in other watershed in Upper Blue Nile such as Koga and Gilgel Abay watersheds of Ethiopia (e.g. Bitew and Gebremichael, 2011). Furthermore, there have been subsequent studies that conducted in Blue Nile River basin to evaluate satellite rainfall products in the estimation of rainfall (e.g. Romilly and Gebremichael (2011), Solomon (2013), and Gebremichael (2010)) that concluded 3B42V7 and CMORPH are much closer to the actual rainfall fields in Ethiopian highlands compared to PERSIANN. Despite such growing potential of the predictive accuracy of the satellite rainfall products in reproducing hydrological features, further research is required towards error characterization and uncertainty analysis to indicate the potential improvement in the satellite rainfall products.

Simulations of streamflow from rain gauge, 3B42V7, and CMORPH inputs showed comparatively good performance, while simulation based PERSIANN inputs resulted in small R^2 and NSE value and large negative bias and large RMSE value. Simulation based on rain gauge has relatively high NSE (0.83) and R^2 (0.83) values and small RMSE value and relatively small bias. Simulation based on 3B42V7 and CMORPH has nearly similar NSE and R^2 value (0.66 & 0.7 for 3B42 V7 and 0.59 & 0.7 for CMORPH) respectively. The performances are relatively small in terms of NSE (0.32) and R^2 (0.49) for PERSIANN products.

The 3B42V7 and CMORPH products reproduce the discharge hydrograph reasonably well than PERSIANN. Further analysis and research can focus on error analysis to reveal the issues on peak flood areas. The overestimation of the large flood peaks in the years 2005 & 2006 may be systemic or related to data accuracy.

As noticed on modeling process, there are a number of factors which may interfere in the performance of the accuracy of SWAT modeling. For instance, the filling of missing data gaps in precipitation and observed data was attained by using simple linear interpolated data. This may be sources of bias in the areal data given the limited density of gauged data. Some irregular data series may also be associated to measurement error. It is clear that these types

of errors could be propagated in a modeling process and could affect the model result. In future, research should focus on understanding how the satellite rainfall products affect hydrological process such as soil moisture, Actual evapotranspiration.

CHAPTER FIVE

5. Conclusions and Recommendations

5.1. Conclusion

Satellite-based precipitation estimates with high spatial and temporal resolution and large areal coverage provide a potential alternative source of forcing data to hydrological models in regions where conventional in situ precipitation measurements are not readily available. The Upper Nile basin provides a good example of a case where the use of satellite-derived precipitation could be beneficial. To overcome limitation of lacking robust station data, this research study uses some of the available globally gridded high resolution precipitation datasets to simulate runoff using SWAT model. We limited our analyses to the following specifics: the semi-distributed hydrologic model SWAT; Angar Watershed (3627.6 km²) and three types of satellite precipitation products (3B42V7, CMORPH and PERSIANN).

Hydrologic simulation models are very essential way used to assess hydrological characteristics of watershed. They are efficient tools for evaluating effects and impacts that occur in hydrologic regime. They can be used to find out, predict and interpret what happened and will happen throughout a basin in time and space. They can be used to learn influence of various environmental and other factors that are interrelated to the basin and their impact and adverse impact when they are at risk.

The comparison results of rainfall magnitudes from satellite rainfall products and rain gauges revealed that the new version TMPA 3B42V7 and CMORPH rainfall estimates gave relatively accurate result compared to rain gauge rainfall estimates at daily scale. The result revealed that, in this region, correlations between satellite rainfall values and rain gauge values were good (R Values are ranging from 0.64 to 0.89). Comparatively the highest correlation was observed between rain gauge and 3B42V7 ($R = 0.89$) and Poor correlation was observed between rain gauge and PERSIANN ($R = 0.64$). But PERSIANN is a good data set under monthly and yearly scale. Based on subsequent studies conducted in Blue Nile River basin to evaluate satellite rainfall products in the estimation of rainfall and the result of this study, it is concluded that 3B42V7 and CMORPH are much closer to the actual rainfall fields in Ethiopian highlands compared to PERSIANN.

Our results reveal that the utility of satellite rainfall products as input to SWAT for monthly stream flow simulation strongly depends on the product type. Simulation from all rainfall inputs captured the trend of observed hydrograph. Simulation based on the new version of TMPA 3B42V7 and CMORPH showed consistent and modest skills in their simulations but overestimates the large flood peaks slightly in some years (2005 & 2006). On the other hand, the infrared – based algorithm PERSIANN simulation showed poor skills in reproducing the hydrograph at the outlet of the watershed. One of the reasons may be due to PERSIANN product is relatively less correlated to observed rainfall and other satellite products.

Overall, the results indicate that although some uncertainties exist in these gridded datasets (3B42V7 & CMORPH), the application of these gridded data prove useful for hydrological studies in the absence of station data. It strongly indicates the growing potential of the predictive accuracy of the satellite rainfall products in reproducing hydrological features. The SWAT model also proves to be a good tool in such a modeling approach.

5.2. Recommendations

Generally from this specific study the following recommendations could improve similar research for future work:

- Although the temporal and spatial resolution may be higher, the different sources of errors in these datasets need further investigation and much more work is needed to that end. Nevertheless, the usefulness and suitability of applying these gridded products has been highlighted and it is promising that in areas where there is paucity of station observations, these gridded products can be used well for applications of rainfall runoff modeling.
- The other thing which is highly recommended is that the weather stations should be improved both in quality and quantity in order to improve the performance of the model. Hence, it is highly recommended to establish good meteorological stations and to have high quality of stream flow data.

6. References

1. Arnold, J. G. and Fohrer, N. (2005), SWAT 2000: current capabilities and research opportunities in applied watershed modelling. *Hydrol. Process.*19, 563–572, doi:10.1002/hyp.5611.
2. Arnold, J. G., Allen. P. M. and Bernhardt. G. (1993), A comprehensive surface-groundwater flow model. *Journal of Hydrology*, 142: 47-69.
3. Arnold, J.G., Srinivasan, R., Muttiah, R.S. and Williams, J.R. (1998), Large area hydrologic modeling and assessment, part I: Model development, *J. of American Water Resources Association*, 34, 11, 73-89.
4. Beven, K.J. (2006), A manifesto for the enquiringly thesis. *Journal of Hydrology* 320, 18–36.
5. Bitew, M. M., and M. Gebremichael. (2011), Assessment of satellite rainfall products for streamflow simulation. *Hydrol. Earth Syst.Sci.*, 15, 1147-1155.
6. Cunderlik, J. (2003), *Hydrological model selection for CFCAS project, Assessment of water resource risk and vulnerability to change in climate condition*, University of Western Ontario.
7. Danuso, F. (2002), Climak: A Stochastic Model for weather Data Generation. *Italian Journal of Agronomy* 6 (1): 57-7, 67-68pp.
8. Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., and Ropelewski, C. F. (2007), Validation of satellite rainfall products over East Africa’s complex topography, *Int. J. Remote Sens.*, 28(7), 1503–1526.
9. Dune, T. (1983), Relation of field studies and modeling in the prediction of storm runoff. *Journal of Hydrology*, Vol 65, 24-48.
10. Gassman, P. W., Reyes. M. R., Green. C. H., Arnold. J. G. (2007), The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Direction. *American Society of Agricultural and Biological Engineers* ISSN 0001-2351.
11. Gessese, A. (2008), *Prediction of sediment inflow for Legedadi reservoir using SWAT watershed and CCHEID sediment transport models*. Faculty of Technology, Addis Ababa University.
12. Green W. H. and Ampt G. A. (1911), Studies on soil physics. *J. of Agri. Scie.*, 4, 11-24.

13. Green, W.H. and Ampt, G.A. (1911), Studies in Soil Physics, I: The Flow of Air and Water through Soils. *Journal of Agricultural Sciences* 4: 1-24.
14. Hargreaves G. L., Hargreaves G. H. and Riley J. P. (1985), Agriculture benefits for Senegal River basin. *J. Irrig. and drain. Engr.*, 111, 2, 113-124.
15. Hirpa, F. A., Gebremichael, M., and Hopson, T. (2010), Evaluation of High Resolution Satellite Precipitation Products over Very Complex Terrain in Ethiopia, *J. Appl. Meteorol. Clim.*, 49(5), 1044–1051, doi:10.1175/2009JAMC2298.1.
16. Holvoet, K., van Griensven, A., Seuntjens P. and Vanrolleghem, P.A. (2004), Hydrodynamic modeling with SWAT for predicting dynamic behavior of pesticides. Water and Environment Management Series, *Young Scientist*, 211-21
17. Hsu, K.-l., X. Gao, S. Sorooshian, and H. V. Gupta. (1997), Precipitation estimation from remotely
18. Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B. (2007), The TRMM Multisatellite Precipitation Analysis (TMPA): Quasiglobal, multilayer, combined-sensor, precipitation estimates at fine scale, *J. Hydrometeorol.*, 8, 38–55. in medium watersheds of the Ethiopian highlands
19. Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P.P. CMORPH (2004), A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, 5, 487–503.
20. Ndomba, P. (2002), SWAT model application in a data scarce tropical complex catchment in Tanzania. *Physics and chemistry of the Earth*.
21. Neitsch, S.L., Arnold, J.G., Kiniry, J.R., and Williams, J.R. (2005), *Soil and Water Assessment Tool, Theoretical documentation: Version 2005*. Temple, TX. USDA Agricultural Research Service and Texas A & M Black land Research Centre.
22. O'Connell,P.E. (1991), Recent advances in the modeling of hydrological systems. Kluwer, Dordrecht, 3-30.
23. Priestley C.H.B and Taylor R.J. (1972), On the assessment of surface heat flux and evaporation using large scale parameters. *Mon. Weather Rev.*, 100, 81-92.
24. Ragan, R.M. (1968), An experimental investigation of partial area contributions. Int. Assoc. of Science Hydrology, No76, 241-251. Sensed Information using artificial neural networks. *Journal of Applied Meteorology* 36(9):1176-1190.

25. Santhi, C., Arnold, J.G., Williams, J.R., Dugas, W.A., Srinivasan, R. and Hauck, L.M. (2001), Validation of the SWAT model on a large river basin with point and nonpoint sources, *Journal of the American Water Resources Association*, Vol. 37, No. 5.
26. Setegn, S., Srinivasan, R., Dargahi, B. and Melesse, A. (2008), Spatial delineation of soil erosion vulnerability in the Lake Tana Basin, Ethiopia. *Hydrological Processes, Hydrol. Process.* (2009).
27. Sharpley, A.N. and Williams, J.R., eds. (1990), *EPIC - Erosion/Productivity Impact Calculator: 1. Model Documentation*. U. S. Department of Agriculture Technical Bulletin No. 1768. 235.
28. Soil Conservation Service. (1972), *Section 4 - Hydrology in National Engineering Handbook*. SCS.
29. Solomon (2013), Error in high- resolution Satellite Rainfall Production Streamflow Prediction in Birr Watershed, Ethiopia
30. Solomon, H. (2005), *GIS-Based Surface Runoff Modeling and Analysis of Contributing Factors; a Case study of the Nam Chun watershed, Thailand*. M.Sc Thesis, ITC, the Netherlands.
31. Sorooshian, S., K. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite (2000), Evaluation of PERSIANN system satellite based estimates of tropical rainfall, *Bull. Am. Meteorol. Soc.*, 81, 2035–2046.
32. Turk, F. J. and Miller, S. D. (2005), Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSRE blended data techniques, *IEEE T. Geosci. Remote*, 43, 1059– 1069.
33. Ushio, T. and Kachi, M. (2009), Kalman filtering applications for global satellite mapping of precipitation (GSMaP), in: *Satellite Rainfall Applications for Surface Hydrology*, edited by: Gebremichael, M. and Hossain, F., Springer-Verlag, New York.
34. Van Griensven A., Meixner T. (2004), Dealing with Unidentifiable Sources of Uncertainty within Environmental Models. *Hydrology Journal*.
35. Yilmaz, K. K., T. S. Hogue, K.-L. Hsu, S. Sorooshian, H. V. Gupta, and T. Wagener. (2005), Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting, *J. Hydrometeorol.*, 6, 497–517.
36. Ymeti. (2007), Rainfall estimation by Remote sensing for conceptual rainfall-runoff modeling in the Upper Blue Nile basin

APPENDICES

Appendix I List of tables

Comparison result of average rainfall magnitude over the watershed.

Table 1: Annual rainfall from various rainfall products

Date	Rain Gauge	3B42V7	CMORPH	PERSIANN
2003	1396.70	1583.67	1635.25	1339.38
2004	1513.54	1615.53	1829.34	1494.45
2005	1546.62	1862.42	2035.10	1656.83
2006	1727.90	1938.26	1949.00	1694.11
2007	1711.27	1635.69	1464.65	1400.94
Mean	1579.21	1727.11	1782.67	1517.14

Table 2: Total Seasonal rainfall (June - September)

Date	Rain Gauge	3B42V7	CMORPH	PERSIANN
2003	1213.50	1338.81	1371.34	1178.13
2004	1103.03	1284.53	1401.76	1068.89
2005	1150.38	1336.10	1552.43	1242.60
2006	1261.11	1346.33	1396.96	1174.69
2007	1189.48	1153.35	968.80	899.68
Mean	1183.50	1291.82	1338.26	1112.80

Table 3: Percentage of Seasonal rainfall

Date	Rain Gauge	3B42V7	CMORPH	PERSIANN
2003	86.88	84.54	83.86	87.96
2004	72.88	79.51	76.63	71.52
2005	74.38	71.74	76.28	75.00
2006	72.99	69.46	71.68	69.34
2007	69.51	70.51	66.15	64.22
Mean	75.33	75.15	74.92	73.61

Table 4: Mean monthly rainfall (mm)

Date	Rain gauge	3B42V7	CMORPH	PERSIANN
Jan	8.70	17.99	8.47	1.68
Feb	33.13	14	23.18	20.61
Mar	59.46	19.49	40.32	61.9
Apr	54.20	68.53	49.83	42.48
May	128.57	220.85	148.76	136.62
Jun	295.36	281.42	291.59	317.37
Jul	331.73	295.25	393.79	327.27
Aug	265.46	269.49	324.39	255.64
Sep	290.95	273.28	328.49	212.51
Oct	79.90	107.39	131.04	100.88
Nov	17.77	41.84	29.19	23.26
Dec	14.88	8.86	14.25	17.48

Table 5: Linear correlation coefficient (R) of daily rainfall

R	3B42V7	CMORPH	PERSIANN	Rain Gauge
3B42V7	1	0.89	0.89	0.71
CMORPH	0.89	1	0.84	0.69
PERSIANN	0.89	0.84	1	0.64
Rain Gauge	0.71	0.69	0.64	1

Table 6: Linear correlation coefficient (R) of monthly rainfall

R	3B42V7	CMORPH	PERSIANN	Rain Gauge
3B42V7	1	0.94	0.97	0.96
CMORPH	0.94	1	0.94	0.94
PERSIANN	0.97	0.94	1	0.95
Rain Gauge	0.96	0.94	0.95	1

Table 7: Linear correlation coefficient (R) of monthly streamflow simulation

R	Observed	Rain Gauge	3B42V7	CMORPH	PERSIANN
Observed	1	0.91	0.84	0.84	0.70
Rain Gauge	0.91	1	0.80	0.80	0.75
3B42V7	0.84	0.80	1	0.92	0.84
CMORPH	0.84	0.80	0.92	1	0.84
PERSIANN	0.70	0.75	0.84	0.84	1

Table 8: Performance statistics of model simulation based on various rainfall inputs

	Rainfall inputs	NSE	R²	RMSE	BIAS
Calibration result	Rain gauge	0.83	0.83	3.4	0.94
	3B42V7	0.66	0.7	4.27	0.98
	CMORPH	0.59	0.7	4.89	1.04
	PERSIANN	0.32	0.49	6.07	0.82
Validation result	Rain gauge	0.69	0.72	10.82	0.84
	3B42V7	0.38	0.57	18.1	0.54
	CMORPH	0.33	0.58	18.83	0.52
	PERSIANN	-0.21	0.37	20.46	0.25

Appendix II. List of figures

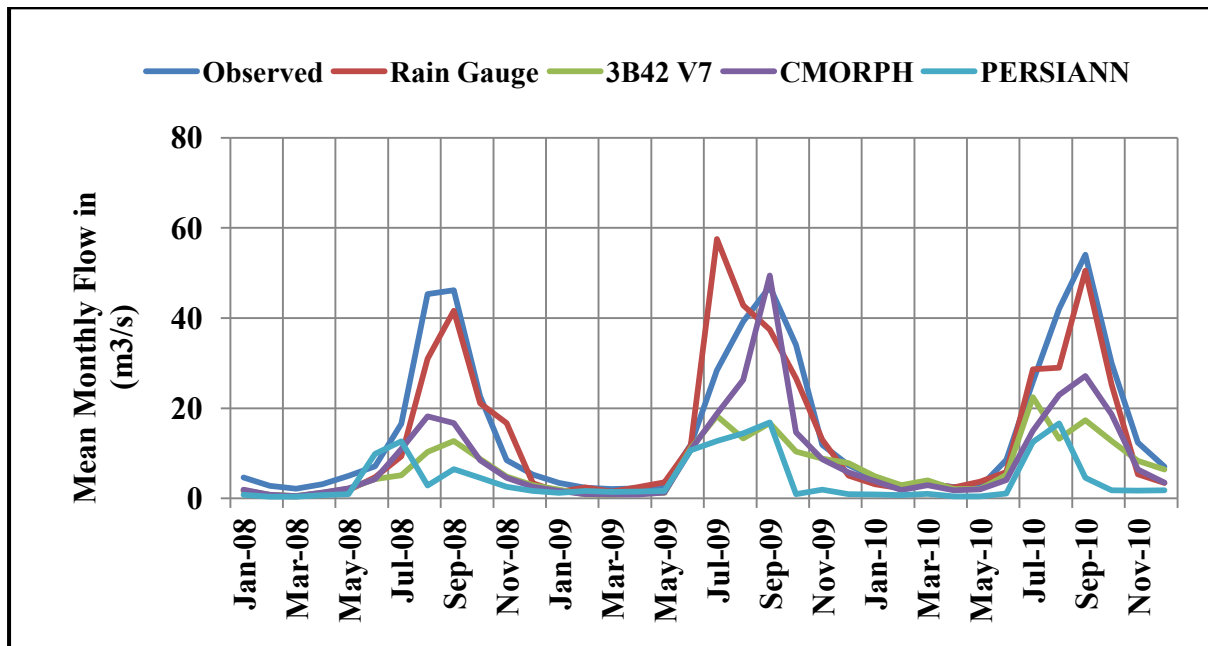
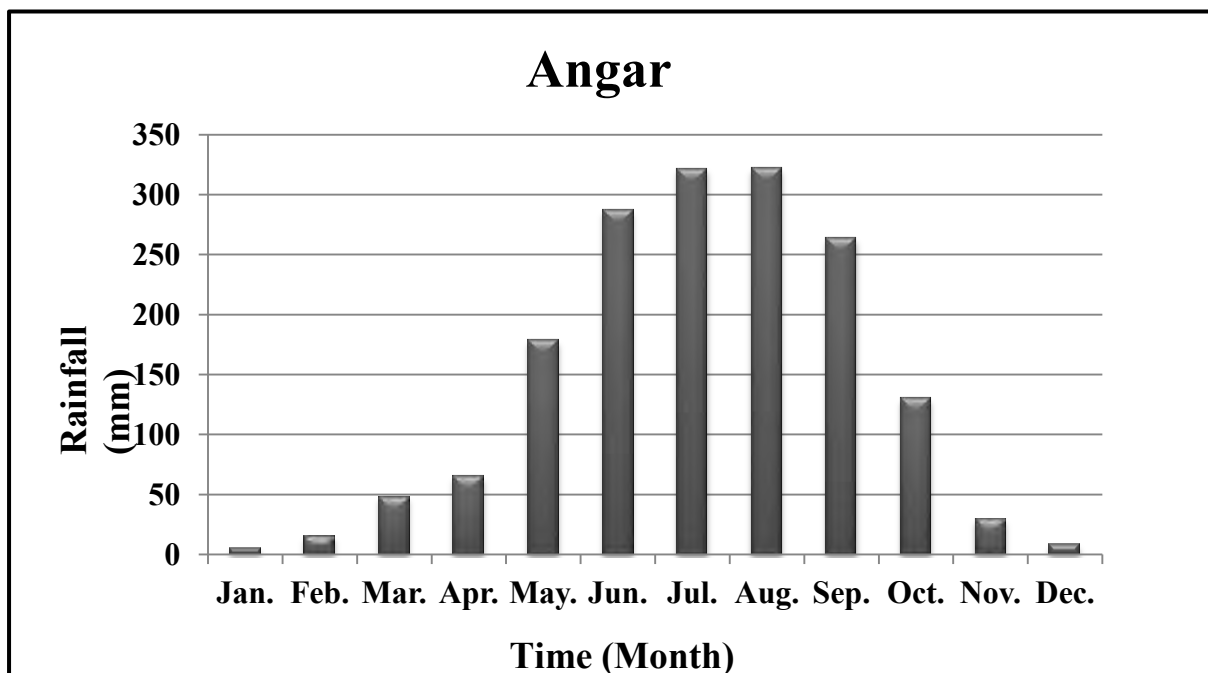


Figure 1 Comparison of Arc SWAT simulated stream flow (based on rainfall input data from various sources in the legend) and observed monthly stream flow during the validation period.



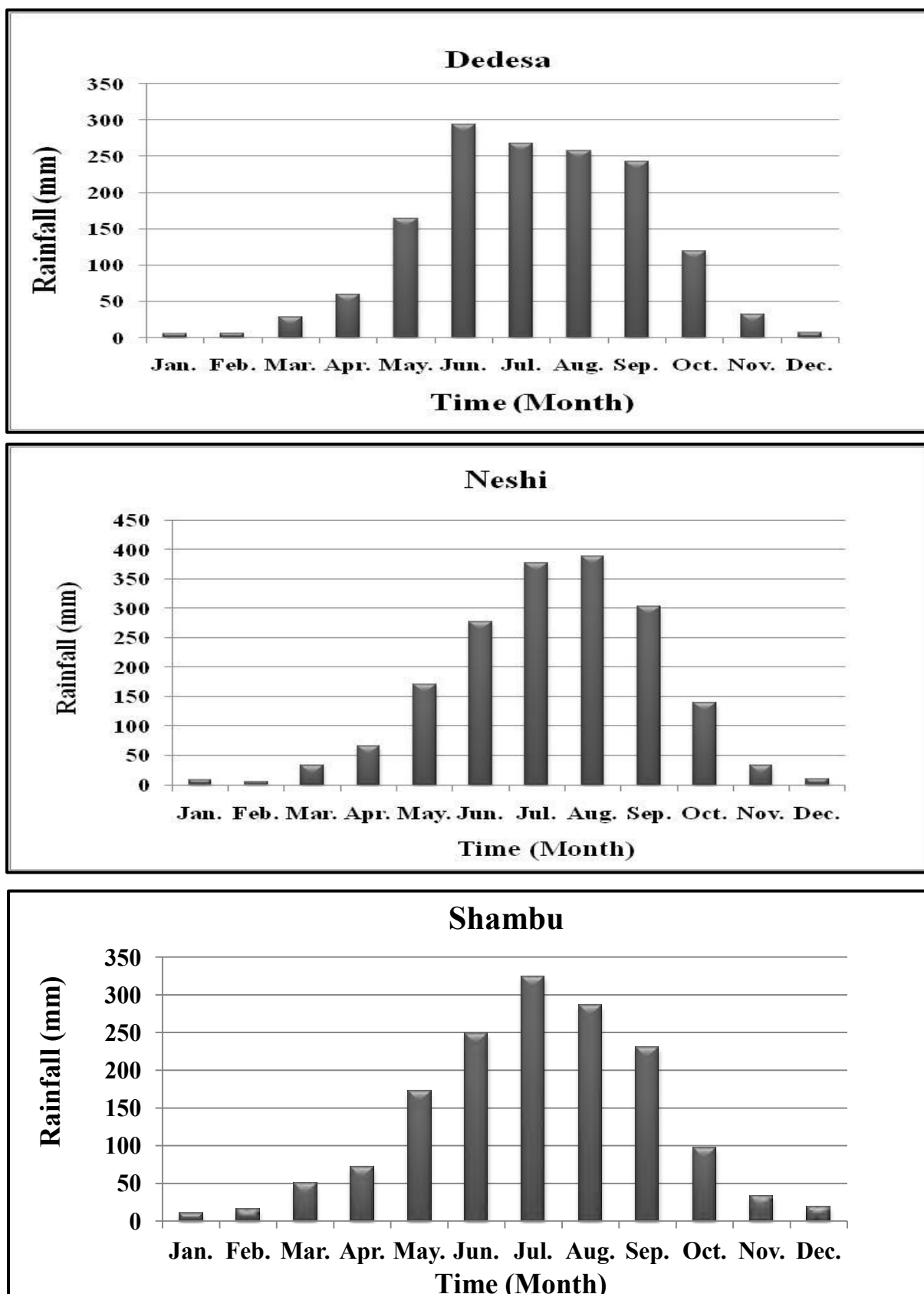


Figure 2 Mean monthly rainfall distribution of selected meteorological stations for the period of 1992 - 2012

Appendix III. Symbols and description of weather generator parameters (WGEN) used by the SWAT model

S. No	Symbol	Description
1	TMPMX	Average or mean daily maximum air temperature for month (o C).
2	TMPMN	Average or mean daily minimum air temperature for month (oC).
3	TMPSTDMX	Standard deviation for daily maximum air temperature for month (oC).
4	TMPSTDMN	Standard deviation for daily minimum air temperature for month (oC).
5	PCPMM	Average or mean total monthly precipitation (mm H2O).
6	PCPSTD	Standard deviation for daily precipitation for month (mm H2O/day).
7	PCPSKW	Skew coefficient for daily precipitation in month.
8	PR_W1	Probability of a wet day following a dry day in the month.
9	PR_W2	Probability of a wet day following a wet day in the month.
10	PCPD	Average number of days of precipitation in month.
11	SOLARAV	Average daily solar radation for month (MJ/m2/day).
12	DEWPT	Average daily dew point temperature in month (oC).
13	WNDVAV	Average daily wind speed in month (m/s).

Appendix IV. Soils parameter and legend used in SWAT model

S. No	Symbol	Description
1	NLAYERS	Number of layers in the soil (min 1 max 10)
2	HYDGRP	Soil hydrographic group (A, B, C, D)
3	SOL_ZMX	Maximum root depth of the soil profile (mm)
4	ANION_EXCL	Fraction of porosity from which an ions are exchanged (mm)
5	SOL_CRK	Crack volume potential of soil (%)
6	TEXTURE	Texture of the layer
7	SOIL_Z	Minimum depth from soil surface to bottom of layer (mm)
8	SOL_BD	Moist bulk density (g/cm^3)
9	SOL_AWC	Available water capacity of soil surface to bottom of the layer (mm/mm)
10	SOL_K	Saturated hydraulic conductivity (mm/hr)
11	SOL_CBN	Organic carbon content (%)
12	CLAY	Clay content (%)
13	SILT	Silt content (%)
14	SAND	Sand content (%)
15	ROCK	Rock fragmented content
16	SOL_ALB	Moist soil albedo
17	USLE_K	Soil erodibility factor (K)

Appendix V. Soils parameters values used in SWAT model

SNAM	CODE	NLAYERS	HYDRP	SOL_ZMX (mm)	ANION_EXCL (mm)	SOL_CRK (%)	TEXTURE	SOL_Z (mm)	SOL_BD (g/cm ³)	SOL_AWC	SOL_K (mm/hr)	SOL_CBN (%)	CLAY (%)	SILT (%)	SAND (%)	ROCK	SOL_ALB	USLE_K
Haplic Alisols	Alh	1	A	1800	0	0	C	400	1.10	0.08	0.50	3.00	45.00	31.00	24.00	4.00	0.14	0.23
		2					C	1000	1.36	0.08	0.30	0.15	62.00	22.00	16.00	5.00	0.14	0.23
		3					C	1800	1.36	0.08	0.30	1.00	70.00	17.00	13.00	7.00	0.14	0.14
Rhodic Nitisols	NTr	1	D	1800	0	0	C	300	1.10	0.08	0.50	2.00	57.00	28.00	15.00	4.00	0.14	0.23
		2					C	900	1.36	0.08	0.30	0.15	74.00	18.00	8.00	5.00	0.14	0.23
		3					C	1800	1.36	0.06	0.30	0.05	81.00	13.00	6.00	7.00	0.14	0.23
Dystric Leptosols	Lpd	1	B	350	0	0	SIC	350	1.20	0.08	0.50	2.00	51.00	40.00	9.00	4.00	0.13	0.20
Haplic Acrisols	Ach	1	D	1650	0	0	C	300	1.10	0.08	0.50	4.00	45.00	29.00	26.00	4.00	0.12	0.12
		2					C	900	1.36	0.08	0.30	0.15	79.00	15.00	6.00	5.00	0.11	0.18
		3					C	1650	1.36	0.08	0.30	0.05	88.00	10.00	2.00	7.00	0.14	0.11
Eutric Leptosols	Lpe	1	B	900	0	0	CL	200	1.41	0.08	0.50	2.00	39.00	35.00	26.00	4.00	0.15	0.25
		2					C	900	1.36	0.08	0.30	0.15	56.00	25.00	19.00	5.00	0.15	0.25
Haplic Nitisols	NTh	1	D	200	0	0	C	200	1.10	0.08	0.50	3.00	54.00	31.00	15.00	4.00	0.15	0.25
Eutric Vertisols	Vre	1	C	1800	0	0	C	300	1.10	0.08	0.50	2.00	68.00	25.00	7.00	4.00	0.13	0.20
		2					C	800	1.36	0.08	0.30	0.15	74.00	19.00	7.00	5.00	0.11	0.08
		3					C	1800	1.36	0.08	0.30	1.00	74.00	17.00	9.00	7.00	0.11	0.08

Appendix VI. Water balance parameters values of Angar watershed

Name	Parameter Description	Water balance parameter annual average values for Angar watershed			
		Rain gauge	3B42V7	CMORPH	PERSIANN
PRECIP	Rainfall (mm)	1579.2	1727.1	1782.7	1517.1
SQ	Surface runoff (mm)	626.24	603.98	703.86	414.64
LQ	Lateral soil (mm)	36.04	26.67	32.59	20.73
SHAL AQ	Ground water (mm)	82.07	63.65	85.66	52.14
REVAP	Evaporation from shallow aquifer, soil & plant (mm)	79.33	77.89	84.46	79.8
DEEP AQ RECH	Deep aquifer recharge (mm)	8.31	9.39	10.76	8.31
TOTAL AQ	Total aquifer recharge (mm)	166.2	187.78	215.26	166.17
TOTAL YLD	Total water yield (mm)	733.66	684.46	810.4	478.69
PERCOLATION	Percolation out of soil (mm)	155.9	178.14	203.51	157.33
ET	Actual evapotranspiration (mm)	746.2	903.5	824	913.8
PET	Potential evapotranspiration (mm)	1184.2	1144.5	1179	1257.7
TRAN. LOSSES	Transmission losses (mm)	10.69	9.84	11.71	8.82