



ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY

EVALUATION OF CLIMATE FORECAST AND SYSTEM REANALYSIS

SATELLITE WEATHER DATA FOR THE PREDICTION IN THE

UNGAUGED BASIN: CASE STUDY OF OMO-GIBE RIVER BASIN,

ETHIOPIA

A MASTER'S THESIS

BY

MINTIWAB ADDIS ALEMAYEHU

DEPARTMENT OF CIVIL ENGINEERING

COLLEGE OF ARCHITECTURE AND CIVIL

ENGINEERING

SEPTEMBER, 2021



ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY

**EVALUATION OF CLIMATE FORECAST AND SYSTEM REANALYSIS
SATELLITE WEATHER DATA FOR THE PREDICTION IN THE
UNGAUGED BASIN: CASE STUDY OF OMO-GIBE RIVER BASIN,
ETHIOPIA**

BY

MINTIWAB ADDIS ALEMAYEHU

A Thesis Submitted as a Partial Fulfillment to the Requirements for the
Award of the Degree of Master of Science in Civil Engineering
(Hydraulic Engineering)

TO

**DEPARTMENT OF CIVIL ENGINEERING
COLLEGE OF ARCHITECTURE AND CIVIL ENGINEERING**

SEPTEMBER, 2021

DECLARATION

I hereby declare that this thesis entitled "Evaluation of Climate Forecast and System Reanalysis (CFSR) Satellite Weather Data for Hydrological Models, the case of OMO-GIBE River Basin, Ethiopia" was done by myself, with the guidance of my advisor, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted, in whole or in part, for any other degree or professional qualification.

Name: Mintiwab Addis Signature,  Date 12/09/2021

Witness

Advisor : Dr. Adanech Yared.....

Signature 

date16/09/2021

APPROVAL

This MSc thesis entitled "Evaluation of Climate Forecast and System Reanalysis (CFSR) Satellite Weather Input Data for Hydrological Modelling, the case of Omo-Gibe River Basin, Ethiopia" has been approved by the following examiners after the thesis presentation for the Master of Science Degree in Hydraulic Engineering.

Signed by Examining Board:

Adanech Yared Tillo (PhD) [Signature] 16/09/2021
Advisor: Signature Date

Dr. Brook A [Signature] 17/09/2021
Internal Examiner: Signature Date

Dr. Sirak Tekleab [Signature] 16/09/2021
External Examiner: Signature Date

Yohannes Haagos [Signature] 17/09/2021
Chairperson: Signature Date

Abenzer T. [Signature] 17/09/2021
DGC chairperson: Signature Date

Sisay Demaku (PhD) [Signature] 17/09/2021
College Dean/ Associate Dean for GP: Signature Date



ABSTRACT

Hydrological data inadequacy at ungauged watersheds has been a problem for modeling the water resources of Omo-Gibe Basin, Ethiopia. Climate Forecast and System Reanalysis (CFSR) Satellite climate products can be utilized as an alternative source of climate data in such regions where the usual ground observed climate data are insufficient. Therefore, the aim of this study was to evaluate the hydrological performance of CFSR weather data as an input to the HBV-Light hydrological model over the Omo-Gibe River Basin, Ethiopia. The evaluation was executed in three different steps. Initially, the CFSR data was compared to ground climate data using efficiency descriptors (ENS , R^2 , $PBIAS$) and categorical statistics (Probability of Detection and False Alarm Ratio) for the same time window of 1987 to 2010. Secondly, the runoff predictive performance of the CFSR driven HBV -Light model was calibrated and validated at gauged catchments. Finally, streamflow was generated at ungauged catchments through the method of regionalization. The soil descriptors (Soil permeability, Soil available water content, Soil bulk density), vegetation descriptors (NDVI) and topographical descriptors were some of the data used in this study. Prior to the regionalization, Principal component analysis (PCA) was carried out on physical catchment characteristics (PCCs) to improve the regression model between the model parameters and PCCs. The result of direct comparison revealed that CFSR correlated to ground observed rainfall data at catchments with an efficiency of (R^2 , $NSE \geq 0.7$) except for Walga, Wabe, Gorombo, and Sokie-Weybo where CFSR have shown reasonable efficiency with an acceptable bias of 10%, 8%, and -11%. However, the CFSR efficiency was increased after bias correction applying the linear scaling bias correction method. CFSR data-driven rainfall-runoff model has shown a good performance (R^2 , $NSE \geq 0.7$) in simulating the observed flow at well-recorded catchments. The performance of the regional model verified in the validation catchments was found satisfactory and recommended to be modified introducing any dataset not included in this study like anthropogenic activities, geological and mining data.

Keywords/phrases; CFSR, HBV-light, Regionalization, Ungauged catchments, Omo-Gibe

ACKNOWLEDGMENTS

First of all, my precious honor towards almighty **GOD** who always help me and become my strength throughout my life.

I am greatly thankful to my respected Advisor Dr. Adanech.Yared for her supervision, support, and encouragement. I would also want to thank Mr. Nigussie for his friendly support in sharing HBV-Light software. Finally, I would like to express my genuine gratitude and appreciation to Ambo University for giving me this chance and sponsoring me to specialize. Likewise, I am greatly thankful to individuals and organizations who contributed directly or indirectly to this research work.

TABLE OF CONTENTS

DECLARATION.....	II
APPROVAL	III
ABSTRACT	IV
ACKNOWLEDGMENTS	V
TABLE OF CONTENTS	VI
LIST OF TABLES	IX
LIST OF FIGURES	X
LIST OF SYMBOLS AND ABBREVIATIONS	XI
CHAPTER 1 INTRODUCTION	13
1.1 Background	13
1.2 Statement of the Problem	15
1.3 The objective of the research	16
1.3.1 General objective	16
1.3.2 Specific objectives	16
1.4 Research questions	16
1.5 Scope of study	16
1.6 Significance of the study	17
CHAPTER 2 LITERATURE REVIEW	18
2.1 Satellite Rainfall Products (SRP)	18
2.2 Satellite-Based Rainfall Estimation Methods	19
2.3 Review on some of Satellite Rainfall Products (SRP)	19
2.3.1 PERSIANN-CDR satellite rainfall products	19
2.3.2 TRMM-3B42v7 Rainfall Products	20
2.4 Global Re-analysis Weather Data	20
2.4.1 Climate Forecast System Reanalysis (CFSR)	21
2.4.2 Bias correction for CFSR data	22
2.4.3 Previous study on satellite weather data	23
2.5 Hydrological models	24

2.5.1	Historical developments of hydrological models	24
2.5.2	Hydrological Model selection	25
2.5.3	Classification of hydrological models	25
2.5.4	Previous study with HBV-light model	30
2.5.5	Previous study with HBV-light model in Ethiopia	30
2.6	Ungauged catchments	31
2.7	Regionalization.....	32
2.7.1	Regionalization approaches	32
2.7.2	Regional model development	33
2.7.3	Principal component analysis (PCA)	33
CHAPTER 3	MATERIALS AND METHODS	35
3.1	Description of the Study Area	35
3.1.1	Location	35
3.1.2	Climate	36
3.1.3	Topography	36
3.1.4	Land use	37
3.1.5	Drainage sub-basins and River system of Omo-Gibe River Basin	37
3.2	Data collected and materials used in the study	37
3.2.1	Data collected.....	37
3.2.2	Materials used	44
3.3	Data analysis	44
3.3.1	Rainfall data quality check	44
3.3.2	Evaluation of CFSR for Orographic effect	48
3.3.3	Bias correction for CFSR data	49
3.3.4	Comparison of areal CFSR data to areal ground observed data	49
3.3.5	Evaluation of rainfall detection capacity of daily CFSR data	51
3.3.6	Description of HBV-Light model	52
3.3.7	HBV-Light model components and structures	52

3.3.8	HBV-Light Model Input data setup	54
3.3.9	Sensitivity analysis of CFSR and meteorological data	55
3.3.10	Model Calibration	56
3.4	Model Performance Evaluation	57
3.4.1	Physical Catchment Characteristics (PCCs)	57
3.4.2	Principal component analysis (PCA)	58
3.5	Regionalization.....	61
3.5.1	Developing Regional Regression Equation	61
CHAPTER 4	RESULTS AND DISCUSSION	63
4.1	Comparison of the areal CFSR and ground observed datasets	63
4.2	Runoff simulation results in the Omo-Gibe River Basin	66
4.2.1	Sensitivity analysis of HBV-Light model	66
4.2.2	HBV-Light model Daily calibration and validation results	69
4.3	The established Regional Model	72
4.4	Validation of the Regional Model	74
4.5	Flow estimation at the ungauged catchments	75
CHAPTER 5	CONCLUSIONS AND RECOMMENDATIONS	77
5.1	Conclusions	77
5.2	Recommendations	79
REFERENCES	81

LIST OF TABLES

Table 3.1: Rain detection capability and False Alarming Ratio	51
Table 3.2: Parameters range used in the sensitivity analysis	56
Table 3.3: Physical Catchment Characteristics (PCCs) used in the regionalization...	58
Table 3.4: KMO and Bartlett's Test	59
Table 3.5: Eigen analysis of the Correlation Matrix	61
Table 4.1: The comparison using statistical metrics	65
Table 4.2: The comparison using categorical metrics	65
Table 4.3: Optimized parameter values at gauged catchments (1989-2006).	68
Table 4.4: Optimized model parameter from CFSR data	68
Table 4.5: Result Summary of daily flow simulations using HBV model	71
Table 4.6: Summary of HBV model daily simulation result	71
Table 4.7: The regional model developed for MPs and PCCs links.	73
Table 4.8: Model parameters predicted at the gauged catchments	75
Table 4.9: Validation of the regionalized model at validation catchments.....	75
Table 4.10: Model parameters derived for ungauged catchments	76

LIST OF FIGURES

Figure 3.1: Location map of study area	35
Figure 3.2: Mean monthly rainfall of some stations (1987 -2010).	36
Figure 3.3: Meteorological stations and River system of Omo-Gibe Basin	38
Figure 3.4 : Distribution of CFSR Grid data at Omo-Gibe Basin	39
Figure 3.5 Mean Annual Rainfall and Mean Annual Evapotranspiration map	40
Figure 3.6: NDVI map of the Omo-Gibe River Basin	41
Figure 3.7: Soil Permeability (SAT_K) and Bulk Density (SOL_BD) map	42
Figure 3.8 Soil Available Water Content map of the Basin	43
Figure 3.9: Homogeneity test for meteorological stations	45
Figure 3.10: Consistency of rainfall data in the stations	46
Figure 3.11: Monthly Average Discharges at a some of gauging stations	48
Figure 3.12: Evaluation of CFSR data for Orographic effect	48
Figure 3.13: Schematic structure of HBV-Light model (Seibert, 2005)	52
Figure 3.14: Group of homogenous catchments classified using PCA	60
Figure 4.1: Monthly comparisons of the rainfall datasets.....	64
Figure 4.3: Sensitivity analysis	67
Figure 4.4: Daily calibration @ Tunjo sub-catchment (1992 -2004)	69
Figure 4.5 Daily validation of CFSR data at Tunjo (2005 -2010)	70
Figure 4.6: Comparison of the stream flow simulation	74
Figure 4.6: Hydrograph at the Sharma catchment (Ungauged catchment)	76

LIST OF SYMBOLS AND ABBREVIATIONS

ARS	Agricultural Research Service
AWC	Available Water content
BETA	Parameter in Soil Moisture Routine in the HBV Model
CFSR	Climate Forecast System Reanalysis
CPC	Climate Prediction Center
DA	Drainage Area
DEM	Digital Elevation Model
ET _o	Reference Evapotranspiration
FAO	Food & Agriculture Organization
FAR	False Alarm Ratio
FA	Flow Accumulation
FC	Parameters defining maximum soil moisture storage
FD	Flow Direction
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement
GIS	Geographic Information System
HBV	Hydrological Byråns Vattenalansavdelning
LLP	Length of longest flow Path
ME	Mean Elevation
MAPE	Mean Annual Potential evapotranspiration
MAR	Mean Annual Rainfall
NDVI	Normalized Difference Vegetation Index
MPEF	Meteorological Product Extraction Facility (),
MPEG	Multi-Sensor Precipitation Estimate–Geostationary

MoWIE	Ministry of Water, Irrigation and Electricity
NMSA	Ethiopian National Meteorological Services Agency
NASA	National Aeronautics and Space Administration
NRLB	Naval Research Laboratory's blended product
NCEP	National Centers for Environmental Protection
NOAA	National Oceanographic, Atmospheric Administration
NDVI	Normalized Difference Vegetation Index
NSE	Nash-Sutcliffe efficiency
PCA	Principal Component Analysis
PCC	Physical catchment characteristics
POD	Probability of Detection
PBIAS	Percent bias
R ²	Coefficient of Determination
SB_D	Soil Bulk Density
SAT_K	Soil permeability
TWI	Topographic Wetness Index
TIROS-1	Television Infra-Red Observation Satellite
TRMM	Tropical Rainfall Measuring Mission

CHAPTER 1 INTRODUCTION

1.1 Background

Climate data is the most critical parameter for watershed modeling in the water cycle, improving the short term, medium, and long-term weather forecasts, and climate monitoring. Effective planning and utilization of water resources require sufficient climate data, but these data are usually obtained from rain-gauge stations of *in situ* monitoring systems which are expensive to preserve in Africa due to complex topography, economy, and deserts as the measurement of precipitation are difficult in areas with poor road networks (Pombo & de Oliveira, 2015).

A typical difficulty in modeling watershed hydrology is obtaining accurate climate input data, as the network of observation stations for rainfall is sparse and unevenly distributed in developing countries. Land-based climate 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 (Fuka et al., 2013). Rainfall data is a crucial resource in many socio-economic activities, particularly for those African countries which are prevalently dependent on the rainfed agriculture that took care of agribusiness (Worqlul *et al.*, 2015).

Recent advances in multi-satellite rainfall estimates have allowed the use of high-resolution satellite rainfall products in hydrological modeling for runoff simulations and predictions. Those high-resolution satellite climate data are sufficiently available and attracting the hydrologists' attention especially in developing countries where climate radars are missing and routine meteorological stations are inadequate with short time duration data (Bitew & Gebremichael, 2011).

By contributing 90% of the runoff to lake Turkana (located downstream in Kenya), Omo-Gibe River Basin is the second largest river system in Ethiopia (Jillo et al., 2017). However, representative hydrometeorological data in such a large basin of river system for hydrological modeling has been an active problem to the limited availability of *in situ* observation stations. The high-resolution globally available CFSR weather forecasts blended with ground meteorological data have the potential to overcome the above-mentioned problem.

Climate Forecast System Reanalysis (CFSR) is the new dataset freely downloadable from the National Center of Environmental Prediction (NCEP) and currently covers from 1979 to the present, at a horizontal resolution of approximately (38 km) with 64 vertical layers. CFSR data are planned, accomplished, and released as worldwide high-resolution joining atmosphere-ocean-land surface-sea ice structure to supply the most excellent gauge of the state of these coupled spaces and all-inclusive accessible for each hour. Results showed that utilizing the CFSR precipitation to force a watershed model provides stream discharge simulations that are better than models forced using traditional weather gauging stations.

The harmonious application of these data sets would help us to (a) input precipitation and temperature and vegetation information into a hydrological model and calculate the soil moisture and surface evapotranspiration using the water and energy balance equations (b) the measured soil moisture and surface temperature can be used in two ways (i) to calibrate certain model parameters (ii) to verify the output of the model through validation (Fuka et al., 2013).

Even though Satellite Rainfall Products are serving as a substitution for scarce climate data, they are sometimes subjected to a variety of potential errors, which originate from the discontinuous revisit time of observing sensors and weak relationships between remotely sensed signals and rainfall rate. Recent Researchers Habib *et al.*(2012), Shrestha *et al.*(2017) indicated that satellite-based rainfall estimates are sometimes not reliable and consensus as the estimates require correction. Orographic rainfall is one of the common rainfall types in the study basin and happens due to cloud formation at the mountain face with the lifting of moist carrying wind and air mass blowing from waterbodies. According to Dinku *et al.* (2008), this orographic cloud may lead to precipitation while the cloud top is still relatively warm and may not be detected by the satellite sensors. Hence to certify a justified utilization of Satellite Rainfall Product (SRP) a systematic validation is required before using for hydrological modeling. There are two methods for validating SRP: either by comparing to ground observed rain gauge data, or through hydrological modeling using the satellite data as an input (Worqlul *et al.*, 2014).

Therefore, the purpose of this study was to evaluate the performance of CFSR data in three ways: First, direct comparison of areal average CFSR data to its areal average observed rainfall of each catchment in the basin. Second, evaluation of CFSR through hydrological modeling using HBV-Light model at the gauged catchments, thirdly,

prediction of flow at ungauged catchments and so as to fill the gap of data scarcity in Omo-Gibe River Basin, Ethiopia.

1.2 Statement of the Problem

Ethiopia is a large country and known to be a water tower of East Africa but most River Basins have less coverage of hydro-meteorological gauging stations. Omo-Gibe River Basin is one of those basins with ungauged sub-watersheds at its north western and southern parts.

In Omo-Gibe River Basin, the ground observed hydrometeorological gauging stations are poorly recorded and long-time data from only a few well-documented weather observations have constrained efficient planning and management of water resources and hampers early flood warning properly as they are rare and even nonexistent. As there is no complete information of inflow from each catchment and total outflow from the basin together with uneven spatial and temporal distribution of climate variables also a problem for modeling water resources in the Basin. However, the basin has significant economic, social, and ecological importance within and neighboring countries.

World meteorological observation (WMO, 2003) reported that African rain-gauge density is 8 times lower than (WMO) recommendation. The Omo-Gibe River Basin is also rich in multipurpose water resources development raising the economy of the country such as cascade hydropower plants. These plants are continuously supplied by the rivers at upstream subsequent watersheds along with V-shaped valley reaches and suitable topographies.

However, In the lower Omo-Gibe Basin, the non-existent of hydro-meteorological stations has initiated to search for another alternative source of data for sustainable use of these water resources in the basin. The general lack of in situ precipitation measurements in the southern part of study area necessitated the use of CFSR data, modeling the hydrology and its features using satellite-generated climate data as a meteorological data gap filling after evaluating its hydrological performance. This study was thus intended to in depth evaluation of CFSR data starting from pair wise comparisons, to hydrological modeling through stream flow simulation.

Therefore, the main objective of this study was aimed to evaluate the Climate Forecast and System Reanalysis (CFSR) for use as an input in HBV -Light hydrological modeling in the catchments of the Omo-Gibe River Basin, Ethiopia.

1.3 The objective of the research

1.3.1 General objective

The main objective of this study was to compare CFSR data and climate data ground observed rainfall data for flow prediction in ungauged watersheds of Omo-Gibe River Basin, Ethiopia.

1.3.2 Specific objectives

1. To compare (CFSR) data against ground observed rainfall data
2. To evaluate the performance of CFSR data through streamflow simulation using the HBV-Light hydrological model.
3. To understand the rainfall- runoff processes in gauged watersheds

1.4 Research questions

1. Is there any significant difference between CFSR climate data and ground observed climate data?
2. Are CFSR rainfall products effective in simulating stream flows in calibration to HBV-Light hydrological model for large-scale watersheds?
3. Are CFSR data-driven regionalized HBV model parameters able to predict streamflow at ungauged catchments?

1.5 Scope of study

The domain of this study was Omo-Gibe River Basin. For this study, the scope was evaluating and assessing the hydrological performance of CFSR against ground observed meteorological data after calibrating and validating at gauged catchments against the observed flow, and therefore, applying the CFSR data for ungauged

catchments to generate stream flow in the basin was the main target of this study. The HBV hydrological model was used.

1.6 Significance of the study

This research will help other experts, initiatives, and researchers in the areas of water resources planning and management work by providing relevant information on satellite generated and reanalyzed climate data (CFSR) for hydrological modeling that further helps for feasibility studies, flood management including early warning systems, design of best management practices to control contaminant runoff, sizing of hydraulic structures for extreme events, obtaining the streamflow at the outlet of ungauged catchments in data-limited basins as an alternative data source.

It is supportive to climate modelers and end-users for understanding the errors and vulnerabilities expected and how they broadcast in hydrological modeling. Gives information on how to access satellite data to in-filling missed & predicting values of daily precipitation & temperature data for ungauged catchments. Reliable continuous streamflow forecasting is an important factor in watershed planning and sustainable water resource management, as it is instrumental in obtaining a deeper sense of flow variability in ungauged basins. Furthermore, the estimation of flood peaks, low flow characteristics, and FDCs can be derived once the synthetic continuous flow time series is generated. Streamflow data is also used in the design of critical engineering structures such as highways, drainage systems, reservoirs, etc.

CHAPTER 2 LITERATURE REVIEW

2.1 Satellite Rainfall Products (SRP)

In 1960, satellite-based remote sensing was launched from the first weather and earth observant technology known as meteorological satellite Television Infra-Red Observation Satellite (TIROS-1). the study of the earth's atmosphere and oceans using data obtained from these remote sensing devices has advanced rapidly. In the 1970s rainfall estimation using Infra-Red (IR) sensors was launched on geostationary platforms to track cloud movement and advance climate and weather prediction (Tan et al., 2018).

The primary scope of weather satellites monitoring is to provide information on climate occurrence, amount, and spatial distribution over the globe continuously from all areas including those inaccessible to gauges and radars for various applications in meteorology, climatology, hydrology, and environmental sciences (Shrestha *et al.*, 2017).

These weather satellites have been serving as a source for radiation signals reflected to or emitted from the ground and atmosphere. This estimation is also a method of deriving qualitative (image) and quantitative (value) using visible and infrared techniques through indirect relationships between solar radiance reflected by clouds, that is, the clouds brightness temperature and rainfall. InfraRed (IR) precipitation estimates are created using the microwave estimates for calibration microwave and IR estimates are combined to fill the gaps; and (4) the data are rescaled to monthly totals where by gauge observations are also used indirectly to remove bias (Berhanu et al., 2015).

Nowadays, Satellite-derived rainfall estimates have become a powerful tool for supplementing ground-based rainfall estimates. Recently, Earth observation data for environmental or societal purposes have become readily available through Earth observation (EO) satellites and data distribution systems. Some of the freely available spatially distributed rainfall estimates are the Tropical Rainfall Measuring Mission (TRMM), EUMETSAT's Meteorological Product Extraction Facility (MPEF), Multi-Sensor Precipitation Estimate–Geostationary (MPEG), the Climate Forecast System Reanalysis (CFSR), the NOAA/Climate Prediction Center morphing technique

(CMORPH), precipitation estimation from remotely sensed information using artificial neural network (PERSIANN), the Naval Research Laboratory's blended product (NRLB), and more (Worqlul et al., 2014).

Polar-orbiting satellites travel in a circular orbit from pole to pole orbiting at an altitude of about 800 km and use microwave (MW) channels (Shrestha et al., 2017). Orbits of these satellites are such that they pass the equator at the same local time on each orbit, providing about two overpasses each day. Some Satellite rainfall products such as TRMM, CMORPH, PERSIANN with different spatial and temporal resolutions have emerged as an alternative or supplement to conventional precipitation observations (Habib et al., 2012).

2.2 Satellite-Based Rainfall Estimation Methods

Satellites can be classified based on their purposes. Rainfall estimating satellites are primarily meteorological satellites known to be Geostationary satellites and Polar-orbiting satellites. Geostationary satellites provide continuous observation of the earth's surface and provide data on half-hourly or even lesser durations. Imageries obtained from these satellites are mainly visible (VIS) and IR at a resolution of about 4 km, with information on clouds, collected once every 15- 30 minutes. There are several operational geostationary meteorological satellites in orbit such as the (MTSAT, GOES, METEOSAT, FY series, and INSAT).

2.3 Review on some of Satellite Rainfall Products (SRP)

2.3.1 PERSIANN-CDR satellite rainfall products

Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks- Climate Data Record provides daily rainfall estimates at a spatial resolution of 0.25 degrees in the latitude band 60°S -N from 1983 to the near-present. The precipitation estimate is produced using the PERSIANN algorithm on GridSat-B1 infrared satellite data, and the training of the artificial neural network is done using the National Centers for Environmental Prediction (NCEP) stage IV hourly precipitation data. PERSIANN-CCS performs fairly well under cold clouds ($T_b < 253K$), extending the algorithm's capability to capture warmer rainfall requires more investigation. Employing multi-spectral images (Pombo et al., 2015).

2.3.2 TRMM-3B42v7 Rainfall Products

The Tropical Rainfall Measuring Mission (TRMM), a joint US-Japan satellite mission, was launched in 1997 to monitor tropical and subtropical precipitation and to estimate its associated latent heating covering the latitude band 50 °N-S. The TRMM satellite carries five rain-measuring instruments: Microwave Imager (TMI), the Visible Infrared Scanner (VIRS), the Lightning Imaging Sensor (LIS), the Clouds and Earth 's Radiant Energy System (CERES), and the Precipitation Radar (PR) which is the first spaceborne precipitation radar (Huffman et al., 2007). Estimates are provided at relatively fine scales (0.25°×0.25°,3-h) in both real and post-real time to accommodate a wide range of research applications. However, the errors inherent in the finest scale estimates are large (NCAR, 2018).

2.4 Global Re-analysis Weather Data

Re-analysis data are multilayer global gridded representations of weather using a systematic approach of synthesizing blending both observational and numerical model data, spanning an array of climate variables to produce datasets for climate monitoring and research. These data are created via unchanging data assimilation schemes and models which inject all available observations every 6-12 hours over the period being analyzed (Saha et al., 2010). The combination of data and modeled fields provides scientists with multi-decadal information on weather patterns, climate variability, and change variables from modeled surface fields, such as precipitation, evaporation, and radiation fluxes, to components of the stratosphere (Fuka et al., 2014).

Global reanalysis of weather data provided by the United States and Europe is currently used for various hydrological applications around the world. some of them are National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR); the NCEP and US Department of Energy (DOE) NCEP/DOE; the NCEP and the National Center of Atmospheric Research (NCAR) NCEP/NCAR; the European Centre for Medium-Range Weather Forecasts (ECMWF) RA-15/40; Japan Meteorological Agency (JMA) JRA-25; and National Aeronautics and Space Administration (NASA) (HU et al., 2017).

2.4.1 Climate Forecast System Reanalysis (CFSR)

The CFSR is a third-generation globally available reanalysis product and high resolution, blending atmosphere-ocean-land surface-sea ice system to provide the best estimate of the state of these coupled domains. The CFSR has also been extended as an operational, real-time product into the future with appealing strengths include (1) joining of atmosphere and ocean during the prediction of the 6-hour guess field; (2) an interactive sea-ice model; and (3) assimilation of satellite radiances by the Grid-point Statistical Interpolation (GSI) scheme over the entire period (National oceanic and atmospheric administration,1979-2011). The CFSR atmospheric spatial resolution is approximately 38 km with 64 levels extending from the surface to 0.26 hPa. Satellite-based radiance observations were bias-corrected with spin-up runs at full resolution, considering variable CO₂ concentrations. This procedure enabled smooth transitions of the observation record due to evolutionary changes in satellite observing systems (Penman, 2020).

CFSR data consider the alteration in CO₂, aerosols, and trace gas variations in the atmospheric model and assimilate the radiance measurements from a series of National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites (Zhang et al., 2020).

The CFSR atmospheric, oceanic, and land surface output products are available at an hourly time resolution and a 0.312 x 0.312-degree latitude and longitude resolution (NOAA). Application importance of CFSR dataset NCEP climate forecast system reanalysis (CFSR) data set for small to medium-sized watershed can qualify and be the best choice with the following criteria (Fuka et al., 2014).

- (1) the dataset should be open and available, including all climate variables.
- (2) a spatial resolution needs to be 38 km; and
- (3) the length of records should include adequate historical coverage to allow model calibration and validation, and extend to the present.

Weather is often monitored at a location outside the watershed to be modeled, sometimes at a long distance from the watershed as a result, the available records may not meaningfully represent the weather occurring over the watershed. An additional complication is that rain gauge data effectively point measurements which may represent precipitation poorly across a watershed particularly if there are high hydroclimatic gradients (Fuka et al., 2013).

CFSR can be utilized to substitute precipitation inputs in hydrological modeling. These data have their advantages when determining different forms of precipitation such as snow, hail, rainfall, and the appropriate relationship between radar reflectivity and rain rate, thus, there is no need to consider the additional method to estimate weather conditions for watershed-scale modeling, one possibility is the use of multilayer global gridded representations of weather known as reanalysis data sets.

2.4.2 Bias correction for CFSR data

Satellite rainfall products may not detect the rainfall from the warm clouds as the cloud-tops would be too warm for IR thresholds, and there will not be much ice aloft to be detected by PM sensors (Habib *et al.*, 2012).

National Oceanic and Atmospheric Administration's (NOAA) is a popular data source for satellite-based rainfall estimates that uses multiple passive microwave (PMW) satellites and Thermal infrared (TIR) images for estimation of precipitation and these satellite products are mostly prone to errors as they are estimated from secondary sources (for instance, cloud top brightness temperature) are subjected to a bias of under or overestimation of the rainfall event when compared with Insitu rainfall data. For example, A TIR sensor provides useful information on storm clouds based on top cloud temperature by assuming that relatively cold clouds are associated with thick and high clouds that tend to be associated with the production of high rainfall rate underestimates warm rain and misidentifies cirrus clouds as rain (Romilly & Gebremichael, 2011). Passive Microwave sensors are not available on geostationary satellites, which makes them have longer latency.

Worqlul *et al.* (2017) evaluated CFSR for Orographic and convective effects by plotting the long-term annual average CFSR rainfall against station elevation to see the rainfall–elevation relation. Two clear trends were observed. The first one shows a 50 mm of rainfall increase for every 100 m elevation increase and the second trend observed was a 125 mm rainfall increase for every 100 m elevation increase. These two relations can be explained by stations likely affected by a combination of orographic and convective rainfall. Orographic rainfall is one of the common rainfall types in the study basin and happens due to cloud formation at the mountain face with the lifting of moist carrying wind and air mass blowing from waterbodies. According to Dinku *et al.* (2008), this orographic cloud may lead to precipitation while the cloud top is still

relatively warm and may not be detected by satellite sensors. Hence bias correction of CFSR should be carried out for the study basin.

2.4.3 Previous study on satellite weather data

Dile & Srinivasan (2014) evaluated CFSR data for hydrologic modeling in data-scarce watersheds of the Blue Nile River Basin, Ethiopia. The Soil and Water Assessment Tool was used to evaluate the performance of CFSR weather with that of conventional weather in simulating observed streamflow at four river gauging stations in the Lake Tana basin. Dile & Srinivasan (2014) suggested that the use of CFSR data set without model calibration showed reasonable performance at Gilgel Abay and Gumera River gauging stations at a monthly time step using two performance indicators, Nash–Sutcliffe Efficiency (NSE) and Percentage Bias (PBIAS) coefficients. The result in an NSE value of more than 0.75 showed the very good performance of the model in these gauging stations. The PBIAS value for Gilgel Abay also indicated very good performance, while the PBIAS value for the Gumera showed good model performance.

Vu Thi Thom (2018) evaluated four gridded rainfall products, including CFSR, APHRODITE, PERSIANN, and TRMM, as input to the SWAT model to simulate streamflow against rain-gauge data over the Srepok River Catchment in Vietnam. Amongst the four different datasets, the TRMM and APHRODITE data show their best match to rain gauge data in simulating the daily and monthly streamflow with satisfactory precision in the 2000–2006 periods. The result indicated that the TRMM and APHRODITE data have potential applications in driving hydrological models and water resources management in data-scarce and ungauged areas in Vietnam.

Berhanu et al.(2016) used the Climate forecast system reanalysis (CFSR) dataset provided by the National Center for Environmental Prediction (NCEP) to study the rainfall variability pattern in Ethiopia using the fuzzy overlay technique for multi-indices of rainfall. The study showed that this dataset captures the rainfall pattern in Ethiopia with some magnitude of bias. Bias correction of the CFSR dataset with a linear scaling technique was applied. the performance after correction increased to NSE value of (0.8), and PBIAS of (1.3) and concluded that the efficiency of the recently released CFSR dataset in capturing the daily rainfall patterns of Ethiopia is good. Its spatial coverage of approximately 38 km grid size is also an advantage for showing the spatially distributed rainfall for different applications. The simulations

with CFSR and conventional weather yielded minor differences in the water balance components in all but one watershed, where the CFSR weather simulation gave much higher average annual rainfall, resulting in higher water balance components.

Both weather simulations gave similar, annual crop yields in the four administrative zones. Finally, it was suggested that data-scarce regions such as remote parts of the Upper Blue Nile basin, CFSR weather could be a valuable option for hydrological predictions where conventional gauges are not available.

Tolera et al. (2018) evaluated CFSR data with the conventional weather data in the Upper Awash. CFSR driven SWAT model to compare the performance of the two weather datasets at simulating monthly streamflow. Model evaluation statistics showed that the CFSR global weather data performed similarly to the conventional weather data for simulating the observed streamflow at Melka Kunture station. At Keleta station, where the conventional data is scarce, the CFSR performed better and good option for areas where no reliable weather data exists for hydrological modeling.

2.5 Hydrological models

Hydrological models are a simplified, conceptual representation of the components of the hydrologic cycle and characterizations of the real-world system through mathematical abstraction. It also involves highly nonlinear processes, complex interactions and high spatial variability at basin scale. There are different forms of hydrological models and are primarily developed for a better understanding of the hydrologic processes and prediction of hydrologic phenomena in a watershed to establish baseline characteristics whenever data is not available (Gosain et al., 2009).

2.5.1 Historical developments of hydrological models

The hydrological cycle and its dynamic processes have greatly attracted the attention of many authorities and researchers on water resources planning programs, in which understanding and prediction of the impact of changes in climate and the long-term forecasting of the water cycle made to focus on hydrological modelling. Hydrology has a long history dating back to several millennia (Biswas, 1970). However, the birth of hydrologic modeling can be traced to the 1850s when Mulvany (1850) developed a method for computing the time of concentration and hence the rational

method for computing peak discharge which is still used for urban drainage design, Darcy who conducted experiments on flow-through sands and developed what is now referred to as Darcy's law which laid the foundation of quantitative groundwater hydrology, and Fick's first law which states that under steady-state conditions, the diffusive flux is proportional to the concentration gradient (spatial) which laid the foundation of water quality hydrology formulating a mathematical model which represents hydrologic processes and the interaction between them (Gosain et al., 2009).

2.5.2 Hydrological Model selection

A wide range of hydrological models are used by different researchers; however, the applications of those models are highly dependent on the purposes for which the modeling is made (Mcintyre, 2013). According to the Cunderlik and Simonovic, (2007), choice of hydrological models depend mainly on the requirement and needs of the research or project under interest and the following as criteria: -

- a) Required output of the model
- b) Availability of input data
- c) Hydrological processes that need to be modeled to estimate the outputs adequately
- d) Prices and availability of the model.

There are many different reasons why modeling of the rainfall-runoff processes of hydrology is needed. The main reasons behind this are a limited range of hydrological measurement techniques and a limited range of measurements in space and time. Many rainfall-runoff models are carried out purely for research purposes as a means of enhancing knowledge about hydrological systems, other types of models are developed and employed as tools for simulation and prediction aiming ultimately to allow decision-makers to improve decision-making about hydrological problems (Beven, 2000). Among many hydrological models, the conceptual semi-distributed model HBV-Light was selected for this study based on the above-mentioned criteria.

2.5.3 Classification of hydrological models

Model classifications are generally based on the method of representation of the hydrological cycle or a component of the hydrologic cycle. The rainfall-runoff hydrological models can be classified as: lumped or distributed, deterministic or stochastic and conceptual or physically based models.

According to Chow et al. (1988), stochastic and deterministic models are often considered to be at the top level of the classification tree, under the way they treat the randomness of hydrologic phenomena. Stochastic models use local hydrometric data to predict flows. These models allow for some randomness that results in different outputs and are based on analysis of past events, commonly rainfall and river discharge. Deterministic models generally produce a single output of runoff for a given rainfall under identical physical environments. Without going into too much detail, deterministic hydrologic models can be classified into three main categories (Cunderlik & Simonovic, 2007).

A. Lumped models: lumped hydrologic model simulation evaluated only at the outlet of the basin that is without explicitly accounting for the response of individual sub-basins and parameters do not vary spatially within the basin. Parameters of lumped models often do not represent physical features of hydrologic processes and usually involve a certain degree of empiricism. According to Haan et al. (1994), the impact of spatial variability of model parameters is evaluated by using certain procedures for calculating effective values for the entire basin and the most commonly employed procedure is an area-weighted average. Lumped models are not usually applicable to event scale processes. If the interest is primarily in the discharge prediction only, then these models can provide just as good simulations as complex physically-based models Water Balance Model (WATBAL), Snowmelt Runoff Model (SRM), Identification of unit Hydrograph and Components from Rainfall, Evaporation and Streamflow data (IHACRES) are examples of lumped hydrological models (Beven, 2019).

B. Distributed models: distributed hydrological model parameters are fully allowed to vary in space at a resolution usually chosen by the user. Distributed modeling approach attempts to incorporate data concerning the spatial variation of parameters together with computational algorithms to evaluate the influence of this distribution on simulated precipitation-runoff behavior. Distributed models require large amounts of data for parameterization in each grid cell (Beven, 2000). However, the governing physical processes are modeled in detail, and if properly applied, they can provide the highest degree of accuracy. For instance, HYDROTEL, MIKE11/SHE, and WAT FLOOD are distributed models.

C. Semi-distributed models: parameters of semi-distributed models are partially allowed to vary in space by dividing the basin into several smaller sub-basins. Semi-

distributed model structures are more physically based than the structure of lumped models and less demanding input data than fully distributed models. Semi-distributed models can be grouped into Kinematic Wave theory models and probability distributed models. According to (Beven, 2019). The Wave theory models are simplified versions of surface and/or the subsurface flow equations of physically-based hydrologic models. In the case of the probability distributed models, spatial resolution is considered by using probability distributions of input parameters across the basin. Examples of semi-distributed models are SWAT HEC-HMS, HBV(Bergström,2001), and TOPMODEL (Cunderllk, 2003).

SWAT model

The Soil and Water Assessment Tool (SWAT) model is one of the river basin or watershed scale model developed by the United States Department of Agriculture - Agricultural Research Service (USDA-ARS) in Temple, Texas during 1970. The SWAT model has been applied in many studies around the world, especially in research related to hydrology, erosion, climate, soil, temperature, plant growth, nutrients, pesticides and land management (Arnold et al., 1998). The SWAT model is physically based, semi-distributed, continuously simulating stream flow, sediment yield, nutrients, pesticides, and agricultural management in watersheds with varying soils, land use, and management conditions over long periods (Neitsch, et al, 2011).

The major advantage of the SWAT model is its accessibility with detailed online documentation, user groups, video tutorials, international conferences and a unique literature database (more than 2700 papers) are available. This all makes the tool user-friendly, one of the best known and most widely used tools to develop water quality models at the watershed scale and its demonstrated capability almost in all basins of Ethiopia. The model is continuously improved and supported by the core developmental team and as a response to shortcomings demonstrated by many users (Czapar et al., 2005). This results in the development of new tools, e.g., GIS interface with pre-and post-processing and statistical evaluation tools. In addition, a trend to interface SWAT with other environmental or economic models enlarges its application range (Gassman et al., 2007).

The limitation of the 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 and sediment routing (Catherine Kuhn, 2014).

HEC-HMS model

The Hydrologic Engineering Center-Hydrological Modeling system (HEC-HMS) is a semi-distributed rainfall–runoff simulation model developed by the U.S. Army Corps of Engineers at the Hydrologic Engineering Center. HEC-HMS conceptually representing watershed behavior as different components of runoff processes depending upon the information needs of the hydrological study to accurately predict catchment outflows from upstream sub-catchments along with the drainage network. There are four major components for a complete setup of HEC-HMS (1) Basin model; (2) Meteorological model; (3) Control specification; and (4) Input data (time series, paired data and gridded data to be properly linked with each other for accurate operation. The Basin model in HEC-HMS Model Setup contains the hydrologic element and their connectivity that represents the movement of water through the drainage system (Scharfenberg, 2016). The arch hydro tool and HEC-GeoHMS, an Arc view extension developed by the U.S. Army Corps of Engineers (USACE) was employed to create the basin model background map file and to delineate the sub catchments from the Digital Elevation Model (DEM) using Arch GIS 10.4.1. The sub basin physical characteristics such as longest flow lengths, centroidal flow length and slopes derived from DEM are used for estimating the initial hydrologic parameters to be used as an input of HEC-HMS model. The background maps provide a spatial context for the hydrologic elements composing basin model. They are commonly used for showing the boundaries of a watershed or the location of streams (USACE, 2005).

WATBAL model

An Integrated Water Balance (WATBAL) is a lumped conceptual model developed in the Stockholm Environment Institute (SEI), USA which represents the water balance in the use of continuous functions of relative storage to represent surface outflow, sub-surface outflow and evapotranspiration (Cunderlik & Simonovic, 2015). It has essentially two main modeling components. The first is the water balance component that uses continuous functions to describe water movement into

and out of a conceptualized basin. The second is the calculation of potential evapotranspiration using the well-known Priestly Taylor radiation approach. The mass balance is written as a differential equation and storage is lumped as a single conceptualized bucket with the components of discharge and infiltration being dependent on the state variable relative storage (Cunderlik & Simonovic, 2015).

HBV-Light model

Hydrological Byråns Vattenavdelning (HBV) model is a conceptual semi-distributed water balance model for continuous daily simulation of catchment runoff (Windstorm et al., 1997). The HBV model was originally developed by SMHI in the early 70's to assist hydropower operations aiming to create conceptual hydrological model with reasonable demands on computer facilities and calibration data. HBV model is computerized catchment model that converts precipitation, potential evaporation, and snowmelt, if applicable, into streamflow/reservoir inflow by simulating the natural hydrological processes for continuous simulation of runoff, hydrological forecasting, design flood computations, and climate change studies (Seibert, 2005).

The first version of HBV light was programmed in 1995 and then continuously updated by Jan Seibert (at Oregon State University, the Swedish University of Agricultural Sciences, Stockholm University, and the University of Zurich). HBV-model has been further developed by the SMHI (Swedish Meteorological and Hydrological Institute) and has become widely used for runoff simulations in Sweden (Bergström, 2001).

The version HBV-light was developed at Uppsala University in 1993 using Microsoft Visual Basic and has become widely used in education at several universities. The important feature of this model is simple to use Windows-version for research and education. The HBV model has become widely used and exists in several versions where its development started in the 1970s (Swedish Meteorological and Hydrological Institute). HBV-light model was used and tested for its runoff modeling capacity in more than 40 countries including Ethiopia. Input data includes precipitation, air temperature, long-term average monthly estimates of evapotranspiration, runoff (for calibration), and basin geographical information (Marc, 2012).

2.5.4 Previous study with HBV-light model

Seibert, (1999) applied HBV model to eleven catchments in the NOPEX region in Sweden and establish a model parameter to watershed descriptors and modeling runoff from ungauged catchments and therefore discuss the physical basis of the model. The optimized parameter sets were decided per each catchment from the Monte-Carlo procedure, common efficiency, and a fuzzy test for uncertainty reduction through different objective functions. The runoff generation of the model well predicted the observed runoff statics and catchment features were best related to six of the 13 parameters. The simulations of runoff using the established regional parameter sets were tested with variable results. Parameter sets that gave an acceptable agreement between observed and simulated hydrograph ($R_{eff} > 0.8$) could be found for all catchments with little unfortunate results for the two smaller catchments.

Parajka et al.(2005) used HBV-Light model to examine the relative performance of a range of methods for transposing catchment model parameters to 320 ungauged catchments of Austria calibrating from 1987 to 1997 and verifying the model for the period 1976–1986. For the calibration period, the median Nash-Sutcliffe model efficiency (NSE) of daily runoff is 0.67 as compared to $NSE = 0.72$ for the at-site simulations. For the verification period, the corresponding efficiencies are 0.62 and 0.66.

2.5.5 Previous study with HBV-light model in Ethiopia

A study on the HBV-light model in different parts of Ethiopia was also made by many researchers and the results showed that the HBV-light model can be taken as one of a good hydrological model. Some of them were described below.

Worqlul et al. (2017) have assessed the performance of commonly used satellite rainfall products Climate Forecast System Reanalysis (CFSR) and Tropical Rainfall Measuring Mission (TRMM) 3B42 version 7 as input to a semi-distributed hydrological model HBV for daily stream flow simulation in the Gilgel Abay and Main Beles Basins, Ethiopia. They concluded that HBV model has explained approximately 80 % of the observed flow variation. Rainfall estimate from the CFSR has also captured the observed flow though model calibration with a good NSE and on average the CFSR runoff simulation has captured approximately 75 % of the variation of the observed flow. HBV model through model calibration have

responded for the extra rainfall of CFSR satellite rainfall estimate it has compared to the gauged rainfall. In HBV model, the maximum soil moisture storage parameter (FC) was too large indicating a deeper hydraulically active soil increasing the storage capacity of the soil. The simulation by the CFSR data for HBV model was able to capture the peak flows better than the runoff simulation by the gauged rainfall. So, the CFSR data might be more suitable to predict extreme events when using either PED or HBV models.

Jillo et al. (2017) applied HBV-light model to study the seasonal variability of water balance within the Omo-Gibe River Basin. The model predicted the water balance components, the seasonal variability of hydrological response, the spatial augmentation of aridity with the rainfall regime of all calibrated catchments in terms of the Kling-Gupta Efficiency (KGE) measure.

In this study, the new version 'HBV-light' was applied which corresponds to the HBV-6 version described by Bergstrom et al. (2001) with two slight changes. The new version executes a warming-up for the startup period at which the variables progress from standard opening values to their precise values considering meteorological conditions and parameter values. Besides, the restriction of using only integer values for the routing parameter MAXBAS has been widened.

2.6 Ungauged catchments

Ungauged catchments are part of drainage areas in the basin where hydrometric stations are not available or they became inactive and stream flow data are inadequate (in quantity and quality) of hydrological observations and as a result, model parameters are usually transferred from hydrologically homogenous gauged catchments. Runoffs are either poorly measured or absent. One possible method is the use of conceptual rainfall-runoff models. However, in such models, most parameters are not measurable but have to be estimated by calibration using at least observed runoff data Seibert, (1999). However, still the transfer of runoff series in time by the use of models is rather difficult. Therefore, model parameters and any information for hydrological modeling in such catchments are usually estimated from gauged catchments of similar characteristics either by extrapolating or regionalization. International Association of Hydrological Science (IAHS) dedicated a decade (2003-2012) to the challenging issue of Prediction in Ungauged

Basins (PUB) and defined it as the prediction or forecasting of the hydrological responses of ungauged or poorly gauged basins and their associated uncertainty. The typical approach is to look for relationships between optimized model parameter values and catchment characteristics. Parameter sets can then be compiled for ungauged catchments from measurable variables (Sivapalan, 2003).

2.7 Regionalization

Rainfall-runoff modeling of hydrological watersheds is a process by which streamflow is simulated using different modeling criteria. It also needs a model parameter optimization at gauged catchments by fitting to their observed streamflow. However, for ungauged catchments, it is difficult to determine model parameters using HBV-Light through calibration since observed stream flows are not available for calibration. Hence in such cases, it is important to apply regionalization (transferring information from gauged catchment to the ungauged catchment by defining a mathematical relationship between the calibrated(gauged) and ungauged catchments based on their physical characteristics. Therefore, regionalization is referred to the prediction of model parameters to any site in the basin using spatial and climatological pieces of evidence to generate flow data using rainfall as input (Gebru, 2009).

As it was mentioned in section 1.2, the lower part of the Omo-Gibe River Basin is ungauged but has an important hydrological resource surrounding the basin. Therefore, in such a large data-scarce area, a conceptual run-off model was set up for regionalization of model parameters and flow generation from the CFSR data-driven HBV model. HBV-Light has about eight model parameters that require optimization at gauged streams. These parameters should be determined for ungauged catchments too.

2.7.1 Regionalization approaches

Different approaches can be used for hydrological predictions in ungauged basins. the choice of catchments from which information is transferred is usually based on some sort of similarity. The most commonly used methods are Spatial proximity, Sub-basin mean method, Area ratio method, and regional model method.

The spatial distance method is applied when the ungauged catchments found near to the donor catchments based on the idea that catchments that are close to each other will likely have a similar runoff regime since climate and catchment conditions will often only vary slightly towards the boundary (Merz et al., 1999). So, the assumption is made that catchments are highly homogeneous with respect to geographic location. Elias (1995) used kriging and the nearest neighbor technique in their study and concluded that geographical proximity does not guarantee hydrological similarity. The area ratio method can be used for simulation of streamflow in ungauged basins assuming that the streamflow contribution from each sub-catchment to the total catchment yield is proportional to a ratio of the catchment area or other attributes (Schreider et al., 2002). The sub-basin mean method has the idea behind that the catchments will collectively have the mean of similar characteristics and then the mean of model parameters represent the area. The Regional model method detailed in sub-section 3.5 was applied in this study.

2.7.2 Regional model development

The determination of the mathematical relationship between the spatial and temporal distribution of the model parameters for each catchment uniquely in the basin (parameter estimation) scheme consists of regression equations that relate each model parameter to the physical catchment characteristics (land surface, topographical, and climatological characteristics) of both ungauged and gauged catchments (Parajka et al., 2005). To characterize the values of the model parameters for each catchment through a regression model, the optimized model parameters were first calibrated and validated at the gauged catchments.

2.7.3 Principal component analysis (PCA)

Run-off estimation at a particular location is dependent on several catchment characteristics but these Physical catchment characteristics are correlated among themselves. Besides, some hydrological components are more predominant in their information content than other components. Therefore, PCA is a popular statistical technique that transforms the original correlated variables into uncorrelated components (also called orthogonal components or principal components). These components are the linear transformations of the original variables of an orthogonal axis (Wuttichaikitcharoen & Babel, 2014). PCA can therefore remove the correlation among the independent variables (catchment characteristics) by forming

another set of uncorrelated variables and possibly reduce the dimension of the data matrix before regressing them on the model parameters in the regional model. PCA is therefore a technique used for reducing redundancy of catchment descriptors and find the more effective parameters for prioritization of the watershed and concluded that the results of PCA reflect a good look on the prioritization of watershed.

Sharma et al., (2014) applied Principal component analysis to 13 dimensionless geomorphic parameters on 8 sub watersheds of Kanhaiya Nala watershed tributary of Tons River located in Part of Panna and Satna district of Madhya Pradesh, India, to group the parameters under different components based on significant correlations. Therefore, Principal Component loading is applied to get a better correlation and group the parameters into physically significant components in the regional model.

CHAPTER 3 MATERIALS AND METHODS

3.1 Description of the Study Area

3.1.1 Location

The study area is one of the major river basins in Ethiopia and is situated in the southern part of the country. The basin covers an area of 79,000 km² with a length of 550 km and an average width of 140 km. The geographic location is 4°00'N & 9°22'N latitude and 34°44' E & 38°02' E longitude and a mean altitude of 2800masl. The unique characteristic of this basin is its complex topography separated abruptly into the highlands in the north and lowlands in the south. The western watershed is the range of hills and mountains that separate the Omo-Gibe Basin from the Baro-Akobo Basin. To the north and northwest, the basin is bounded by the Blue Nile Basin with a small area in the northeast bordering the Awash Basin. (Richard Woodroof and Associates, 1996).

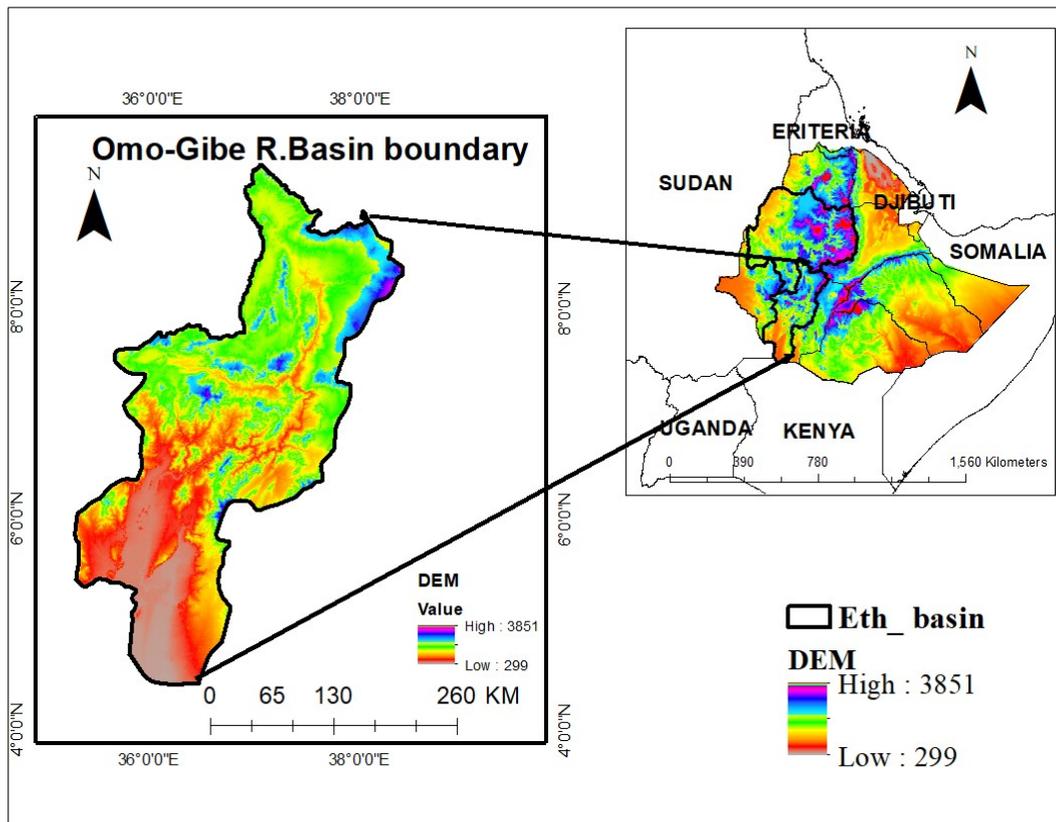


Figure 3.1: Location map of study area

3.1.2 Climate

The climate of the Omo-Gibe River Basin varies from a tropical humid one in the highlands to a hot arid climate in the southern part of the floodplain. The climate in the entire part is tropical sub-humid. Precipitation in the Basin diverges from over 1900 mm per annum in the north-central areas to less than 300mm per annum in the south. Furthermore, the rainfall seasonality is unimodal for the northern and central parts of the basin and bimodal for the south. The average yearly temperature in the basin fluctuates from 16⁰c in the plateaus of the north to 29⁰c in the lowlands of the south. As the country, Ethiopia is a tropical region, the Inter-Tropical Convergent Zone (ITCZ) is the principal factor that influences its weather system. The seasonal rainfall distribution in the Basin arises out of the yearly exodus of the ITCZ. The Rainfall pattern strongly decreases from north to south of the watershed particularly less than 300 mm/year near Lake Turkana (Richard Woodroof and Associates, 1996).

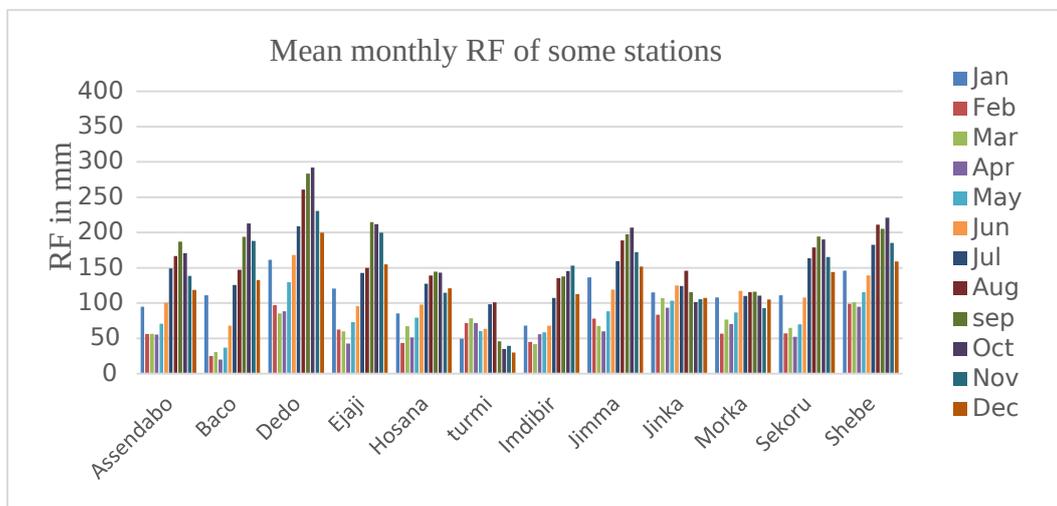


Figure 3.2: Mean monthly rainfall of some stations (1987 -2010).

3.1.3 Topography

The topography of the Omo-Gibe River Basin as a whole is characterized by its physical variation. The northern two-thirds of the basin has mountainous to hilly terrain cut by deeply incised gorges of the Gojeb, and Gilgel-Gibe Rivers, while the southern one-third of the basin is a flat alluvial plain punctuated by hilly areas. The northern and central half of the basin lies at an altitude greater than 1500m.a.s.l with a maximum elevation of 3360m.a.s.l (located between Gilgel-Gibe and Gojeb tributaries), and the plains of the Lower-Omo lies between 400-500 m.a.s.l.

The headwaters of the Great-Gibe River are at an elevation of about 2200 m.a.s.l. Although there are some important tributaries from different directions, the general direction of flow of the Gibe River is towards the Omo River and then to Lake Turkana a fault feature, filled with alluvial and lacustrine sediments of recent origin associated with the Great Rift Valley (Jillo et al., 2017).

3.1.4 Land use

Most of the northern catchments of the Omo-Gibe Basin is under extensive cultivation with increased land pressure, meaning the expansion of cultivated areas into increasingly marginal lands at the expense of woodlands. Deforested areas are now confined to areas too steep and inaccessible to farm. The main gorges of the basin are relatively unpopulated and support a cover of open woodland and bushland with grasses, the eastern part of the basin has some of the most densely populated and intensively farmed areas in the country. The south of the basin is more sparsely populated with a greater population of natural vegetation, though even here the forest is decimated at an alarming rate (Sahle *et al.*, 2019).

3.1.5 Drainage sub-basins and River system of Omo-Gibe River Basin

Recently, river flows are being measured at different points in the basin. Of these all the stations, except one, measures flow in small sub-basins. The exception is the gauging station on the Gibe at Abely where the catchment area is 15,804 km² and represents about 20 % of the basin area. The northern part of the catchment has many tributaries from the northeast, the largest ones being the Wabe and the Walga. These drains largely cultivated land, much of it with rather impeded drainage; this is an area where erosion processes are important. The Tunjo and Gilgel-Gibe are rivers also draining mainly cultivated land from the southwest; having a higher proportion of more permeable soils than the Walga and Wabe catchments (Richard Woodroof and Associates, 1996).

3.2 Data collected and materials used in the study

3.2.1 Data collected

To achieve the proposed objectives, having relevant data is mandatory. These data are satellite-produced weather data (CFSR), Meteorological data, Hydrological data & data of catchment characteristics of the basin.

3.2.1.1 Ground meteorological data

In the study basin, the meteorological stations located in and around the watershed for the required period of records of different classes were collected from the Ethiopian National Meteorological Agency (NMA). Meteorological data that were collected for estimation of streamflow include rainfall, temperature, relative humidity, topographic and vegetation information, wind speed, and sunshine hour. Uneven distribution and data incompleteness in the stations have challenged the performance of meteorological data to use as a reference for CFSR data evaluation. Therefore, the selection criteria of the stations were based on their long-time data availability and their relative impact on the study area (weight) around with representative coverage and Class (that was given by the agency).

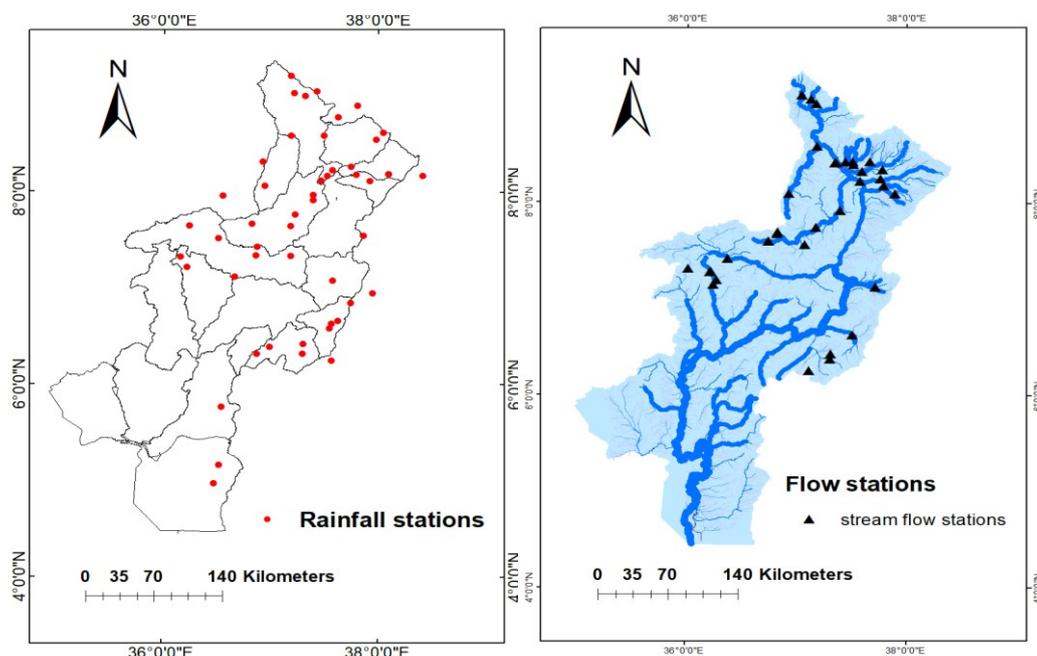


Figure 3.3: Meteorological stations and River system of Omo-Gibe Basin

3.2.1.2 River Discharge Data

River discharge data of the basin were collected from the Ministry of Water, Irrigation, and Energy (MoWIE), Basins Development Authority (BDA), Hydrology and Water Quality Directorate (HWQD) for a required period of years. These data were then used for hydrological modeling (sensitivity analysis, calibration, and validation of the model). There are 33 operational flow gauging stations in the upper and central parts of the basin, most of which are located in small streams that could not be representative of the runoff dynamics in the basin, and besides many have

only short records with many missing data and available from 1979 to 2010 in most stations. The southern, arid sub-basin are ungauged (Jillo et al., 2017). There is also a considerable fluctuation of flow due to rainfall and temperature seasonal variation (Saito et al., 2018). The stream flow gauging stations used in this study were shown in Figure 3.2.1.1.

3.2.1.3 Climate Forecast and System Re-analysis (CFSR) data

The CFSR dataset include daily rainfall, maximum and minimum temperature, wind speed, relative humidity, and solar radiation, and sunshine hour, downloaded from the globally accessible from <http://cfs.ncep.noaa.gov/cfsr> for a bounding box of 4° 00" N - 9°22" N latitude and 34°44"E - 38° 24" E longitude for the period 1979 to 2010 on daily basis. The data set was then clipped for only the watershed boundary of the Omo-Gibe River Basin with the support of GIS application. CFSR has a spatial resolution of 38km*38km (0.32° * 0.32°) grid size.

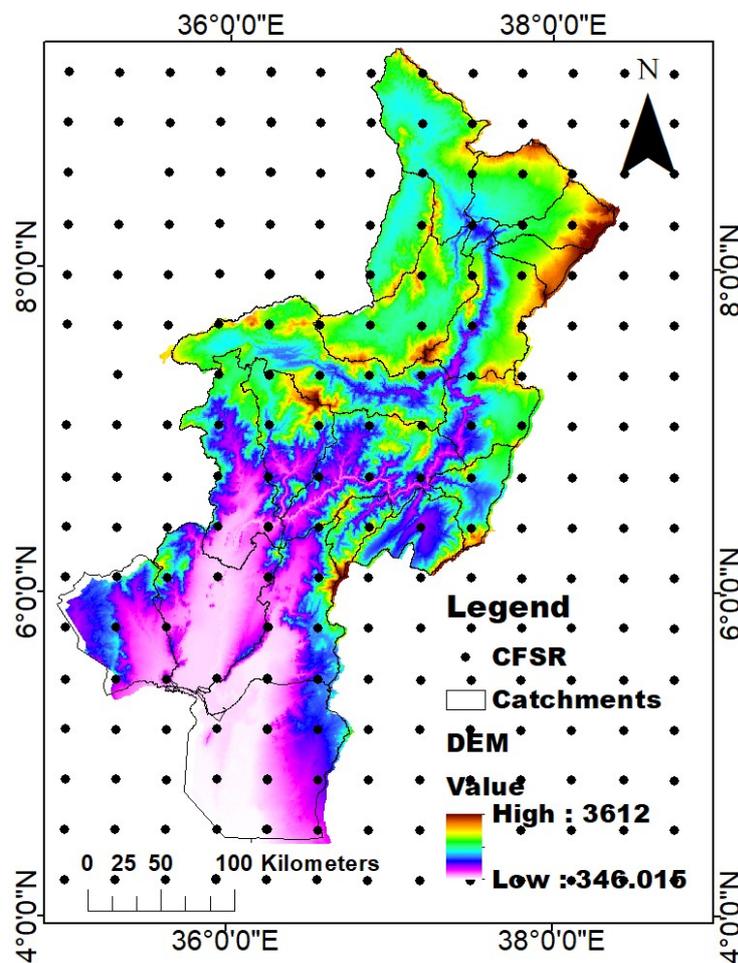


Figure 3.4 : Distribution of CFSR Grid data at Omo-Gibe Basin

3.2.1.4 Mean Annual Rainfall and Mean Annual Potential Evapotranspiration map

Runoff depends on the intensity of rainfall. More the rainfall, more will be a runoff. If the rainfall intensity is very less and it rains as light showers then much of the water will be lost in infiltration & evaporation resulting less runoff. The mean annual rainfall distribution map and mean annual evapotranspiration maps were kindly provided by International Water Management Institute (IWMI) in a raster format and extracted for each sub-catchment using Arc Map. These maps are important indicators of the spatial variability of climate in the basin that directly affects the rainfall-runoff relations and the surface energy balance that how much water is removed from each watershed. As shown in Figure 3.6. The maps were also used for grouping the catchments based on the similarity of climatic indices in the regionalization process since catchments with similar attributes have a comparable contribution to runoff generation. The climate of the Omo-Gibe Basin is characterized as tropical and humid in the north and north western part and hot and arid in the southern part with spatiotemporal variations of five rainfall regions and temperature observed across the basin, three of them having a unimodal and one is a bimodal rainfall regime. Tesfaye (2011) has used the MAR and MAPE map indices in the regionalized model in Omo-Gibe Basin and got reasonable relation with model parameters.

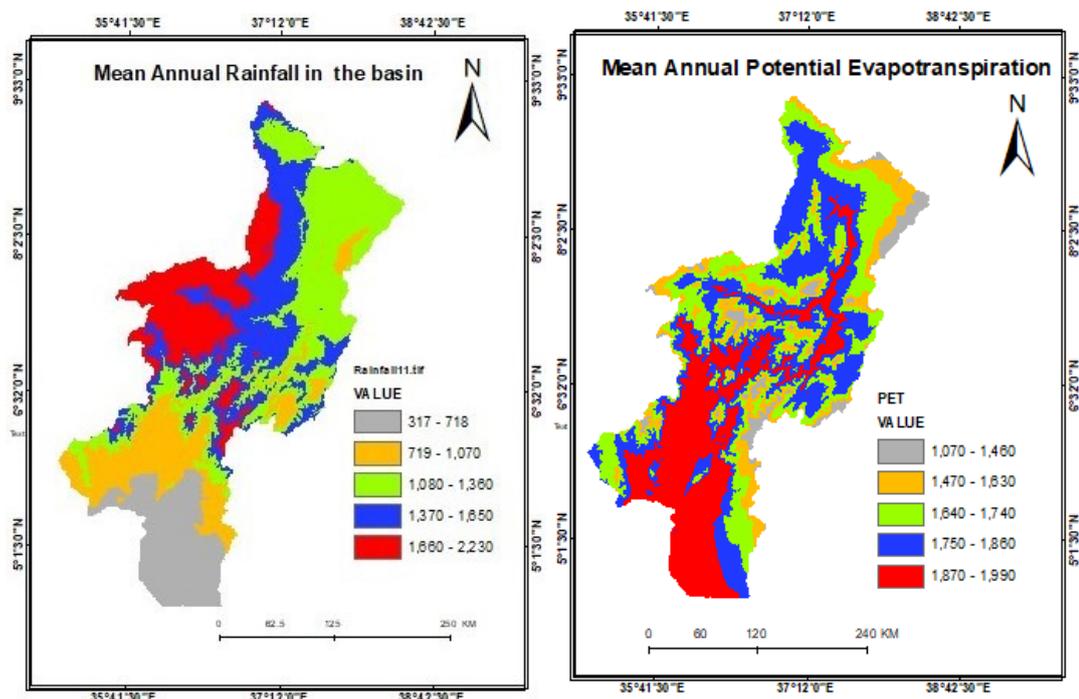


Figure 3.5: Mean Annual Rainfall and Mean Annual Evapotranspiration map

these maps were also used to compare the mean annual rainfall and potential evapotranspiration indices calculated from the CFSR and meteorological data applying the Hargreaves method.

3.2.1.5 The Normalized Difference Vegetation Index (NDVI)

Catchments were classified based on their land use and land cover indices using the NDVI map obtained from IWMI using the raster calculator in ArcMap for each catchment ranging from -1 to 1. As the presence of vegetation cover increases the water holding capacity of the soil and reduces the peak runoff by causing the retardation of the overland flow. As shown in Figure 3.7 below, the central-western part of the basin boundary area has extensive tracts of healthy, dense vegetation forest confined to areas too steep and inaccessible to farm with NDVI values of 0.68 to 0.98. The Gibe, Gojeb, and Omo gorges are relatively unpopulated and support a cover of open woodland and bushland through the inaccessible area, such as where the Addis Ababa to Jimma road crosses the Gibe Gorge, sparsely vegetated areas with NDVI of 0.2 to 0.5. The Omo-Gibe River bodies and lower parts with rocks, clouds, and bare soil have negative NDVI values.

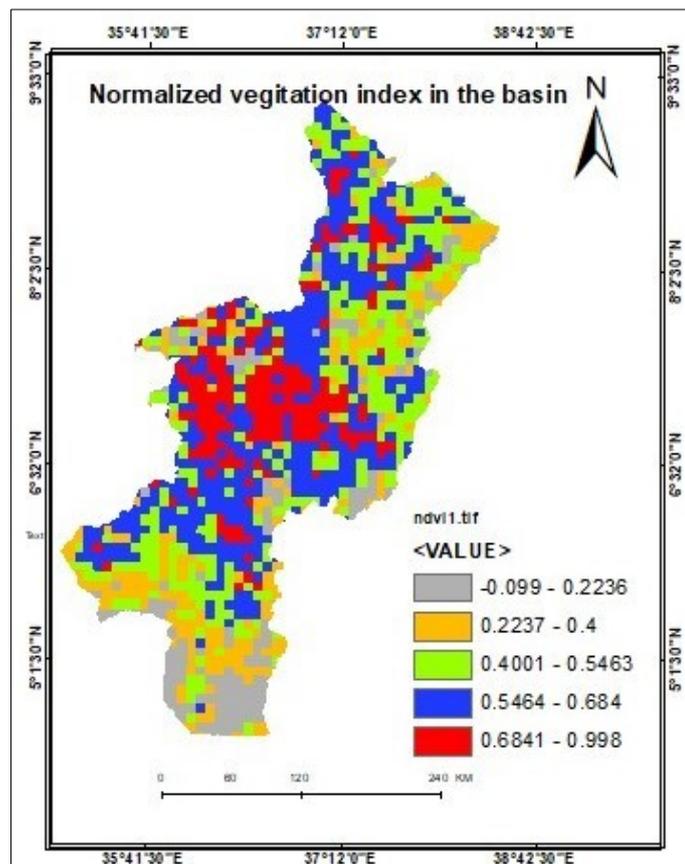


Figure 3.6: NDVI map of the Omo-Gibe River Basin

3.2.1.6 Soil maps

Rainfall-runoff model parameters were usually sensitive to soil descriptors of the catchments that govern the hydrological process. The soil maps of the study basin were obtained from IWMI and used as a potential index for each catchment include (Soil permeability, Soil Available Water Content, Soil bulk density).

Soil permeability (SAT_K) map

The permeability index of the study area ranging from 7.42 m/s to 34.5 m/s indicating the spatial soil water movement and relates to the response routine of the upper soil zone (ALPHA) of the model. Hundecha & Ouarda (2008) have shown the correlation of ALPHA with land use and soil permeability. In Booij et al. (2007) ALPHA got a positive correlation coefficient with permeability index and physiographic descriptors. (Jillo et al., 2017) established a positive correlation of the permeability index and slope direction with (ALPHA) in the study basin.

Soil bulk density (SOL_BD) map

Soil bulk density of the study area ranging from a minimum of 1.1 kg/dm³ to a maximum of 1.34 kg/dm³ with an average of 1.7 kg/dm³ is relevant information with soil profile depth, due to changes in organic matter content, porosity, and compaction accounting for horizon mass.

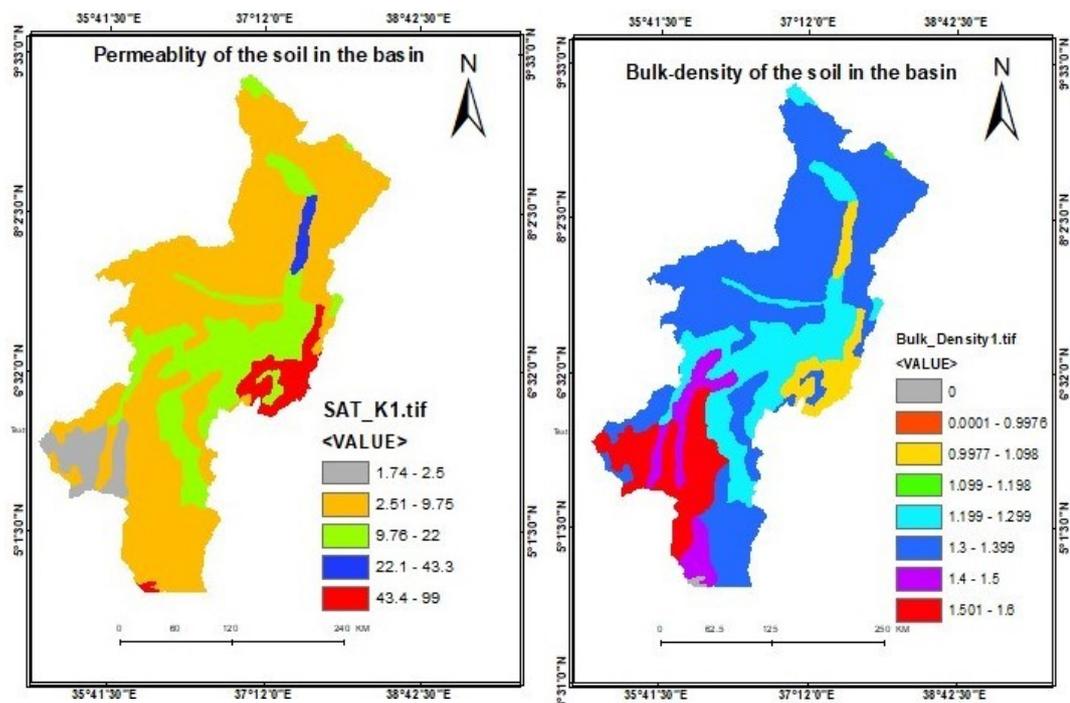


Figure 3.7: Soil Permeability (SAT_K) and Bulk Density (SOL_BD) map

Available Soil Water Content (SOL _AWC) map

The HBV-Light model has been reported to cover a wide range of soil heterogeneity conditions and parameters that controls the model parameters in the rainfall-runoff process function with empirical parameters, BETA and FC. BETA is an index of soil heterogeneity in water content in the basin, A BETA value of zero implies that a basin entirely lacking in the water holding capacity in the soil whereas a high BETA value indicates such homogenous condition that the whole catchment is regarded as a bucket that overflow Simultaneously when the field capacity is reached (Bergström et al., 2001).

Available soil water content is related to climate variables (temperature, wind speed, and humidity) that may control the model parameters in the runoff processes.

As shown in the Figure 3.2.8 below, the soil water content decreases from northern catchments (high summer rainfall intensity, wet, humid temperate) to southern catchments (low rainfall intensity, arid, high evapotranspiration) leading to low streamflow towards north south. Temperature, wind speed, and humidity affect evaporation and transpiration rates, thus infiltration rate, and soil moisture regime, following change in runoff volume. On the other hand, these soil characteristics,

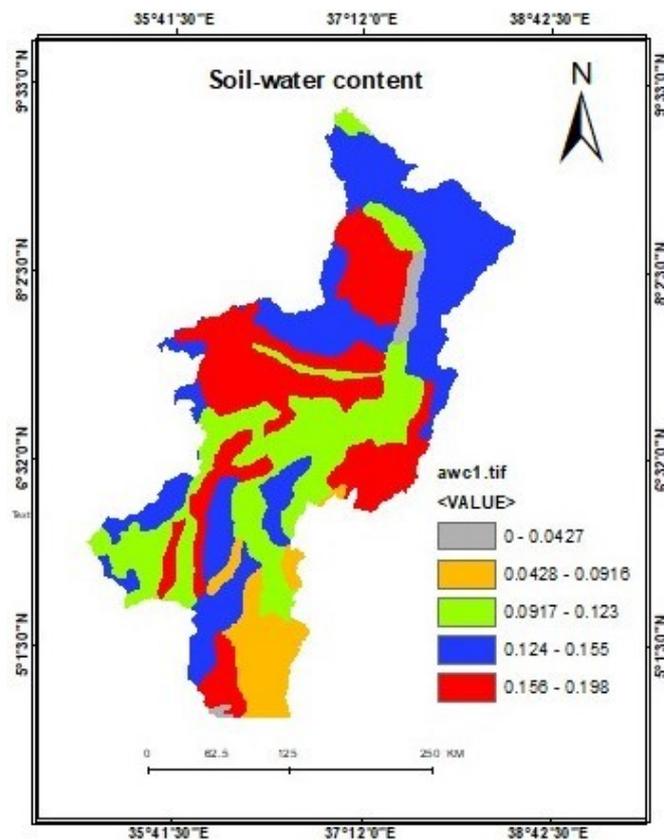


Figure 3.8 Soil Available Water Content map of the Basin

vegetation cover and slope angle are among the natural factors controlling the proportion of precipitation that is converted to runoff in a given landscape, and the time it takes for runoff to enter a stream.

3.2.1.7 Digital Elevation Model (DEM) data

The Digital Elevation Model (DEM) used in the Omo-Gibe River was obtained from a digital elevation map of the study area that was prepared using the Shuttle Radar Topography Mission (SRTM) with a resolution of 30 m x 30m. Then the DEM was loaded into ArcGIS 10.41 in a grid format as a digital representation of the land surface elevation and extraction of catchment indices such as elevation, slope, river networks for hydrological modeling.

3.2.2 Materials used

- ✚ ARC-GIS 10.4.1, to obtain hydrological and physical parameters and spatial information.
- ✚ HBV-Light model, for rainfall-runoff simulation from both datasets.
- ✚ SPSS software: statistical package for developing multiple linear regression between model parameters and physical catchment characteristics including PCA analysis.

3.3 Data analysis

3.3.1 Rainfall data quality check

Homogeneity check

The homogeneity test of rainfall stations was performed to identify a change in the statistical properties of the time series data which was caused by either natural or man-made factors (Wijngaard *et al.*, 2003). For this study, a homogeneity test has been carried out to the neighboring stations that are supposedly homogeneous. The non-dimensional of the month 's value was obtained as: -

$$P_i = \frac{\bar{P}_i}{\bar{P}_{av}} * 100. \dots\dots\dots(3.1) \dots\dots\dots$$

where, P_i represent non-dimensional value of rainfall for the month i , \bar{P}_i represent over years averaged monthly rainfall at the station i , and \bar{P}_{av} is the over year average yearly rainfall of the station. In this study, the homogeneity of rainfall

stations was described below in the graph. In Figure 3.3.1, the selected rainfall stations were non-dimensionalized and plotted together to analyze their homogeneity. The same mode and pattern of the stations were observed and maximum rainfall falls of the stations between May to September in all stations which shows the homogeneity of the stations. All the values of Pi show the result of homogeneity analysis and the time-series observational data of relatively selected stations are homogenous.

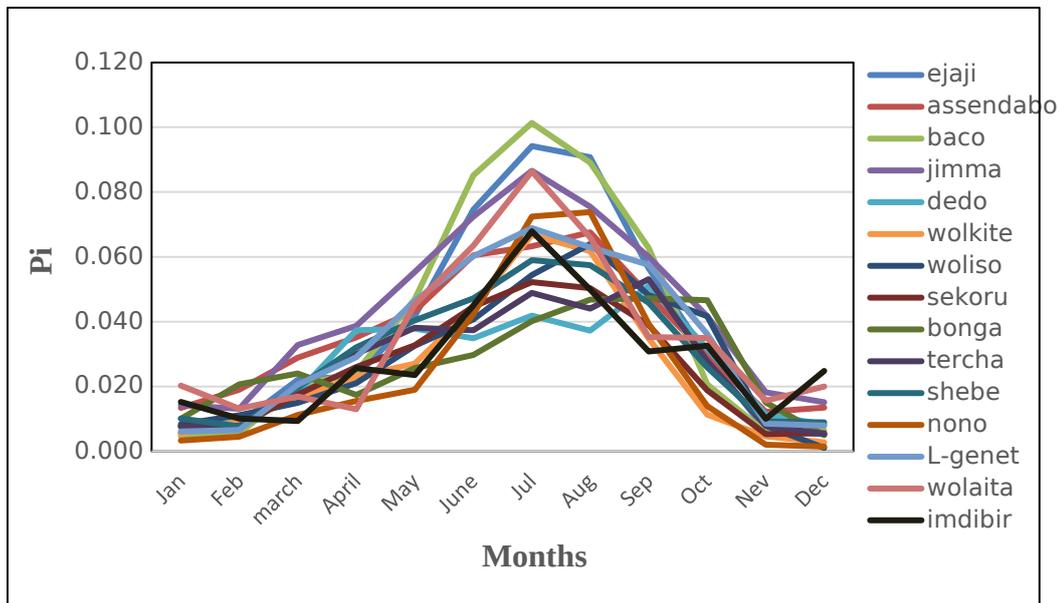


Figure 3.9: Homogeneity test for meteorological stations

Check for Consistency

Double mass curve analysis was carried out to test the internal consistency of a time-series of observational data graphically for identifying and adjusting the inconsistency if any in a station record by comparing its time trend with those of adjacent stations in such a way that the accumulated totals of the gauge in question were compared with the corresponding totals for a representative group of nearby gauges (Giambelluca *et al.*, 1986). When a significant change in the regime of the curve is observed, it reveals that rainfall data is inconsistent at that station and it should be corrected by using Equation 3.2 below (Subramanya, 1998).

$$P_{adj} = \frac{M_C}{M} * P_i \dots \dots \dots (3.2) \dots \dots \dots$$

Where P_{adj} is the adjusted rainfall at any period, P_i is originally recorded rainfall at the period, M_c is the correct (straight-line) slope of the double mass curve and M is the original slope of the double mass curve.

From the double mass curve analysis (Figure 3.3.2), the lines are fairly smooth with no station displaying a long-lasting break in slope. Therefore, and the stations used in this study have not undergone a significant change and therefore no need for further correction of data consistency.

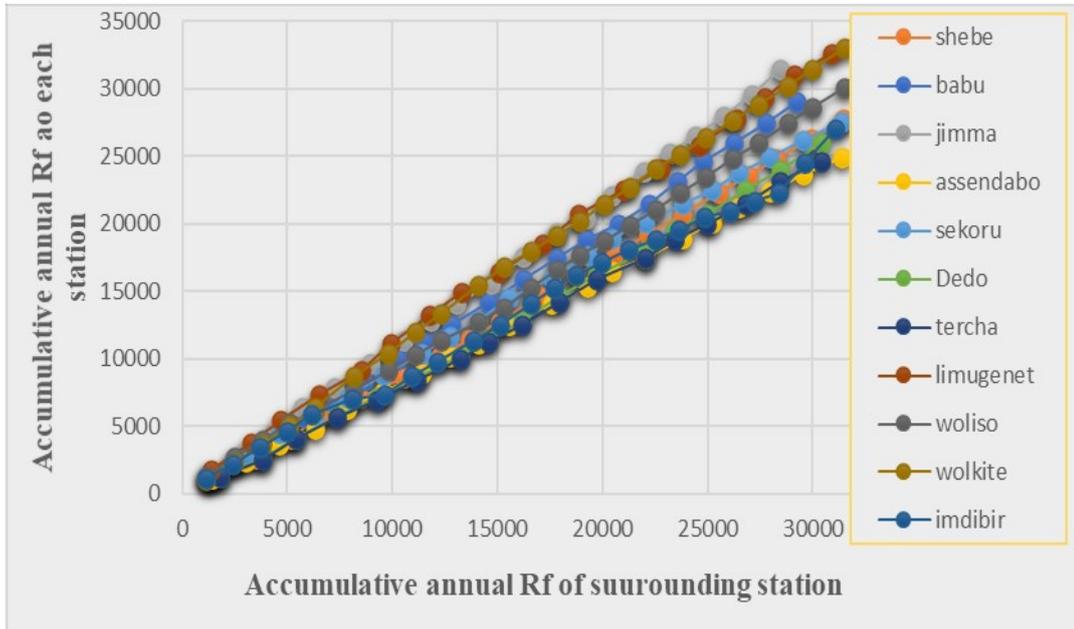


Figure 3.10: Consistency of rainfall data in the stations

Filling in Missing Weather Data

The missed data in the rainfall records of gauged stations usually estimated from the neighboring stations by applying the arithmetic mean method, normal-ratio method, and simple linear interpolation methods. The normal-Ratio method was used for this study because the mean annual precipitation of the adjacent stations exceeded the station in question by more than 10% and it is formulated as

$$P_x = \frac{1}{N} \left(\frac{N_x}{N_A} P_A + \frac{N_x}{N_B} P_B + \frac{N_x}{N_C} P_C + \dots + \frac{N_x}{N_N} P_N \right) \dots \dots \dots (3.3)$$

Where, P_x is the precipitation for the station with a missed record

P_A, P_B, P_C, P_N the corresponding precipitation at the index stations and
 N_A, N_B, N_C, N_N are the long-term mean monthly precipitation at the index stations and station x

Streamflow Data Analysis

The selected hydrological stations with their catchments are in (Appendix D, Table 6). Filling and extension of missed records were carried out by developing a correlation between the station with missed data and any of the nearby stations, whichever gives good correlation streamflow with the properties of historical observed data (Maidment, 1995).

Test for Outliers

The presence of high and low outliers in the measured streamflow data and some unrealistic records such as negative flows, constant observations for successive days tested under the assumption that the logarithms of the original series are normally distributed. For the natural logarithm of the variable, the upper and lower limits of outliers are given as:

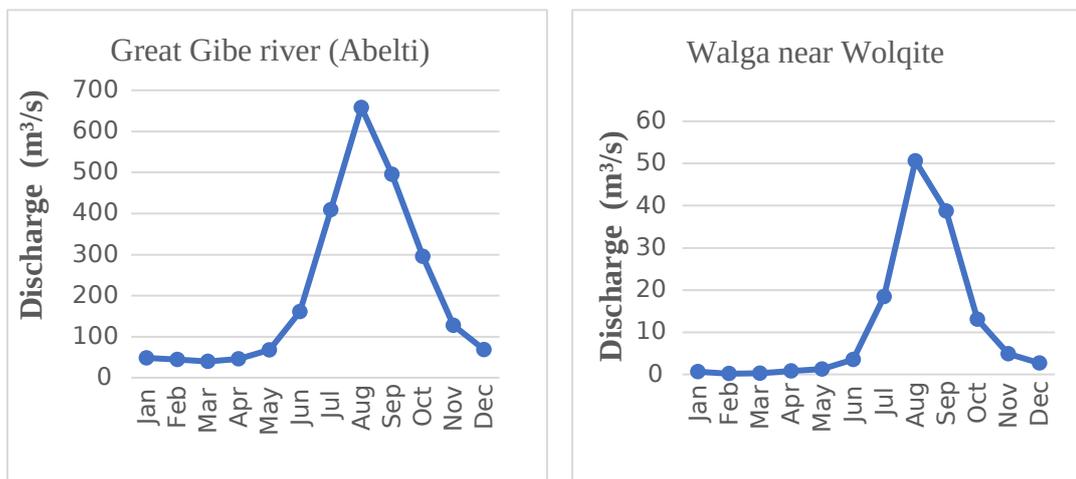
$$X_{higher} = \exp(\bar{\chi} + K_n * S_y) \dots \dots \dots (3.4)$$

$$X_{lowerr} = \exp(\bar{\chi} - K_n * S_y) \dots \dots \dots (3.5)$$

where, $\bar{\chi}$ and S_y are the mean and standard deviation of the natural logarithm of the original variable respectively. K_n is the frequency factor representing the outlier test statistic that depends on the sample size n as shown in equation 3.6.

$$K_n = -0.9043 + 3.345 \sqrt{\log n} - 0.4046 \log n \dots \dots \dots 3.6$$

Some flow gauging stations with the mean monthly flows have shown below in the figure.



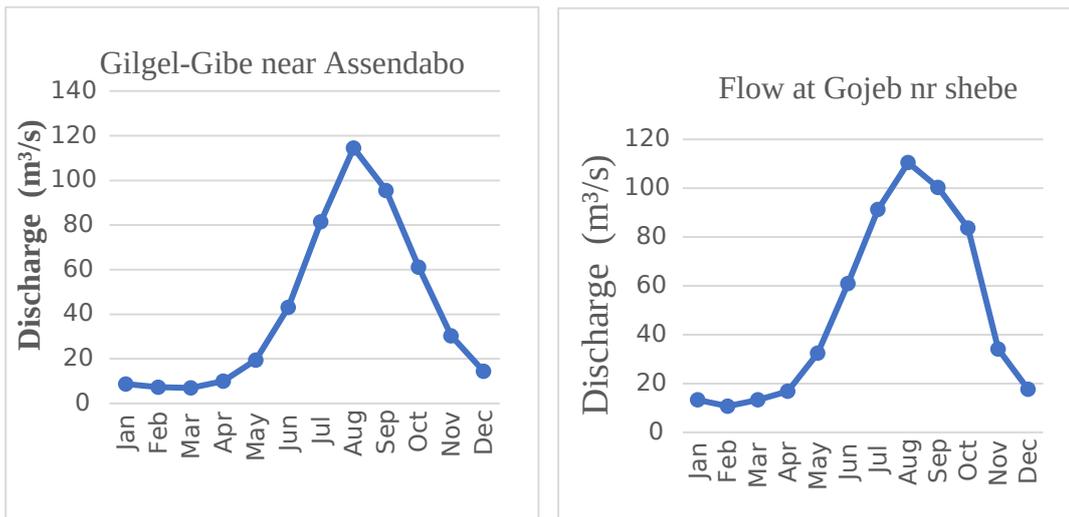


Figure 3.11: Monthly Average Discharges at a some of gauging station (1987-2010)

3.3.2 Evaluation of CFSR for Orographic effect

In order to investigate the effect of orographic rainfall in the estimation of CFSR and meteorological data in each catchment of study area, the long-term mean annual CFSR and ground observed rainfall from 1987 to 2010 was plotted against catchment elevations to see the rainfall–elevation relations. As the result shown in the Figure 3.12, its observed that there is no any significant correlation between the mean annual rainfall estimates of the two data sets and mean elevation and consequently no need for precipitation lapse adjustment to the HBV-Light model simulations conducted by CFSR data.

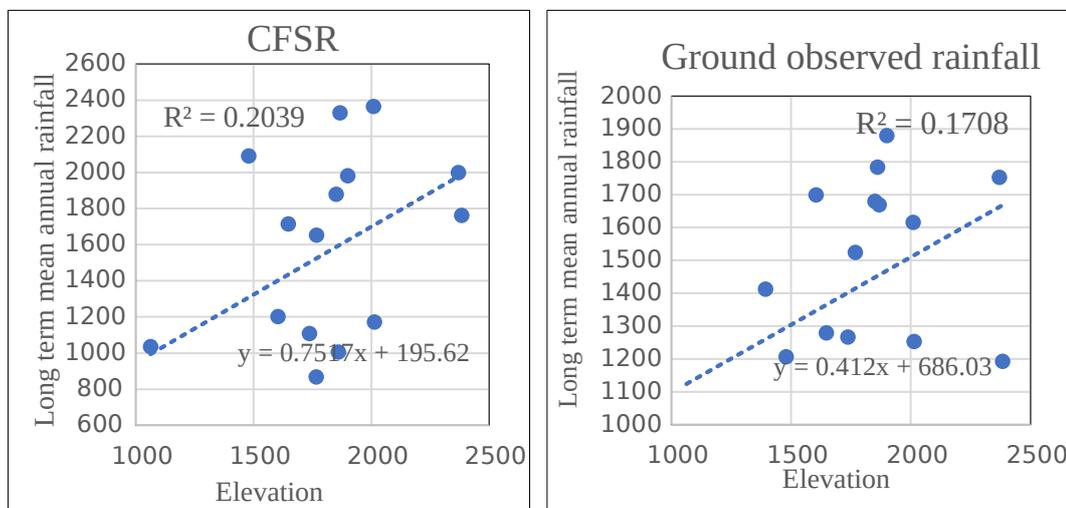


Figure 3.12: Evaluation of CFSR data for Orographic effect

According to Saha et al., (2010), CFSR is configured with blending model gravity wave drag based on Kim and Arakawa (1995) approach and sub-grid scale mountain blocking following Lott and miller (1997) algorithm for coupling atmospheric, oceanic, and surface - modeling component and elevation gradient.

3.3.3 Bias correction for CFSR data

For this study, from among other bias correction methods most of which are statistical, the linear scaling (LS) bias correction method that focuses mainly only on the magnitude of the data excluding the pattern and trend of the dataset was chosen based on findings of Berhanu et al. (2016), Worqlul et al. (2014), who were showed that CFSR can reproduce the observed rainfall pattern in the Blue Nile basin, Ethiopia but overvalued or undervalued only the magnitude and therefore, the adjustment was only made to conform the monthly average of adjusted values to the observed ones as described in equation 3.6. The change factor for precipitation is a multiplier that is computed from the ratio of the monthly mean of the observed to the raw dataset.

$$P_{dcorr} = P_{dsat} \frac{P_{obs}}{P_{sat}} \dots\dots\dots(3.6)$$

P_{dcorr} represent corrected daily satellite rainfall data, P_{dsat} represent long-term mean monthly raw satellite rainfall data. P_{obs} represent long-term mean monthly observed rainfall data.

3.3.4 Comparison of areal CFSR data to areal ground observed rainfall data

Areal estimates of daily CFSR data were compared to the areal estimates of interpolated ground rainfall estimates at each catchment in the basin from 1987 to 2010 years. Areal estimates of satellite rainfall grid generated from CFSR data were determined by weighting each CFSR station in proportion to its spatial coverage of 1444 km². Each CFSR station is multiplied by its area and the sum of those products divided by the total area of the catchment. The sum of the area of each CFSR station gives the area which is equivalent to the total watershed area. Ground rainfall observations are interpolated using the Thiessen polygon method. There are nineteen catchments extracted in the basin including ungauged catchments and the comparisons were carried out on the daily and monthly time steps using time series-

based metrics include the regression coefficient (R^2), Nash–Sutcliffe measure of efficiency (NSE), and percent bias (PBIAS).

The coefficient of determination (R^2): - is the square of the Pearson product–moment correlation coefficient and describes the proportion of the total variance in the observed data that can be explained by the model. It’s value is an indicator of the strength of the relationship between the observed rainfall and CFSR rainfall. R^2 ranges from zero to one with higher values of one indicating better agreement (Legate and McCabe, 1999) and formulated as:

$$R^2 = \frac{\sum_{i=1}^n (Q_{sim} - \overline{Q_{sim}})(Q_{obs} - \overline{Q_{obs}})}{\sum_{i=1}^n (Q_{sim} - \overline{Q_{sim}})^2 \sum_{i=1}^n (Q_{obs} - \overline{Q_{obs}})^2} \dots\dots\dots(3.7)$$

Percent bias (PBIAS): - evaluates the ratio of the mean value of the CFSR data to the ground rainfall gauge by measuring the average tendency of the simulated data to be larger or smaller than their observed counterparts. It shows overestimation or underestimation in the magnitude. The value of perfect much (no PBIAS) is zero. A value greater than zero indicates an underestimation of the observed rainfall, negative value rises from the overestimation of CFSR (Gupta, 1999).

$$PBIAS = \frac{\sum_{i=1}^n (Q_{obs})_i - (Q_{sim})}{\sum_{i=1}^n (Q_{obs})_i} * 100 \dots\dots\dots(3.8)$$

The Nash-Sutcliffe efficiency (NSE): - describes how well the CFSR agrees to the observed rainfall. The model is taken as very good performing when ENS is 0.75 to 1 and fair to good performing model when is 0.6 to 0.8 (Nash and Sutcliffe, 1970).

$$Ens = 1 - \frac{\sum_{i=1}^n (Q_{obs})_i - (Q_{sim})^2}{\sum_{i=1}^n (Q_{obs})_i - (Q_{obs})_i^2} \dots\dots\dots(3.9)$$

Where: - Q_{obs} is observed discharge in the simulation case and ground observed rainfall in the comparison case, $\overline{Q_o}$ mean of observed discharge in the simulation case or mean of CFSR estimate in the comparison case, Q_{sim} is simulated discharge, $\overline{Q_s}$ the mean of simulated discharge in the simulation case or mean of CFSR in the comparison case.

3.3.5 Evaluation of rainfall detection capacity of daily CFSR data

In this study, the misses and hits of CFSR data were evaluated using two verification methods, the Probability of Detection (POD) and The False Alarm Ratio (FAR) based on the contingency Table 3.1 (Shrestha et al., 2017). Therefore, the assessment was conducted for the visualization of only rain detection capability of the CFSR regardless of its magnitude against observed data as follows.

The Probability of Detection (POD): as is illustrated in table 3-2, it indicates what fraction of the observed rainfall was just detected by CFSR. A perfect score of POD is 1 and a poor score is 0.

$$POD = \frac{D}{D + B} \dots\dots\dots(3.10)$$

The False Alarm Ratio (FAR): measures the fraction of rain detections that were falsely alarmed (rainfall did not occur) and vice versa. A perfect score is 0.

$$FAR = \frac{C}{D + C} \dots\dots\dots(3.11)$$

- 📊 “Hit” shows that CFSR detects rainfall of a certain value as the observed does.
- 📊 ‘False alarms’ showed that CFSR estimated oppositely to the Insitu measurement.
- 📊 ‘Misses’ represent CFSR failed to detect a rainfall of a certain day and
- 📊 ‘Correct negatives’ show correctly estimated no rain events by the CFSR product as the Insitu measurement (Shrestha, 2017).

Table 3.1: Rain detection capability and False Alarming Ratio

		Satellite rainfall	
		No rain	yes rain
Observed rainfall	No rain	Correct negative (A)	False alarm (C)
	yes rain	Miss(B)	hit (D)

Ringard et al., (2015) recommended that for flood forecasting and hydrological modeling purposes, it is important to avoid underestimations of rainfall events and rainfall amounts, and then avoid PBIAS > 0 and low POD. In contrast, for drought

monitoring, overestimations must be avoided, and then avoid PBIAS < 0 and high FAR.

3.3.6 Description of HBV-Light model

Hydrological Byr°ans Vattenavdelning (HBV) model is a conceptual semi-distributed water balance model for continuous daily simulation of catchment runoff. The important feature of this model is simple to use Windows-version for research and education. HBV-light model was used and tested for its runoff modeling capacity in more than 40 countries including Ethiopia. Input data includes precipitation, air temperature, long-term average monthly estimates of evapotranspiration, runoff (for calibration), and basin geographical information (Bergstorm, 1995).

3.3.7 HBV-Light model components and structures

HBV-light consists of a snow routine representing snow accumulation and snowmelt by a degree-day method, a soil routine where groundwater recharge and actual evaporation are computed as functions of actual water storage, a response routine with three linear reservoir equations, and a routing routine using a triangular weighting function (Marc, 2012).

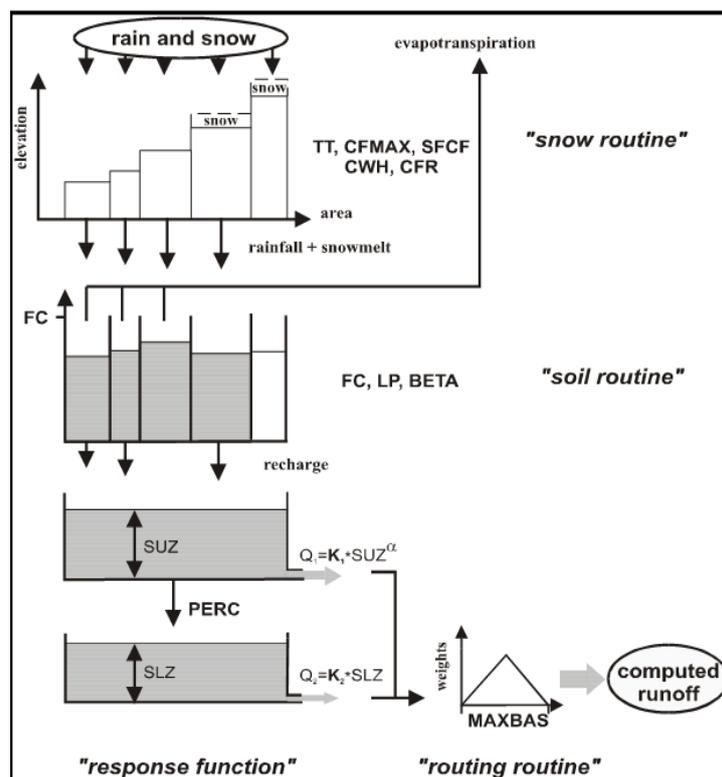


Figure 3.13: Schematic structure of HBV-Light model (Seibert, 2005)

Snow routine

Precipitation is simulated to be either snow or rain depending on whether the temperature is above or below a threshold temperature, TT [°C] (SMHI-2008 manual version 6.2.). Input data are daily air temperature and precipitation. The output is the effective precipitation as rainfall and snowmelt which is fed as input into the soil moisture routine. Since the study area does not experience any snowfall during the year, hence this component was avoided.

Soil Moisture Routine

It is in the soil moisture subroutine that effective rainfall is fed into the soil moisture routine and runoff formation occurs against the actual evapotranspiration. It depends upon the soil infiltration rate and intensity of rainfall based on three parameters. Beta (β), FC, and LP. β controls the contribution to the response function from each millimeter of rainfall. Depending on the relation between maximum soil moisture-holding capacity [FC] and simulated soil moisture [SM], recharges will be generated. If the infiltration rain satisfies the soil moisture it will produce the recharge otherwise it will percolate to the lower zone

$$R = IN * \frac{SM^\beta}{FC} \dots\dots\dots (3.12)$$

Where: - R is recharge [mm], IN is infiltration [mm], SM is simulated soil moisture content[mm], FC is field capacity β is a beta empirical coefficient. Direct runoff has a direct relationship with soil infiltration where simulated soil moisture is inversely related to the max soil moisture-holding capacity.

$$Cf = Cflux * [\frac{FC - SM}{FC}] \dots\dots\dots (3.13)$$

Where: - Cf is capillary firing and Cflux is correction value capillary firing. The soil water either rises upward as capillary rises or percolates to deep lower zone depend on soil available water content.

Response routine

Runoff from the groundwater boxes is calculated as the summation of two or three linear outflow equations based on the threshold SUZ [mm]. A fast, a slow, and a very slow runoff component that is influenced by recession coefficients KO, K1, and K2. The fast component Q0 contributes only to a runoff if a threshold parameter UZL in water storage is exceeded (Seibert, 2005).

$$Q_0 = K_0 * uz^{[1+\alpha]} \dots \dots \dots (3.14) \dots \dots \dots$$

Where: - Q_0 is a quick runoff, K_0 is recession coefficient and UZ is storage in the upper zone and α is Alfa (measure from non-linear of flow in the quick runoff).

Base Flow Routine; The lower reservoir zone contributes to base flow and depends on its storage depth.

$$Q_1 = K_1 * LZ \dots \dots \dots (3.15) \dots \dots \dots$$

Where Q_1 is delay/base flow, K_1 is recession coefficient and LZ is lower zone storage from the upper zone through percolation.

$$Ea = Ep \left(\frac{SM}{LP*FC}, 1 \right) \dots \dots \dots (3.16)$$

3.3.8 HBV-Light Model Input data setup

The HBV-Light model requires a time series of daily rainfall, daily temperature, and long-term potential evapotranspiration together with the daily time series of streamflow for calibration. Areal rainfalls for each of the catchments were computed multiplying by the weight of each station considered in the analysis from the Thiessen polygon. Precipitation-temperature-discharge (PTQ) file was prepared externally with format file.txt and imported into the model. Daily estimates of rainfall (mm), average daily temperature(°C) per sub-basin, and observed daily streamflow in (mm/day) for the same time baseline (1987-2010) of the two datasets together with long-term potential evapotranspiration (mm/day).

Catchment data

Since the HBV-light model works as a semi-distributed conceptual model, the whole basin is divided into different sub-basins, from a digital elevation map of the area prepared using Shuttle Radar Topography Mission (SRTM) with a resolution of 30 m x30m. The areal average rainfall was calculated as a weighted mean of precipitation stations in and around the catchment using the Theisen polygon method in equation 3.10.

$$R_{areal} = \frac{\sum_{i=1}^n R_i A_i}{A_t} \dots \dots \dots (3.17)$$

Where : R_i is the rainfall at station i A_i is the polygon area of the station i A_t is the total catchment area and n the number of stations.

For estimation of long-term potential evapotranspiration, commonly used methods are the Penman-Monteith method (Penman, 1948), the Priestley-Taylor method, (Priestley & Taylor, 1972), and the Hargreaves method (Hargreaves, 1975) with different data requirements. The Penman-Monteith method is recommended as the sole method for determining reference evapotranspiration (ET_o) when the standard meteorological variables including air temperature, relative humidity, and sunshine hours data are available (Djaman et al., 2018). However, in the study basin, those data are not available in all stations. Therefore, Potential evapotranspiration, ET_o (mm d⁻¹) was calculated by using the Hargreaves method since most of the stations have maximum and minimum temperature in all stations (Hargreaves, 1985). The FAO-56 Hargreaves equation for daily computation is given by

$$ET_o = 0.0023 [T_a + 17.8] * \sqrt{T_{max} - T_{min}} * R_a \dots \dots \dots (3.18)$$

Where: -T_{max} (°C) is the maximum daily temperature, T_{min} (°C) is the minimum daily air temperature, *R_a* (MJ m⁻² d⁻¹) is the extraterrestrial solar radiation.

$$R_a = R_s [kT(T_{max} - T_{min})]^{0.5} \dots \dots \dots (3.19)$$

Where: -R_s is extraterrestrial solar radiation [mm/day] and can be obtained from a table (Samani, 2000). *kT* empirical coefficient recommended 0.17 for semi-arid regions. The parameters *a* (mm/day) and *b* are calibrated coefficients, determined on a monthly or yearly basis by regression analysis or visual fitting.

3.3.9 Sensitivity analysis of CFSR and meteorological data

For this study, sensitivity analysis was carried out manually by changing the value of one model parameter at a time in such a way that the value of each model parameter was increased and decreased up to 60% by 20% interval and those having steep slopes are most sensitive parameters while those having moderate to gentle slopes are less sensitive. As described in Table 3.2, the Parameter range considered in sensitivity analysis for the model was recommended by (Jillo *et al.*, 2017), who previously analyzed the parameters in the specified range in this study basin. HBV-light model parameters can be grouped into (Soil Moisture Routine or volume controlling) (FC, LP, and Beta) that influence the total volume and shape controlling parameters or response routine (K0, K1, K2, PERC) that distribute the calculated discharge in time and MAXBAS (routing routine).

3.3.10 Model Calibration

Manual calibration has been carried out for optimization and development of parameter sets of the HBV-Light model. CFSR and ground observed data were individually calibrated for daily time series of 15 years (1989-2003) and validated for 7 years (2004 – 2010), while 1987 -1988 was used as a warm-up period. The manual calibration has been reported in Abdella (2016) to be the best option for turn-by-turn fine-tuning of parameters and allowing each model parameter to be changed at a time and obtain optimized parameters sets within the defined boundary conditions and therefore reducing the uncertainties, obtain the best fit between simulated and observed response characteristics, Parameters space considered in the sensitivity analysis based on the recommendation by (Jillo *et al.*, 2017), who previously analyzed the parameters of the specified range in this study area.

Table 3.2: Parameters range used in the sensitivity analysis

Table	Description	unit	Minimum value	Maximum value	References
Soil routine					Worqlul et al. (2017)
FC	maximum soil moisture storage	mm	50	750	
LP	soil moisture value above which AET reaches PET	-	0.1	1	(Seibert, 2005).
BETA	Shape coefficient	-	0.1	5	
Response routine					Hundecha & Bárdossy (2004)
UZL	Threshold parameter	mm	0	50	
PERC	threshold parameter	d ⁻¹	1	25	
K0	storage (or recession) coefficient 0	d ⁻¹	0	1	(Jillo <i>et al.</i> , 2017)
K1	storage (or recession) coefficient 1	d ⁻¹	0.01	0.4	Parajka et al.(2005)
K2	storage (or recession) coefficient 2	d ⁻¹	0.001	0.15	
Routing routine					
MAXBAS	length of a triangular weighting function	d	1	7	(Jillo <i>et al.</i> , 2017)

3.4 Model Performance Evaluation

The Nash and Sutcliff Simulation Efficiency criteria (ENS), Correlation coefficient (R^2), and Percent bias (PBIAS) were used to evaluate the performance of the HBV model in predicting the observed streamflow at each catchment for both data sets as defined in equation 3.7, 3.8 and 3.9 respectively. The ENS has been commonly used in many rainfall-runoff modeling to account for model errors in estimating the mean of the observed datasets. ENS allows the model to be compared with the initial variance that is demarcated by the observed datasets (Nash and Sutcliffe,1970). The two datasets were validated independently from 2004 to 2010 at Gilgel Gibe, Tunjo, Upper Great Gibe, Gojeb and from 1999 to 2007 at Warabessa, Walga, Wabe Guma-Denchya, and Gorombo based on the same criteria range in Table 3.2.

3.4.1 Physical Catchment Characteristics (PCCs)

Streamflow dynamics is a function of physical catchment characteristics (PCCs) and these PCCs were mainly related to land use, soil type, climate, and topography. The main point here was to evaluate the relation between model parameters and PCCs since one model Parameter may be affected by more than one PCCs. Therefore, the relations between HBV model parameters and (PCCs) were obtained from simple and multiple linear regressions together with Principal Components Analysis. The sensitivity analysis has played a role in the investigation of the relationships between model parameters and physical catchment characteristics.

For testing the regionalization approach the range of physical catchment attributes were collected. These physical catchment characteristics (PCCs) used in developing regression equations were shown in Table 3.3 and the indices of PCCs for each catchment are detailed in Appendix C. The PCC collection was inclusively based on previous regionalization studies in the study area and the availability of the data in the basin. Nineteen catchments including ungauged catchments each consisting physical characteristics of seven topographical descriptors, three climatic factors, three soil descriptors, four vegetation descriptors (see Appendix D).

Table 3.3: Physical Catchment Characteristics (PCCs) used in the regionalization

PCC	Unit	Symbol
Physiographical descriptors		
Digital elevation model	m	DEM
Average catchment slope	%	S
Catchment area	Km ²	DA
Aspect	-	Aspect
Topographic wetness index	-	TWI
Flow direction	-	FD
Flow accumulation	-	FA
Length of longest flow path	KM	LLP
Climate descriptors		
Average Annual Rainfall	mm	MAR
Average Annual Potential Evapotranspiration	mm	PET
Soil descriptors		
Average Soil Available Water Content	cm/m	S_AWC
Average Soil Bulk Density	Kg/dm ²	S_BD
Average Soil Permeability	m/s	SAT_K
Land cover descriptors		
Average Normalized Difference Vegetation Index	-	NDVI
Percent Cultivation land	%	% CL
Percent Forest land	%	% FL
Percent Grass land	%	% GL
Principal components from PCA		
Principal Component 1	-	PC1
Principal Component 2	-	PC2
Principal Component 3	-	PC3
Principal Component 4	-	PC4
Principal Component 5	-	PC5
Principal Component 6	-	PC6

3.4.2 Principal component analysis (PCA)

PCA has been carried to the correlation matrix of the above PCCs in order to obtain Principal components and reduce any insignificant dimensions in the PCCs through the following steps:

Step 1, Data standardization: The indices of PCCs were first standardized (zero mean) to ensure scale comparability and allow for the equal contribution of each variable in the transformation of the data obtained from measurements of different units.

Step 2, Investigation of the fitness of the PCCs for PCA:

The existence and dominance of common hydrological features that existed among the datasets were assessed through the analysis of covariance matrix using Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy and the ratio of partial correlation to multiple correlations.

Step 3: The sphericity of data was checked using Bartlett’s method to test the suitability of the data for PCA in such a way that the factors can reveal the variability found in the data for a significance level of 95% ($p < 0.05$) and therefore the rejection of the null hypothesis. The table in appendix E also shows the inter-correlation matrix among some physical catchment characteristics and it’s worthless to directly regress the PCCs on the model parameters without solving the multicollinearity problem using the PCA application. From table 3.3, the KMO scores 0.728 with Bartlett’s test of sphericity at 95% significance level and showing the fitness of the dataset for PCA (there existed collinearity among the datasets).

Table 3.4: KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.728
	Approx. Chi-Square	152.311
Bartlett's Test of Sphericity	df	78
	Sig.	.000

Extraction of underlying hydrological features: PCA was then carried out to obtain the new principal components through varimax rotation using the Kaiser’s eigenvalue standards (principal components with eigenvalues greater than 1.00) were extracted and reserved for improving the regression model as the corresponding eigenvectors of these eigenvalues show the direction of the data that

describe the maximum amount of variance and the lines that detect most abundant hydrological information of in the data.

Therefore, PCA outputted from varimax rotation indicated six principal components with an eigenvalue greater than 1.00 which accounted for the total cumulative variance of 81%. The contribution of these variables is shown by their eigenvalues in table 3.5. The first principal component (PC1) contributes the maximum percent of the cumulative variance in the data (31.5%) see table 3.5, The second principal component (PCII) contributes 16.5%, PCIII contributes 14. % and proceeding until the last Principal component (PC6) which contributes the minimum percent of cumulative variance (6.2%). Therefore, this is the way how PCA was applied to find the most effective principal components which improve the regional model.

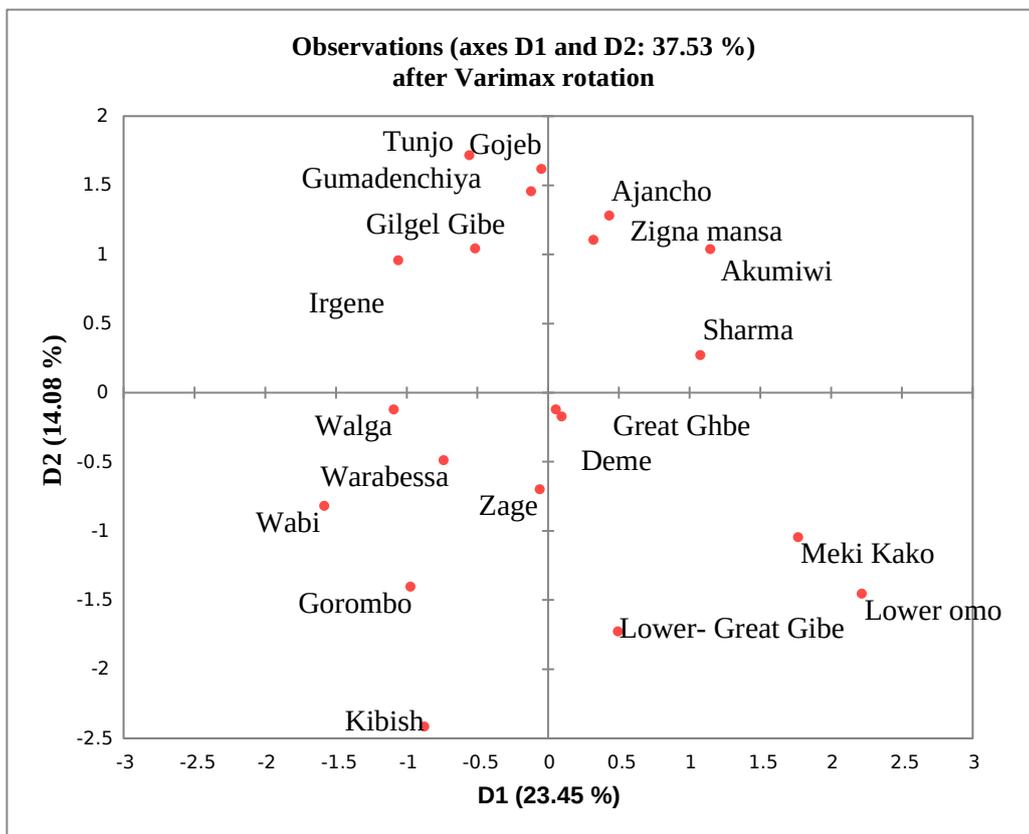


Figure 3.14: Group of homogenous catchments classified using PCA

PCA was also applied to find hydrologically homogenous groups of catchments as shown in Figure 3.14 which could help to group homogenous catchments in the runoff modelling.

But in this study, all the gauged catchments were found the same group and ungauged catchments were found in other group and became very difficult to extrapolate model parameter with in such hydrologically heterogenous catchments.

Table 3.5: Eigen analysis of the Correlation Matrix

Eigenvalue	5.349	2.8	2.4017	1.5708	1.3238	1.0558
Proportion	0.315	0.165	0.141	0.092	0.078	0.062
Cumulative	0.315	0.479	0.621	0.713	0.791	0.853

3.5 Regionalization

In Omo-Gibe River Basin, most of the gauged catchments are located upstream while ungauged catchments are located downstream and the attributes of nearby catchments are still extremely varied which restricted applying the spatial proximity method and similarity of physical catchment characteristics. The study area is highly heterogeneous comprising different hydroclimatic, geological, and topographical properties which will not generate similar hydrological responses. Catchments in the basin for which flow time series are to be estimated have no comparable areas and comparable PCCs. thus, difficult to extrapolate using methods other than the regional model method.

3.5.1 Developing Regional Regression Equation

HBV model has eight parameters for which regression equations were developed, FC, LP, BETA, K₀, PERC, K₁, K₂ and MAXBAS. The regional model was developed from each parameter that best correlated to some of the indices while the other parameter correlates to other indices. Before the regionalization of model parameters to ungauged catchments, HBV-Light model parameters were first optimized at each gauged catchment as described in section 3.3.8. These parameters were then subjected to be expressed as a function of the physical catchment characteristics influencing rainfall-runoff processes in the regression model. Hydrologically acceptable and statistically optimized correlation between model parameters and catchment features was assessed from the stepwise regression. The catchment indices used in developing the model are shown in appendix D. The multiple regression model was written as:

The multiple regression model was written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n$$

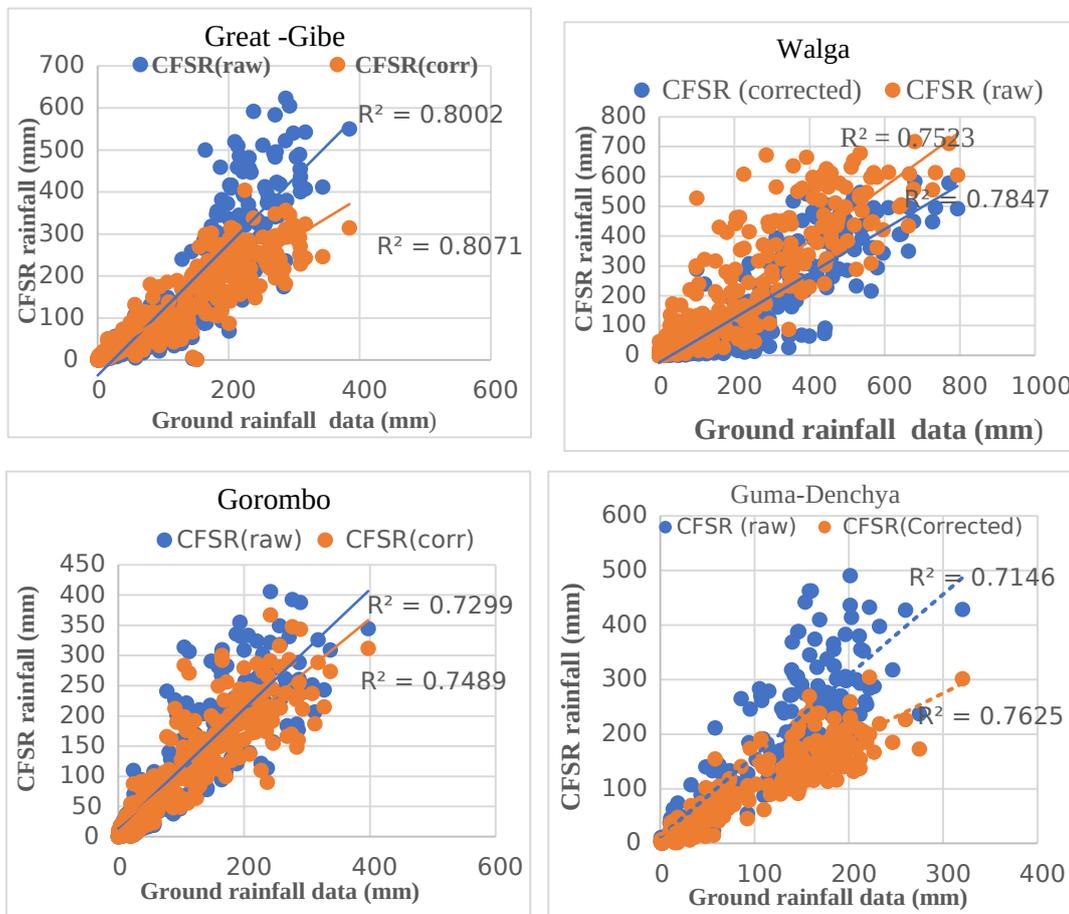
Where Y is the predicted model parameter, $X_1, X_2, X_3, \dots, X_n$ are the set of catchment characteristics, β_0 the intercept of the regression line, $\beta_1, \beta_2, \beta_3, \dots, \beta_n$ the coefficients of catchment features. The overall relation and strength of the above regression equation were tested and selected based on the coefficient of determination, the p-value, and the student's t-test.

CHAPTER 4 RESULTS AND DISCUSSION

In this chapter, the findings of the study were presented for each of the objectives proposed in the general objective section. The results were described step-wise with a brief discussion in reference to other scientific works.

4.1 Comparison of the areal CFSR and ground observed datasets

The correspondence between the areal CFSR precipitation estimates and ground observed data were analyzed for each catchment (see Figure 4.1.1.) assuming the ground observed estimates to be the truth. The results of a comparison using the statistical tests as Coefficient of determination (R^2), The Nash-Sutcliffe efficiency (ENS) and the categorical tests, the Probability of Detection (POD), the False Alarm ratio (FAR) were listed in table 4.1 and 4.2 respectively.



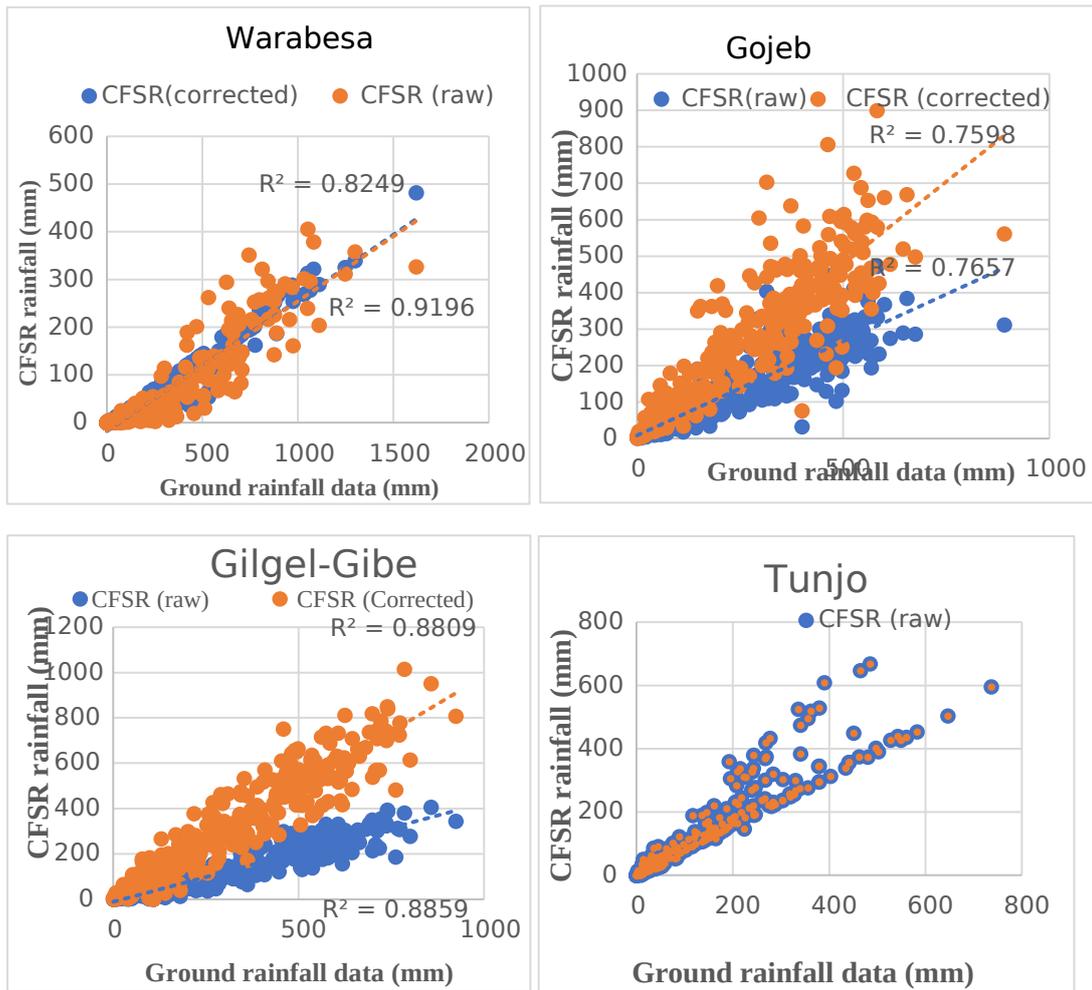


Figure 4.1: Monthly comparisons of the two rainfall datasets (1987 - 2010)

CFSR data reflected more than 70% of the observed variance in daily rainfall of ground observed data at Tunjo, Upper Great-Gibe, Gilgel-Gibe, Gojeb, and Guma - Denchya. At Walga, Wabe, Gorormbo, and Sokie-Waybe catchments. The daily performance of CFSR have shown reasonable correlation to the spread of the observed rainfall. At Demie and Zage catchments, the CFSR however poorly correlated to the observed data for their poor records of the ground observed rainfall data have affected the performance analysis. The application of linear scaling bias correction here slightly improved the efficiency of CFSR data as shown in Figure 4.1. Therefore, CFSR data have generally shown good performance in representing the weather condition of the study basin comparatively to the studies of (Berhanu et al., 2016). The difference is the way of comparison as they evaluated CFSR data based on the point-to-point comparison, but in this study, the CFSR data was evaluated based on

an areal basis. CFSR rainfall estimates are continuous and represent areal rainfall while ground observed rainfall data was obtained from a station measurement at a particular point in location, the use of point data to evaluate gridded climate products have great variation.

Table 4.1: The comparison using statistical metrics

Catchments	Raw. CFSR			Corr. CFSR		
	ENS	R ²	PBIAS (%)	ENS	R ²	PBIAS (%)
Upper Great- Gibe	0.8	0.75	-15	0.85	0.84	-11
Warabessa	0.65	0.66	-13	0.68	0.69	-9
Tunjo	0.92	0.91	0.18	0.93	0.95	0.1
Gilgel-Gibe	0.73	0.70	-17	0.74	0.75	-8
Lower Great Gibe	0.68	0.65	-15	0.70	0.71	-12
Walga	0.71	0.71	11	0.75	0.72	9
Wabe	0.65	0.65	-12	0.68	0.69	-9
Guma-dincha	0.75	0.83	12	0.76	0.78	11
Sokie-waybe	0.69	0.67	-7	0.65	0.69	- 8
Gorombo	0.55	0.6	-20	0.65	0.7	-17
Gojeb	0.7	0.7	-12	0.72	0.7	-7
Demie	0.6	0.62	11	0.65	0.7	7
Zage	0.54	0.56	-15	0.57	0.59	-11

Table 4.2: The comparison using categorical metrics

Catchments	Raw CFSR		Corr.CFSR	
	POD	FAR	POD	FAR
Upper Great- Gibe	0.90	0.02	0.90	0.02
Tunjo	1.00	0.03	1.00	0.03
Warabessa	0.80	0.30	0.80	0.30
Gilgel-Gibe	0.75	0.10	0.75	0.10
Lower Great Gibe	0.80	0.20	0.80	0.20
Walga	0.91	0.04	0.91	0.04
Wabe	0.90	0.04	0.90	0.04
Guma-dincha	0.70	0.00	0.70	0.00
Sokie-waybe	0.50	0.10	0.98	0.10
Gorombo	0.70	0.10	0.98	0.10
Gojeb	0.80	0.08	0.90	0.08
Demie	0.67	0.15	1.00	0.15
Zage	0.60	0.10	1.00	0.10

CFSR showed good quality in Rain Detection (high POD) and the False Alarm Ratio falls in the range of 0.01 to 0.3. The southern catchments (Lower - Omo), where CFSR has shown a slight overestimation (high FAR). Lower – Omo is a floodplain with a hot arid climate where the dry mass of air could result in high lifting condensation levels. Raindrops have to pass via a warm atmosphere before reaching the ground and cause substantial evaporation at a potential level in the dry atmosphere underneath the cloud base and result in high FAR. Thus, although these clouds may not produce rain at the surface, the satellite sensors may detect rain, leading slightly to the observed large FAR values.

Therefore, The R^2 and ENS value between the ground observed rainfall and CFSR monthly average precipitation was larger than 0.60 at all nine meteorological catchments from 1987 to 2010, which showed that the CFSR weather data provide accuracy to build the HBV hydrological model.

4.2 Runoff simulation results in the Omo-Gibe River Basin

The performance of CFSR data in predicting stream flow through calibration and validation to the HBV-light model was tested. The daily runoff simulations at nine gauged catchments and the performance indicators are shown in table.4.5.

4.2.1 Sensitivity analysis of HBV-Light model

LP, and BETA show a significant effect at all catchments based on both performance indicators (ENS and R^2). These parameters were highly sensitive and a slight change in their value impacted in reshaping the hydrographs changing the volumes and peak flows. Fewer sensitive parameters were FC, K1 and MAXBAS. Insensitive parameters were PERC and K0. Similar results were observed in the study of Temesgen, (2019) where FC, LP, and BETA were found the sensitive parameters. K1 and MAXBAS were moderately sensitive for their optimized parameters. It's also observed that the sensitivity level of a model parameter varies from catchment to catchment. i.e., the percentage of sensitivity of parameters for each catchment is different in relation to the physical catchment characteristics of each catchment. The optimized model parameters of CFSR data and ground observed data have comparable values except for FC and LP.

Figure 4.3 shows the sensitivity analysis of both CFSR and ground observed rainfall data at Tunjo, Gojeb, and Gilgel-Gibe catchments, where evapotranspiration is high and therefore LP is a highly sensitive parameter. CFSR predicted the water balance components at Gojeb catchments, that a substantial portion of the moisture is lost by the evapotranspiration and resulted in less runoff and changes the efficiency of the model by reducing the simulated mean annual runoff as the soil moisture is lost by evapotranspiration. In connection to LP, FC and BETA are also sensitive parameters as they all are in the soil routine.

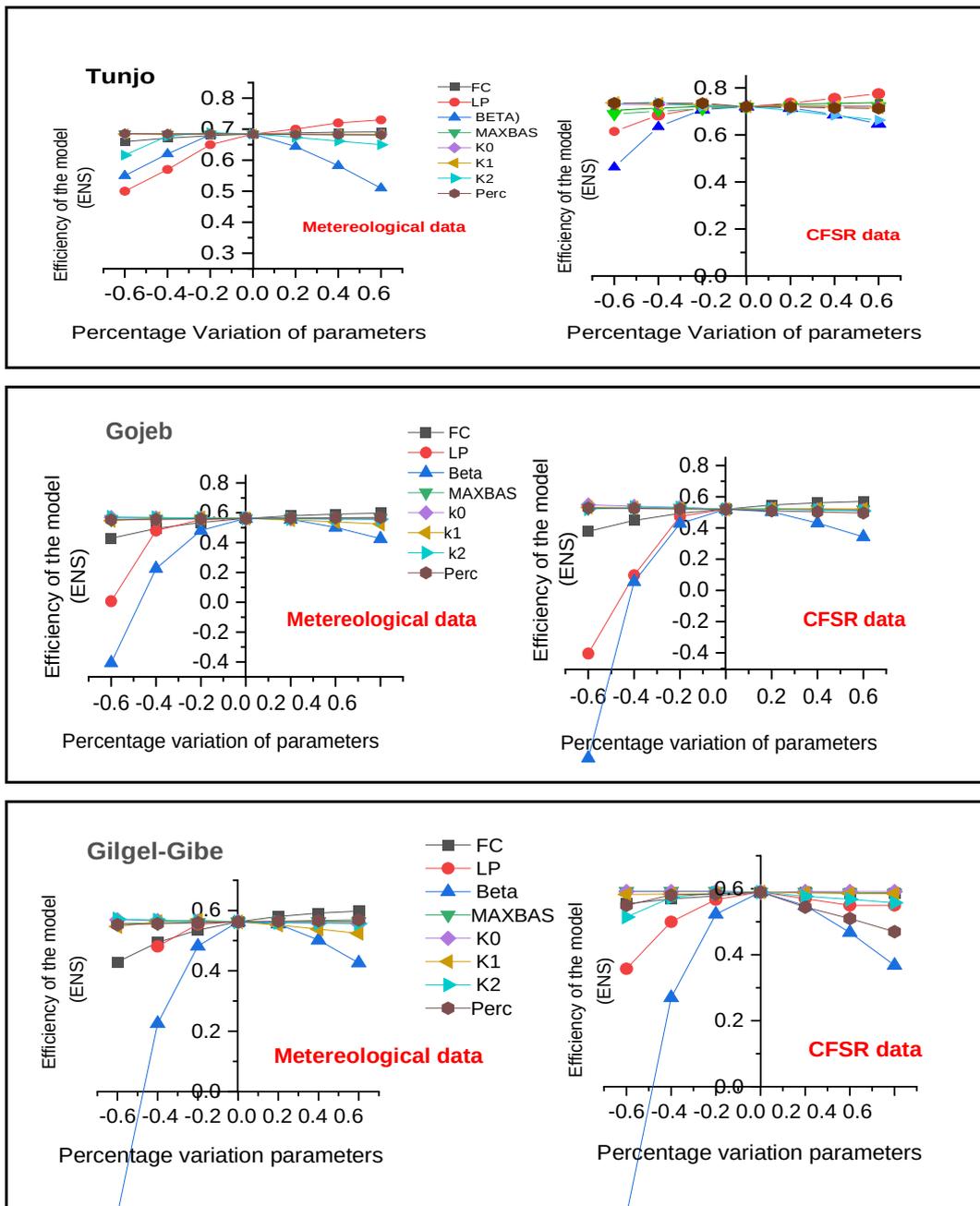


Figure 4.2: Sensitivity analysis

The optimized parameter values of CFSR data and ground observations in the simulation of flow are shown in Table 4.3 and Table 4.4 respectively.

Table 4.3: Optimized parameter values at gauged catchments (1989-2006).

Catchments	The optimized model parameter sets of ground rainfall data-driven HBV model								
	Model parameters	FC	LP	Beta	PERC	UZL	k0	k1	k2
Great Gibe @Abelty	300	1	1	2	20	0.2	0.1	0.05	1
Walga	300	0.8	2	5	20	0.1	0.1	0.1	7
Warabessa	250	0.01	2	20	5	0.05	0.05	0.09	4
Tunjo	350	1	1	4	20	0.001	0.02	0.02	1
Wabe	130	1	0.2	2	15	0.05	0.05	0.05	1
Gorormbo	180	0.45	1	5	10	0.045	0.045	0.05	5
Gilgel-Gibe	550	0.8	2.5	5	50	0.05	0.04	0.05	7
Sokie-weybo	450	0.4	2.5	10	45	0.05	0.1	0.1	5
Demie	520	1	1.5	5	15	0.05	0.02	0.02	4
Guma @ Dincha	450	1	1	3	20	0.05	0.05	0.05	7
Gojeb	500	1	0.7	15	25	0.05	0.05	0.1	4

Table 4.4: Optimized model parameter from CFSR data

Catchments	Best-fit model parameter sets of CFSR data-driven HBV model								
	Model parameters	FC	LP	Beta	PERC	UZL	k0	k1	k2
Great Ghibe nr Abelty	350	0.4	1.2	3	25	0.05	0.045	0.045	6
Walga	350	0.7	4	5	30	0.01	0.02	0.035	4
Warabessa	300	0.5	0.5	5	10	0.3	0.02	0.06	7
Tunjo	400	1	0.6	5	20	0.045	0.045	0.05	5
Wabe	300	1	1.5	1	20	0.2	0.1	0.05	5
Gorormbo	225	0.5	1.5	5	20	0.04	0.04	0.04	4
Gilgel-gibe	650	0.7	2.5	5	10	0.04	0.035	0.035	5
Sokie-Weybo	450	0.5	2.5	5	25	0.05	0.1	0.05	4
Demie	400	1	2	5	20	0.01	0.01	0.01	5
Guma -dincha	500	1	1	3	45	0.2	0.1	0.05	7
Gojeb	650	0.8	1	5	20	0.045	0.01	0.15	6.5

The field capacity (FC) of the calibrated model using CFSR data was slightly larger than the FC value of the model calibrated by gauged rainfall data in all catchments. This could be due to higher rainfall estimation in the CFSR data. Similar results were observed in sensitivity analysis of CFSR data-driven model in the study of Worqlul et al., (2015) that the soil retained greater quantities of water and released it afterward by evapotranspiration and base flow compared to the gauged flow simulation as compared to the gauged. Hence, to reduce parameter uncertainty in the calibration of the model, great attention has been given for the sensitive parameters during the establishment of a regional model which addresses all aspects of the hydrographs in addition to objective functions

4.2.2 HBV-Light model Daily calibration and validation results

HBV-Light model was calibrated at gauged catchments where streamflow has been well updated for calibration of the model. The model has remarkably predicted the observed runoff using CFSR rainfall as input at Upper Great Gibe, Tunjo, Gilgel-Gibe, and Gojeb.

CFSR driven model was able to simulate the water balance components as well as the peak flows better than the observed rainfall-driven model. This was previously expected and described in subsection 4.1.1 as CFSR data captured more than 70% of the variance of the ground observed data even the raw CFSR (without bias correction) have resulted in good efficiency. Figure 4.2.2 and Figure 4.2.3 show a sample daily calibration and validation result of the HBV model at Tunjo and Gilgel-Gibe catchment.

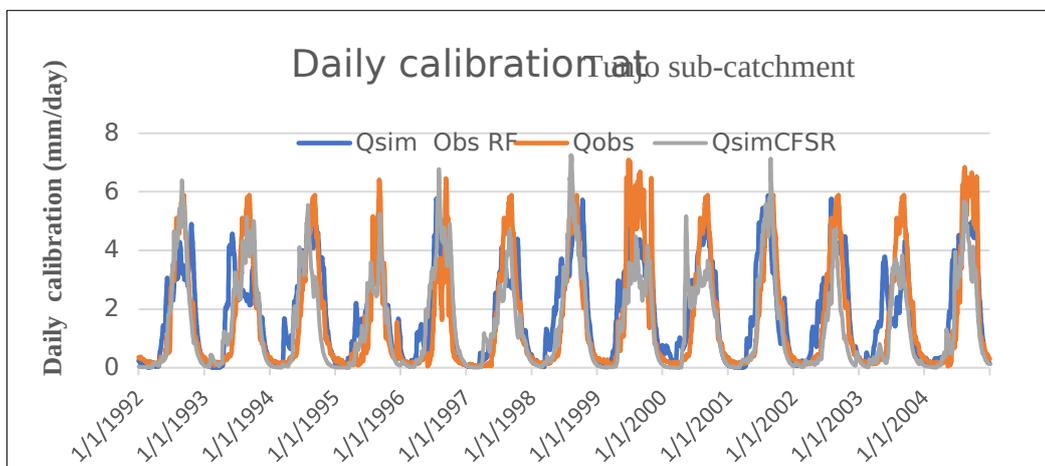


Figure 4.3: Daily calibration @ Tunjo sub-catchment (1992 -2004)

The Great-Gibe watershed is very large and covering a drainage area of 15746km² with a flow station near Abilty. CFSR was able to reproduce the observed flow in such a wide watershed with minimum underestimation of the peak flows which actually occurred by the rise of water level due to sedimentation problem but still, CFSR data predicted the abrupt fluctuation of the peak flows better than the gauged rainfall data at this catchment.

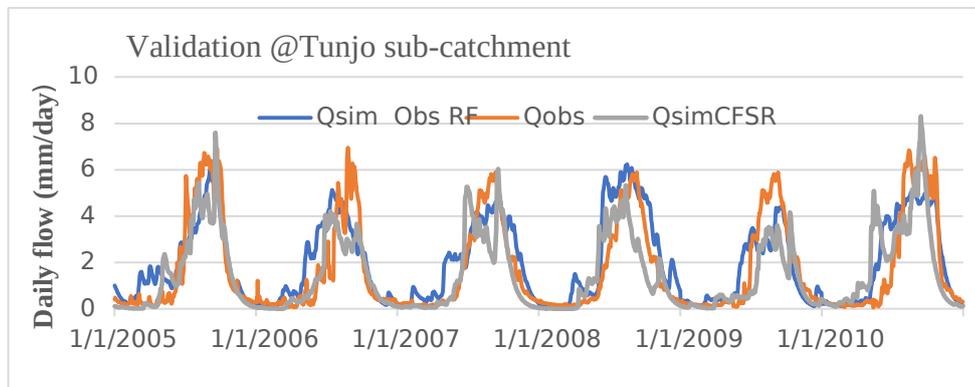


Figure 4.4 Daily validation of CFSR data at Tunjo (2005 -2010)

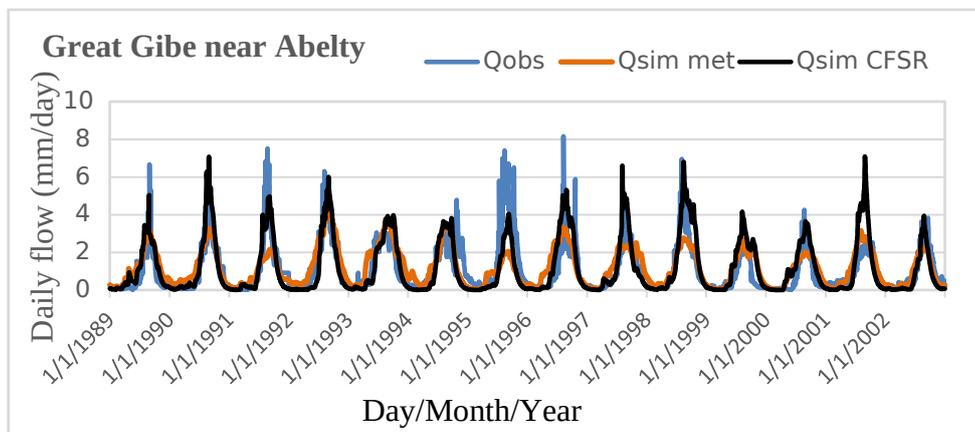


Figure 4.5: Daily calibration at Great Gibe

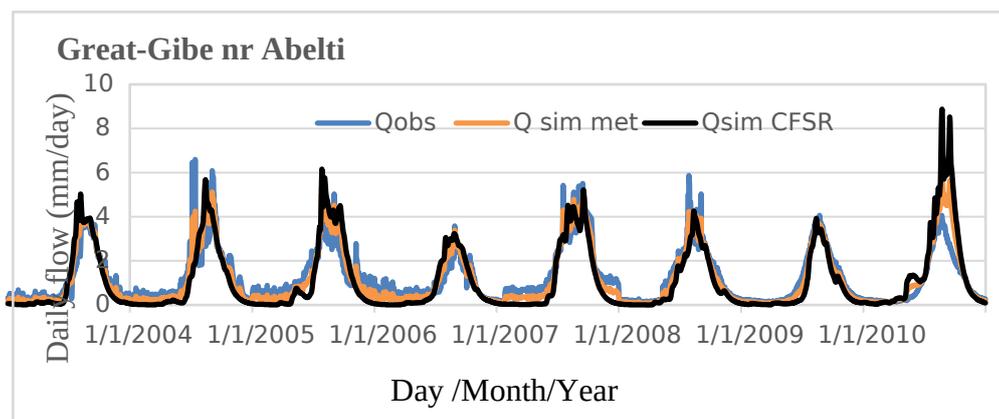


Figure 4.6: Daily validation at Great Gibe near Abelti

The result of HBV model performance in streamflow simulation from CFSR data and ground observed rainfall data were described in table (4-4) and table (4-5) based on objective functions ENS, R², and PBIAS during the calibration period of (January 1-1989 to December 31- 2003) and validation period of (January 1- 2004 to December 31- 2010) respectively. In some catchments, due to data unavailability, the calibration was altered to be started in 1990 and 1992.

Table 4.5: Result Summary of daily flow simulations using HBV model

Performance of HBV model using meteorological data as an input						
Catchments	Calibration			Validation		
	R ²	P BIAS	ENS	R ²	P BIAS	ENS
Great -gibe nr Abelty	0.72	-6.73	0.71	0.75	-5.123	0.69
Tunjo	0.85	3.07	0.85	0.7	-1.75	0.7
Walga	0.66	-11.57	0.65	0.65	-3.05	0.67
Warabessa	0.68	7.2	0.66	0.67	7.5	0.68
Wabe	0.67	-2.37	0.65	0.69	1.6	0.69
Gorormbo	0.65	-10	0.64	0.68	-10	0.69
Gojeb	0.77	6.79	0.76	0.75	9.85	0.75
Gilgel-Gibe	0.8	-9.42	0.79	0.83	-9	0.85
Sokie-weybo	0.48	6.62	0.45	0.4	10.65	0.42
Guma @ Dincha	0.67	-5.54	0.67	0.67	13.95	0.66
Demie	0.5	-15.49	0.5	0.42	-10.95	0.39

Table 4.6: Summary of HBV model daily simulation result

Performance of HBV model using CFSR data as an input						
Catchments	Calibration			Validation		
	R ²	P BIAS	ENS	R ²	P BIAS	ENS
U/Great- Ghibe nr Abelty	0.75	7.05	0.7	0.74	9.53	0.7
Tunjo	0.85	-1.1	0.86	0.84	-1.51	0.82
Walga	0.68	-8.57	0.69	0.68	-5.05	0.66
Warabessa	0.66	10	0.67	0.66	-9.23	0.65
Wabe	0.68	-9.07	0.65	0.67	5.13	0.67
Gorormbo	0.66	-5.32	0.65	0.65	-11.49	0.66
Gojeb	0.77	6.8	0.76	0.78	5.48	0.78
Gilgel-Gibe	0.85	10.17	0.8	0.82	10.85	0.83
Sokie-weybo	0.55	11.23	0.52	0.43	11.65	0.415
Guma @ Dincha	0.65	8.01	0.66	0.67	9.5	0.7
Demie	0.51	-7.49	0.5	0.48	-5.95	0.47

The line plot of CFSR and ground observed rainfall data at all catchments (see Appendix A and B) were prepared for the comparison of the simulated and observed streamflow and indicating the hydrological trend in the data sets (between the simulated from CFSR data and observed flows, and the ground observed data).

4.3 The established Regional Model

For developing a regional model, the optimized parameters at the gauged catchments with relatively good performances during calibration were used. The catchments (Ajancho, Demie, Zage, and Irgine) were screened out from establishing the regional model as their result of the calibration was poor. This could happen be due the sparse coverage in the density of ground observation stations calibrated model parameters could not represent the watershed hydrology and consequently made the model poorly perform.

Fourteen physical catchment characteristics (Appendix D of Table 1, Table 2, and Table 3): three climate descriptors, three soil descriptors, four vegetation descriptors, and four topographical descriptors were prepared to establish the regional model. The correlation of each parameter with the above catchment indices was tested. From the stepwise regression of each group, by observing the correlation coefficient and the physical relevancy of each index to each parameter, those catchment indices that showed good correlation and those that are hydrologically relevant to each model parameter regarding catchment response, though didn't show good correlation, were selected and regressed over each model parameter. The final regression equation was selected after testing its statistical significance through R^2 , t-test and p-value, and hydrological meaning. Therefore, the linear regression equations relating each model parameter to the best describing catchment characteristics were shown in Table 4.10. There are eight model parameters for which regression equations were developed, FC, LP, BETA, ALPHA, PERC, Alpha, K1, K2, and MAXBAS. Accordingly, the maximum water holding capacity (FC) was best correlated with flow direction and PC1, Limit of evapotranspiration (LP) was correlated with PC5. The shape coefficient BETA was best correlated with TWI, Quick runoff coefficient, K0 was best correlated with soil bulk density and PC6. Storage coefficient at upper zone, K1 was best correlated with PC1, PC2, and PC3, likewise, K2 was best correlated with drainage area and NDVI, MAXBAS was best correlated with PC1 and mean

elevation.

Outputs from PCA applying the varimax rotation indicated that six principal components with an eigenvalue greater than 1.00 (see Appendix D, Table 4) accounted for the total cumulative variance of 81% as per their eigenvalues. The first component has explained about 31 percent of the variance in the catchment characteristics, while the second component explained 16.5 percent, and therefore 85 percent of the variance is explained by only the six components. The classification has also indicated that the catchment behavior is more influenced by soil, topographical, and climate descriptors in the first component.

Table 4.7: The regional model developed for MPs and PCCs links.

FC = 812.83 - 28.1837 FD + 25.60 PC1					
MP	Coefficients	Error	Stan	t Stat	P-value
Intercept	812.8333	94.25109	8.624126	0.000134	
FD	-28.1837	8.996605	-3.72133	0.020255	
PC1	25.60588	23.62041	1.084057	0.019967	
LP = 0.7235 + 0.2407 PC5					
Intercept	0.723554	0.062706	11.53879	8.27E-06	
PC5	0.240758	0.102765	2.342797	0.051632	
BETA = 8.4228 - 0.96372 TWI					
Intercept	8.422807	0.994459	8.469741	6.32E-05	
TWI	-0.96372	0.134452	-7.16774	0.000183	
Alpha = -0.152825 + 0.000165 ME + 0.000093 PET - 0.0884 PC3					
Intercept	-0.15283	1.294379	-0.12226	0.907455	
Elevation	0.000165	0.00019	0.865134	0.426516	
PET	0.000093	0.000607	0.154376	0.883352	
PC3	-0.0884	0.048596	-1.8282	0.127062	
Ko = 6.8683 - 5.24401 S _{BD} - 0.03425 PC6					
Intercept	6.868397	1.958284	3.507355	0.012712	
S _{BD}	-5.24401	1.521305	-3.44705	0.013682	
PC6	-0.03425	0.030698	-10.1159	0.007163	
K1 = 0.374 + 0.11738 PC1 - 0.073011 PC2 - 0.08477 PC3					
Intercept	0.373885	0.320551	1.166382	0.096062	
PC1	0.117388	0.197502	0.59436	0.057813	
PC2	-0.07301	0.181283	-0.40273	0.070378	
PC3	-0.08477	0.174567	-0.14861	0.047775	
K2 = 0.034 + 0.0000108 DA - 0.02565 NDVI					
Intercept	0.034586	0.08331	0.415145	0.192472	
DA	1.08E-05	6.43E-06	1.677457	0.014446	
NDVI	-0.02565	0.156141	-0.16429	0.087489	
MAXBAS = 12.24 - 0.00304 PC1 - 0.0035 ME					
Intercept	12.24557	2.75224	4.44931	0.004332	
PC1	-0.00304	0.217258	-5.014	0.098928	
ME	-0.0035	0.001475	-2.37406	0.055214	

4.4 Validation of the Regional Model

For estimating model parameters of ungauged catchments, the regional model has to be tested for its performance initially. The approach of validating the regionalized model consists of applying the regional relationship between the model parameters and the catchment descriptors derived in the calibration set of catchments to the independent catchments within the study area that were not used to derive the regional relationship. The objective functions (ENS and R^2) defined during the calibration period were used to compare the predicted discharge from the regional model with the observed streamflow. Summary of the performances of the regionalized model during the validation period were shown in Table 4.9. A sample hydrograph comparing the observed and simulated discharges of the regional model which have been tested at Lower Great Gibe catchment is shown in

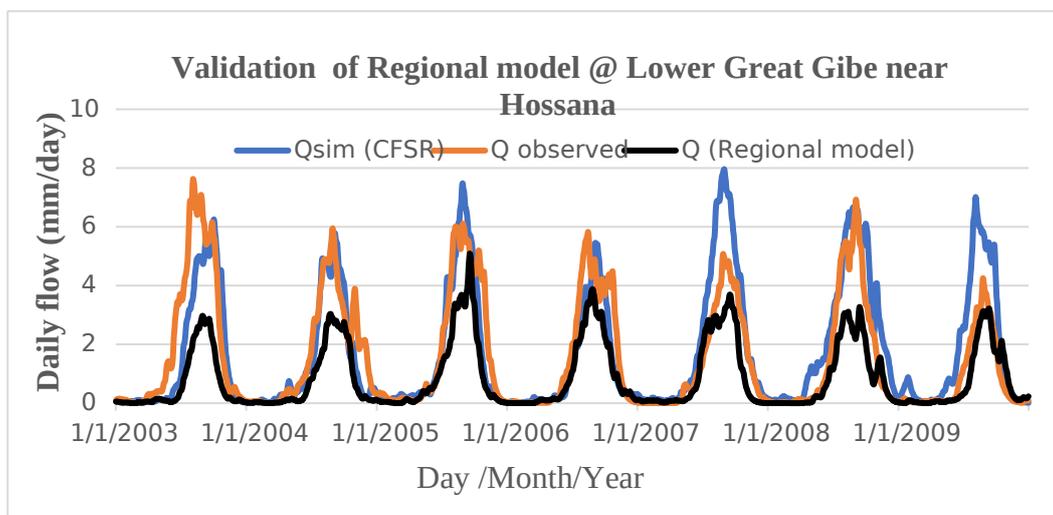


Figure 4.7: Comparison of the stream flow simulation

The performance of the regionalized model in the validation catchments in terms of the Nash-Sutcliffe coefficient show only a slight difference from its performance during the calibration period. This could happen due to the fact that the ranges in the validation catchment descriptors used to validate the regional model found within the corresponding ranges of the descriptors in the of catchments calibration set. However, at Nerie gauge near Jinka town (Lower - Omo) where some of the descriptors (mean annual rainfall, mean annual evapotranspiration, mean elevation, normalized difference vegetation index, soil moisture) were vary from the calibration set but still the performance of the regionalized model in such catchment was found to be satisfactory showing as one

option for applying the regional relationship for extrapolation of the catchment descriptors outside the range within which they were derived.

Table 4.8: Model parameters predicted at the gauged catchments

Model Parameter	FC	LP	BETA	ALPHA	k0	K1	k2	MAXBAS
Upper Great Gibe	344.25	0.55	1.2	0.22	0.03	0.15	0.08	6
Tunjo	327.69	0.79	0.62	0.14	0.07	0.02	0.05	5
Walga	523.54	0.77	3.8	0.1	0.07	0.01	0.05	5
Wabe	272.38	0.92	1.39	0.25	0.1	0.14	0.05	4
Gorormbo	280.69	0.86	0.91	0.46	0.06	0.17	0.03	4
Guma- Dincha	594.22	0.92	1.2	0.13	0.21	0.11	0.06	7
Gilgel-Gibe	591.73	0.6	1.68	0.19	0.01	0.28	0.08	5
Gojeb	613.3	0.65	0.91	0.26	0.09	0.22	0.09	6
Warabessa	577.2	0.53	0.52	0.2	0.29	0.23	0.04	6

Table 4.9: Validation of the regionalized model at validation catchments (2008 - 2010)

Performance Summary of the regionalized model at validation catchments				
Gauging stations	Area (sqm)	Catchments	ENS	R ²
Upper Great Gibe near Baco	288.1	Upper Great Gibe	0.52	0.65
Gibe near Seka	280.4	Tunjo	0.66	0.8
Lower Great Gibe near Hossaina	4228	Lower Great Gibe	0.64	0.79
Kulit near Tedele	350	Walga	0.5	0.63
Sheta near Bonga	190.4	Gojeb	0.57	0.69
Nerie near Jinka	166	Lower - Omo	0.55	0.65

4.5 Flow estimation at the ungauged catchments

After validating the regionalized model at the gauged catchments, the established regional model was used to predict the model parameters of ungauged catchments using their corresponding PCCs. A total of eleven ungauged catchments (Lower Great Gibe, Demie, Zage, Irgene, Upper-Omo, Sharma, Middle Omo, Mey-Kako,

Kibish, Lower-Omo) representing 45 % of the basin area. Three of which are catchments with the unsatisfactory result during calibration while the remaining are totally ungauged and no hydrological data at the outlet. The aim of correlating catchments characteristics to model parameters and developing regional regression equation was described in detail above and the predicted hydrograph for each ungauged catchment was shown in (Appendix C) using the established regional model.

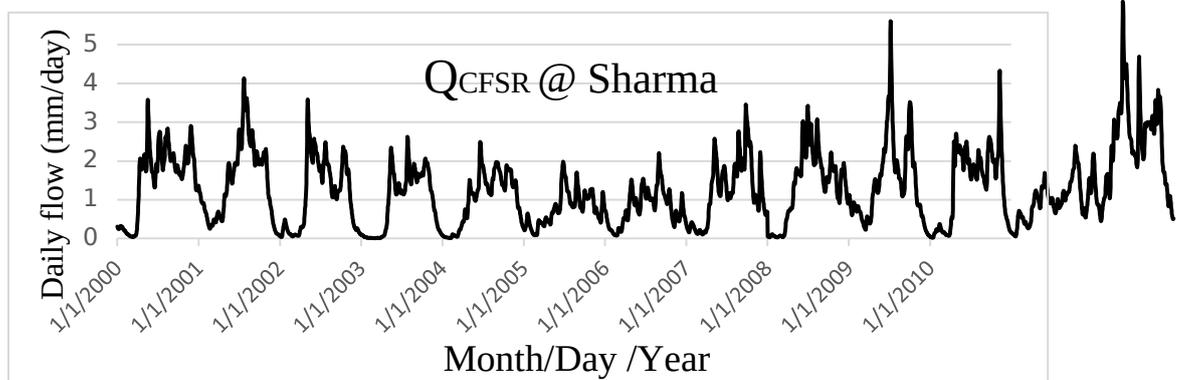


Figure 4.8: Hydrograph at the Sharma catchment (Ungauged catchment)

Table 4.10: Model parameters derived for ungauged catchments

Catchments	FC	BETA	LP	ALPHA	k0	K1	K2	MAXBAS
L/great gibe	393.34	1.49	0.75	0.23	0.56	0.25	0.07	4.84
Demie	546.65	3.76	0.11	0.53	0.85	0.22	0.04	5.00
Zage	480.74	0.86	0.37	0.56	0.85	0.29	0.05	3.92
Irgene	520.45	1.53	0.75	0.62	0.53	0.22	0.04	6.44
ajancho	365.54	0.68	0.86	0.50	0.56	0.50	0.05	6.48
Upper-Omo	489.59	1.32	0.90	0.35	0.50	0.33	0.08	5.80
Sharma	665.47	0.70	0.80	0.07	0.14	0.33	0.07	5.48
Middle-Omo	582.90	0.04	0.56	0.10	0.68	0.48	0.09	3.75
Kibish	543.83	1.39	0.55	0.26	0.08	0.46	0.07	4.01
Meki-kako	256.95	0.23	0.96	0.26	0.06	0.51	0.08	4.33
Lower-Omo	548.16	0.52	0.44	0.19	0.36	0.34	0.13	3.07

CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The main focus of this study was evaluating the performance of wide area covering satellite rainfall product, Climate Forecast and System Reanalysis (CFSR) data as an input to a semi-distributed hydrological model (HBV-Light) and flow estimation at ungauged catchments using CFSR data driven HBV-Light model in Omo-Gibe River Basin, Ethiopia.

Initially, the correlation of the CFSR and the meteorological data was evaluated at each catchment based on ENS, \bar{R} , PBIAS efficiency criteria. The rainfall occurrence detection capacity of CFSR data was also tested using categorical descriptors (POD and FAR).

From the direct comparison performance evaluation, it could be concluded that the daily areal estimates of CFSR data were observed to better correlate to the ground observed average areal rainfall data at well-recorded catchments. The runoff predictive performance of the CFSR data driven HBV- Light model was validated to the observed discharges at gauged catchments and concluded to be a good performance model.

Based on the performance evaluation statistics used (R^2 , NSE, and PBIAS), the CFSR data well simulated the daily streamflow as the ground observed rainfall data at the Upper Great Gibe near Abelty, Tunjo, Gilgel-Gibe, Gojeb, and Walga catchments. At Guma-Denchya, CFSR data driven HBV-Light model better simulated the streamflow than ground observed rainfall data driven HBV-Light model. Therefore, it could be concluded that for better performance of HBV-Light model using CFSR data, the model parameters should be optimized through CFSR data itself (Table 4.5) rather than copying the optimized set of model parameters from ground observed rainfall data driven HBV-Light model (Table 4.6), that is called direct calibration and validation of the HBV-Light model using CFSR data.

The simulation by the CFSR data driven HBV-Light model was able to capture the peak flows better than the runoff simulation by the gauged rainfall. So, the CFSR data can be suitable to predict extreme events when using HBV models.

Therefore, it could be concluded that CFSR data can be the best alternative data for River Basin modeling,

The performance of daily streamflow simulation from the CFSR climate data at larger watersheds (Great Gibe with drainage area of 15746 km²) demonstrated that CFSR data can simulate the stream flow at the larger catchments using HBV-Light model. The observed versus simulated streamflow hydrograph analysis at each catchment using HBV-Light model (Appendices I and II) indicated that CFSR data can predict the seasonal patterns of rainfall occurrences and the timing (the unimodality of the rainfall regime at the northwestern catchments and the bimodality of rainfall in the southern (ungauged catchments)).

The CFSR calibrated model can be used for further analysis of the streamflow responses to land use and climate changes and different management scenarios using the HBV- Light model.

The regional model was validated independently at gauged catchments that were not used in deriving the regional relationship between the model parameters and the catchment descriptors and found to be satisfactory for the flow estimation at ungauged catchments. However, the model works for extrapolation of model parameters to catchments of descriptors within the range considered in deriving the regional relationships, for larger extrapolation, the model may not be robust and the prediction could be uncertain.

The poor performance of the regionalized model could be raised from the anthropogenic activities in the basin which has an important implication on the runoff generation and the regional model of this study has not built by considering anthropogenic influences that led to the occurrence of different runoff generation dynamics within the catchment to regionalize the model parameters. The anthropogenic influences are the natural and human activities such as damming, construction of water retention structures, urbanization, and catchment change and increased abstraction of water for irrigation and industries, and impoundment of water.

The regional model is area specific. i.e., applicable only for the study area they are established for and within the time boundary but the procedure can be applied to other regions.

5.2 Recommendations

Having all the findings and evaluations in the previous section, the following recommendations are broadened for future studies.

CFSR climate data provides complete sets of continuous, areal, climatic datasets for hydrological modeling. These datasets are the best choice for their simplicity to be obtained and used for different functions for hydrological models in providing all climate variables that are essential for the calculation of potential evapotranspiration using Penman-Montieth, as some of these data are being absent in ground observations.

CFSR data, one of satellite rainfall estimate can be subjected to a variety of error sources and the errors in data assimilation, poor association to remotely sensed signals and atmospheric properties that influence the radiation flux. Therefore, it's good to check for and adjust the bias in the CFSR data in reference to ground observed rainfall data of its corresponding catchment before applying it to the hydrological model.

Future study should improve the accuracy the regional model using more data sets that were not included in this study such as data on anthropogenic activities, geological and mining activities, etc. Further improvement of estimation of the parameters may be carried out by incorporating other catchment descriptors. This may recommend a way for future parameter regionalization works.

In this study, PCA is carried out linearly to extract the principal components from physical catchment characteristics. Further study is necessary to consider another classification method that can nonlinearly improve the PCA model structures. The transfer function relating the model parameters with the catchment attributes was set to have a linear form in this study. Future study would test other nonlinear forms of transfer functions with a comprehensive uncertainty analysis included.

The correlation between model parameters especially non-sensitive model parameters (Percolation) needs a geological set of data for the rainfall-runoff modeling from a large number of gauged catchments. Care should be taken while interpreting the hydrological meaning of parameter descriptor relationships found by regressions as high correlation coefficient doesn't assure the hydrological relationship between model parameters and catchment descriptors.

The gauged catchments are found in the northern and northwestern parts of the basin and there are no representative catchments in the southern and southwestern part of the basin yet there is considerable variation in PCCs from northwards to southwards, these challenged for application of model transfer and the regional model development needs further studies using different hydrological models.

REFERENCES

- Abdella, K. (2016). Estimation of discharge for ungauged catchments using rainfall-runoff model in Didessa sub-basin: the case of Blue Nile basin, Ethiopia. *International Journal of Innovations in Engineering Research and Technology (IJIERT)*, 3(9), 62–72.
- Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Petterson, A., & Rummukainen, M. (2001). Climate change impacts on runoff in Sweden - Assessments by global climate models, dynamical downscaling and hydrological modelling. *Climate Research*, 16(2), 101–112. doi.org/10.3354/
- Bergström, S., 1995. The HBV models. Ch. 13, In: Singh, V.P., (Ed.), *Computer Models of Watershed Hydrology*, 443–476.
- Berhanu, B., Seleshi, Y., Demisse, S. S., & Melesse, A. M. (2016). Bias correction and characterization of climate forecast system re-analysis daily precipitation in Ethiopia using fuzzy overlay. *Meteorological Applications*, 23(2), 230–243. <https://doi.org/10.1002/met.1549>.
- Berhanu, S., Alamirew, T., Merkel, B., & Melesse, A. M. (2015). Remote Sens. *Remote Sensing, September*, 11639–11663. doi.org/10.3390/rs70911639
- Beven, K. (2019). *How to make advances in hydrological modelling. 1969*, 1481–1494. <https://doi.org/10.2166/nh.2019.134>
- Bitew, M. M., & Gebremichael, M. (2011). Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands. *Hydrology and Earth System Sciences Assessment*, 1147–1155. doi.org/10.5194 /hess-15-1147-2011
- Booij, M. J., Deckers, D. L. E. H., Rientjes, T. H. M., & Krol, M. S. (2007). Regionalization for uncertainty reduction in flows in ungauged basins. *IAHS-AISH Publication*, 313, 329–337.
- Cunderlik, J. M., & Simonovic, S. P. (2015). Hydrologic Models for Inverse Climate Change Impact Modeling Hydrologic Models for Inverse Climate

Change Impact Modeling. *International Journal of Remote Sensing* ISSN:, June.

Dile, Y. T., & Srinivasan, R. (2014). Evaluation of CFSR Climate data for hydrologic prediction in data- scarce watersheds : an application in the Blue Nile River Basin. *Journal of the American Water Resources Association*, 50(5), 1226–1241. doi.org/10.1111/jawr.12182

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., & Ropelewski, C. F. (2008). Validation of high-resolution satellite rainfall products over complex terrain. *International Journal of Remote Sensing*, Vol. 29(788775834), 4097–4110 Validation. <https://doi.org/10.1080/01431160701772526>

Djaman, K., O’Neill, M., Owen, C. K., Smeal, D., Koudahe, K., West, M., Allen, S., Lombard, K., & Irmak, S. (2018). Crop evapotranspiration, irrigation water requirement and water productivity of maize from meteorological data under semiarid climate. *Water (Switzerland)*, 10(4). doi.org/10.3390/w10040405

Fuka, D. R., Walter, M. T., Macalister, C., Degaetano, A. T., Steenhuis, T. S., & Easton, Z. M. (2013). *Using the Climate Forecast System Reanalysis as weather input data for watershed models*. <https://doi.org/10.1002/hyp.10073>

Fuka, D. R., Walter, M. T., Macalister, C., Degaetano, A. T., Steenhuis, T. S., & Easton, Z. M. (2014). *Using the Climate Forecast System Reanalysis as weather input data for watershed models*. 5623(October 2013), 5613–5623. <https://doi.org/10.1002/hyp.10073>

Gebbru, B. K. (2009). *Grid-based Parameterization of Monthly Water Balance Model for Simulation of River Flow at Ungauged Sites in Abay River Basin*.

Giambelluca, T. W., Nullet, M. A., & Schroeder, T. A. (1986). *Rainfall Atlas of Hawai’i, Report R76, Hawai’i Division of Water and Land Development, Department of Land and Natural Resources*. 267.

Habib, E., Haile, A. T., Tian, Y., & Joyce, R. J. (2012). Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates. *Journal of Hydrometeorology*, 13(6),

1784–1798. <https://doi.org/10.1175/JHM-D-12-017.1>

HU sheng et al. (2017). *Evaluation of the applicability of climate forecast system reanalysis weather data for hydrologic simulation : A case study in the Bahe River Basin of the Qinling*. 27(2014), 546–564. doi.org/10.1007/s11442-017-1392-6

Hundechea, Y., & Ouarda, T. B. M. J. (2008). *Regional estimation of parameters of a rainfall-runoff model at ungauged watersheds using the ' ' spatial ' ' structures of the parameters within a canonical physiographic-climatic spa ... Regional estimation of parameters of a rainfall-runoff model at unga. March 2015*. doi.org/10.1029/2006WR005439

Jillo, A. Y., Demissie, S. S., Viglione, A., Asfaw, D. H., & Sivapalan, M. (2017). Characterization of regional variability of seasonal water balance within Omo-Ghibe River Basin , Ethiopia. *Hydrological Sciences Journal*, 0(0). doi.org/10.1080/02626667.2017.1313419

Legates, D.R., and McCabe, G.J. (1999). Evaluating the use of "goodness- of- fit" measures in hydrologic and hydro-climatic model validation. *Water Resources research* 35, 233-241.

Marc, J. P. (2012). software package Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Zurich Open Repository and Archive University of Zurich Main Library Strickhofstrasse 39 CH-8057 Zurich Wwww.Zora.Uzh.Ch*, 16, 3315–3325. <https://doi.org/10.5194/hess-16-3315-2012>

Mesfin Benti Tolera, I.-M., & Chung, and S. W. C. (2018). Evaluation of the Climate Forecast System Reanalysis Weather Data for Watershed Modeling in Upper Awash Basin, Ethiopia. *Water, Wwww. MDPI*, 10, 725. <https://doi.org/10.3390/w10060725>

Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models, Part I –A discussion of principles, *J. Hydrol.*, 10, 282–290.

Parajka, J., Merz, R., & Bl, G. (2005). *Sciences A comparison of regionalisation*

methods for catchment model parameters. 157–171.

- Penman. (2020). Evaluation the performance of several gridded precipitation products over the highland region of yemen for water resources management. *Remote Sensing*, 12(18). <https://doi.org/10.3390/RS12182984>
- Pombo, S., & de Oliveira, R. P. (2015). Evaluation of extreme precipitation estimates from TRMM in Angola. *Journal of Hydrology*, 523(February), 663–679. <https://doi.org/10.1016/j.jhydrol.2015.02.014>
- Pombo, S., Oliveira, R. P. De, & Mendes, A. (2015). Validation of remote-sensing precipitation products for Angola. *Royal Meteorological Society*, 409(August 2014), 395–409. <https://doi.org/10.1002/met.1467>
- Priestley, C. H. B., & Taylor, R. J. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. *Monthly Weather Review*, 100(2), 81–92. [doi.org/10.1175/1520-0493\(1972\)100\(2\), 81–92](https://doi.org/10.1175/1520-0493(1972)100(2)<81-92>2-B0)
- Ringard, J., Becker, M., Seyler, F., & Linguet, L. (2015). *Temporal and Spatial Assessment of Four Satellite Rainfall Estimates over French Guiana and North Brazil*. 16441–16459. <https://doi.org/10.3390/rs71215831>
- Richard Woodroffe and Associates, 1996. Omo-Gibe River basin Integrated Development Master Plan Study. The Master Plan (Part 1).
- Romilly, T. G., & Gebremichael, M. (2011). *Evaluation of satellite rainfall estimates over Ethiopian river basins. 2008*, 1505–1514. [doi.org /10.5194/hess-15-1505-2011](https://doi.org/10.5194/hess-15-1505-2011)
- Saha, S., Moorthi, S., Pan, H. L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y. T., Chuang, H. Y., Juang, H. M. H., Sela, J., ... Goldberg, M. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8),1015–1057. [doi.org /10.1175/ 2010 BAMS3001.1](https://doi.org/10.1175/2010BAMS3001.1)
- Saito, O., Fürst, C., & Yeshitela, K. (2018). Future land use management effects on

ecosystem services under different scenarios in the Wabe River catchment of Gurage Mountain chain landscape, *Science of the Total Environment* Quantifying and mapping of water-related ecosystem services for enhancing. *Science of the Total Environment*, 646(June). doi.org/10.1016/2018.07.347

Samani, Z. (2000). Estimación de la Evapotranspiración y la Radiación Solar usando un mínimo de datos climáticos. *Journal of Irrigation and Drainage Engineering*, 126(4), 265–267.

Seibert, J. (1999). Regionalisation of parameters for a conceptual rainfall-runoff model. *Agricultural and Forest Meteorology*, ELSEVIER, 99(PII: S0168-1923(99)00105-7), 279–293.

Seibert, J. (2005). HBV light. *HBV Light Version 2 User's Manual*. Number.

Sharma, S. K., Gajbhiye, S., & Tignath, S. (2014). *Application of principal component analysis in grouping geomorphic parameters of a watershed for hydrological modeling* Application of principal component analysis in grouping geomorphic parameters of a watershed for hydrological modeling. February. <https://doi.org/10.1007/s13201-014-0170-1>

Shrestha, M. (2017). Bias correction of climate models for hydrological modelling – are simple methods still useful? *Royal Meteorological Society (RMetS)*, 539(June), 531–539. <https://doi.org/10.1002/met.1655>

Shrestha, M. S., Artan, G. A., Bajracharya, S. R., & Sharma, R. R. (2008). *Using satellite-based rainfall estimates for stream flow modelling: Bagmati Basin* <https://doi.org/10.1111/j.1753-318X.2008.00011.x>

Sivapalan, M. (2003). *Prediction in ungauged basins: a grand challenge for theoretical hydrology*. 3170(June), 3163–3170. doi.org/10.1002/hyp.5155

Tan, M. L., Chua, V. P., Tan, K. C., Brindha, K., Leong, M., Chua, V. P., Tan, K. C., & Brindha, K. (2018). Evaluation of TMPA 3B43 and NCEP-CFSR precipitation products in drought monitoring over Singapore. *International Journal of Remote Sensing*, 2018, 1161, 2089–2104. doi.org/10.1080/01431161.1425566.

Temesgen Ayalew, A. (2019). Rainfall - Runoff Modeling: A Comparative Analyses: Semi Distributed HBV Light and SWAT Models in Geba Catchment, Upper Tekeze Basin, Ethiopia. *American Journal of Science, Engineering and Technology*, 4(2), 34. doi.org/10.11648/j.ajset.20190402.12

Tesfaye, B. (2011). Predicting Discharge at Ungauged Catchments Using Rainfall-Runoff Model(Case study: Omo-Gibe River Basin,Ethiopia). *Addis Ababa University Addis Ababa Institute of Technology School of Graduate Studies Department of Civil Engineering*, 126, 36–49.

Wijngaard, J. B., Tank, A. M. G. K., & Onnen, G. P. K. (2003). *HOMOGENEITY OF 20TH CENTURY EUROPEAN DAILY TEMPERATURE*. 692, 679–692. <https://doi.org/10.1002/joc.906>

WMO. (2003). *Twenty-first status report on implementation of the World Weather Watch: Forty years of World Weather Watch*. (WMO-No. 95, Issue 957). Secretariat of the World Meteorological Organization - Geneva - Switzerland. http://library.wmo.int/pmb_ged/wmo_957.pdf

Worqlul, A. W., Maathuis, B., Adem, A. A., Demissie, S. S., Langan, S., & Steenhuis, T. S. (2014). Comparison of rainfall estimations by TRMM 3B42, MPEG and CFSR with ground-observed data for the Lake Tana basin in Ethiopia. *Hydrology and Earth System Sciences*, 18(12), 4871–4881. <https://doi.org/10.5194/hess-18-4871-2014>

Worqlul, A W, Collick, A. S., Tilahun, S. A., Langan, S., & Rientjes, T. H. M. (2015). *Comparing TRMM 3B42 , CFSR and ground-based rainfall estimates as input for hydrological models , in data scarce regions : the Upper Blue Nile Basin ,*. 2081–2112. <https://doi.org/10.5194/hessd-12-2081-2015>

Worqlul, Abeyou W, Yen, H., Collick, A. S., Tilahun, S. A., Langan, S., & Steenhuis, T. S. (2017). Catena Evaluation of CFSR , TMPA 3B42 and ground-based rainfall data as input for hydrological models , in data-scarce regions :TheupperBlueNile.*Catena*,152,242–251. doi.org/10.1016/j.catena.

Worqlul, Abeyou Wale, Maathuis, B. H. P., Adem, A. A., Demissie, S., & Angeles,

L. (2014). *Comparison of rainfall estimations by TRMM 3B42 , MPEG and CFSR with ground-observed data for the Lake Tana basin in. December.*
<https://doi.org/10.5194/hess-18-4871-2014>

Wuttichaikitcharoen, P., & Babel, M. S. (2014). Principal component and multiple regression analyses for the estimation of suspended sediment yield in ungauged basins of Northern Thailand. *Water (Switzerland)*, 6(8), 2412–2435.
<https://doi.org/10.3390/w6082412>

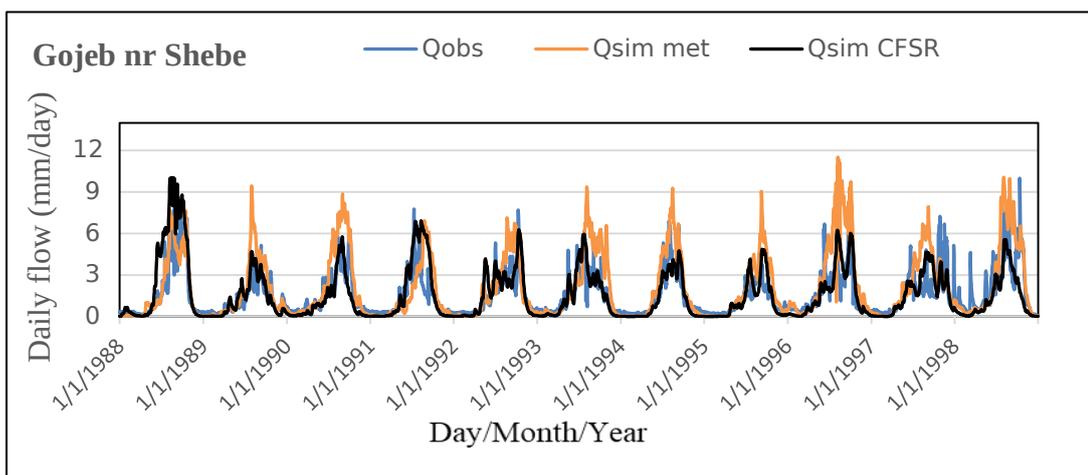
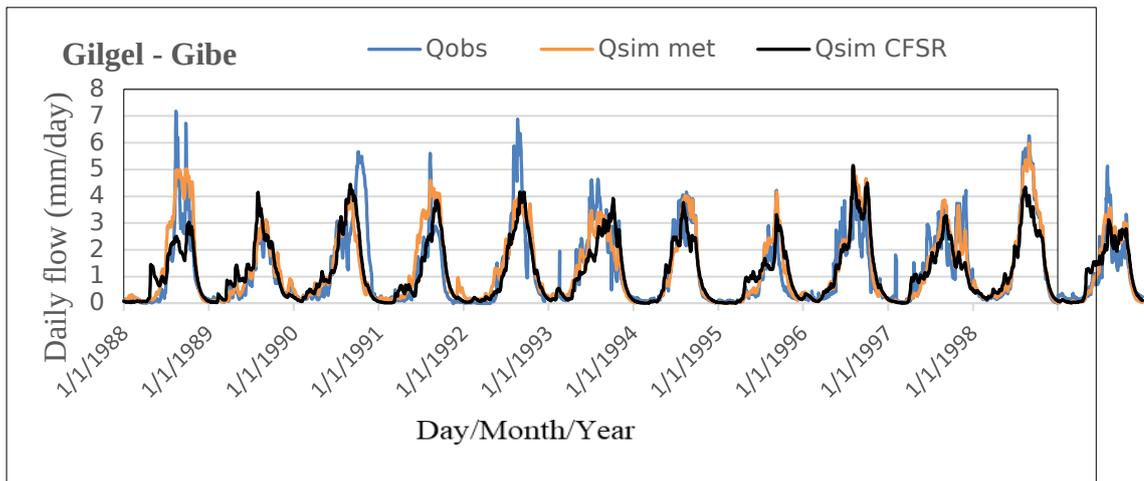
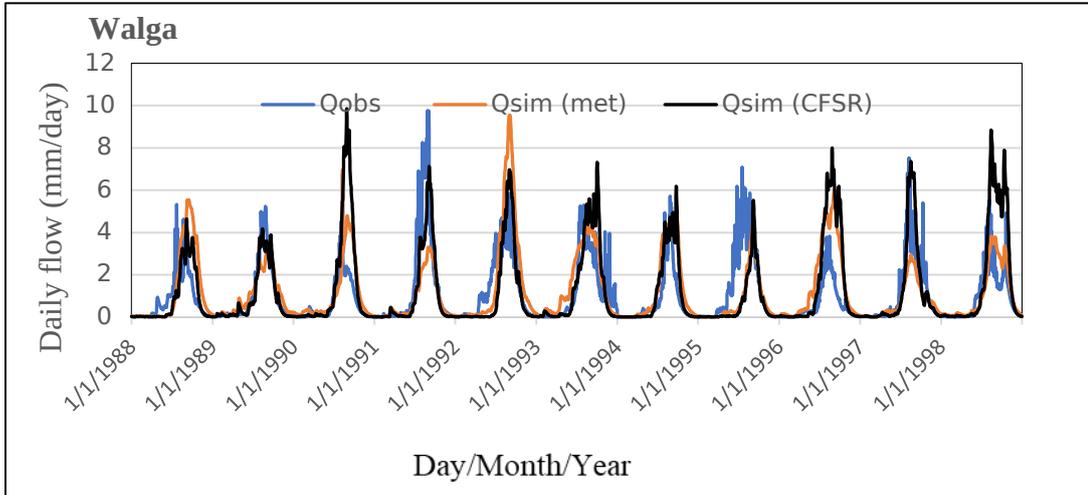
Zhang, L., Meng, X., Wang, H., Yang, M., & Cai, S. (2020). *Investigate the Applicability of CMADS and CFSR Reanalysis in Northeast China.*

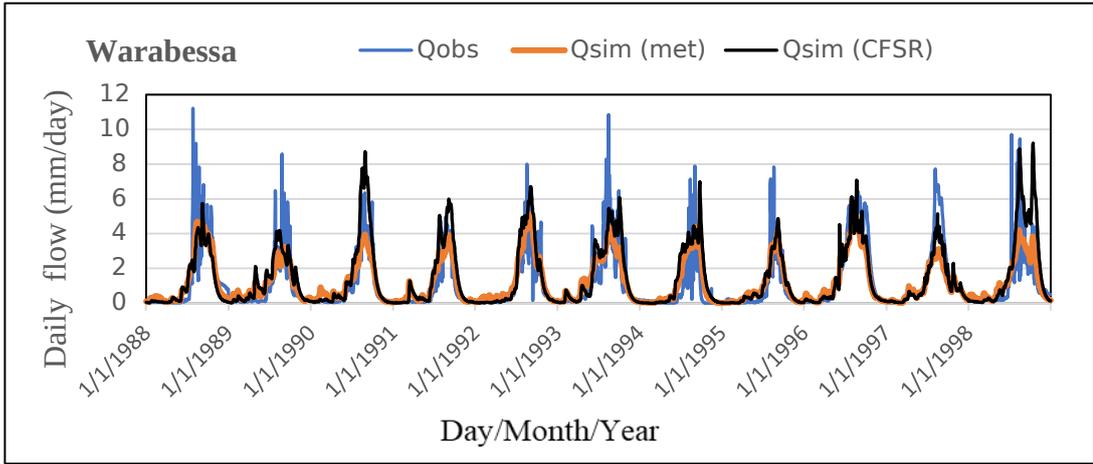
|

APPENDIX A

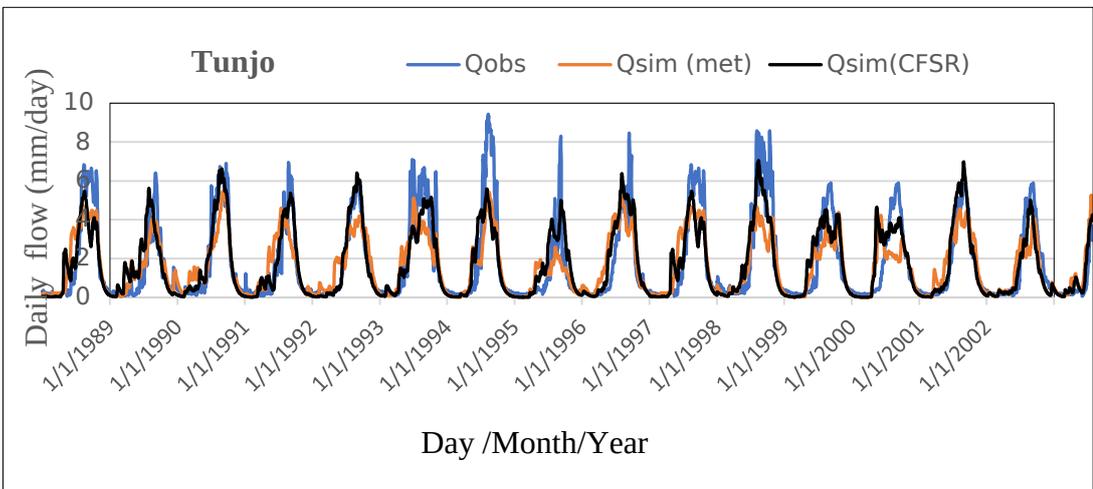
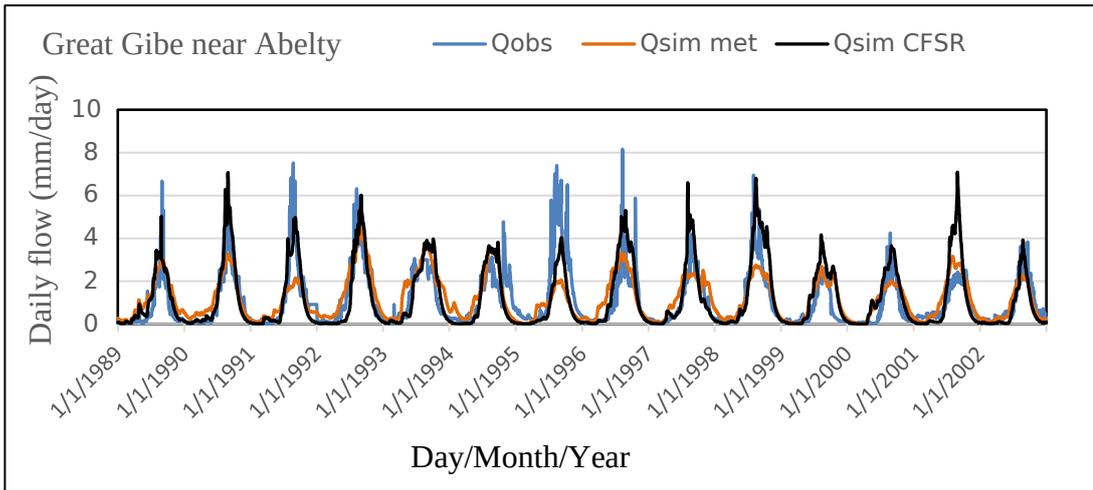
Observed and simulated flow hydrograph of HBV-Light model using CFSR and ground observed climate data as an input at the gauged catchments.

A: Daily calibration (1988 -1998)

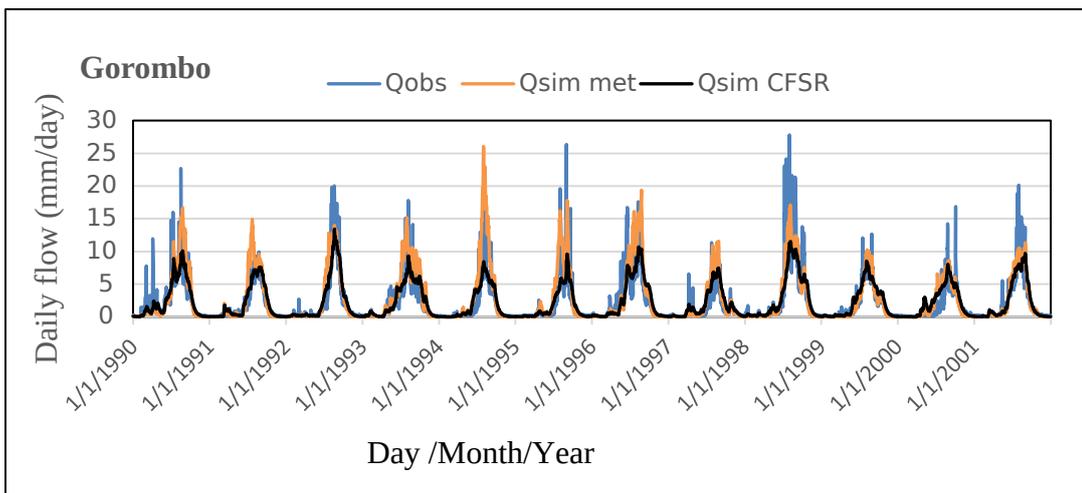
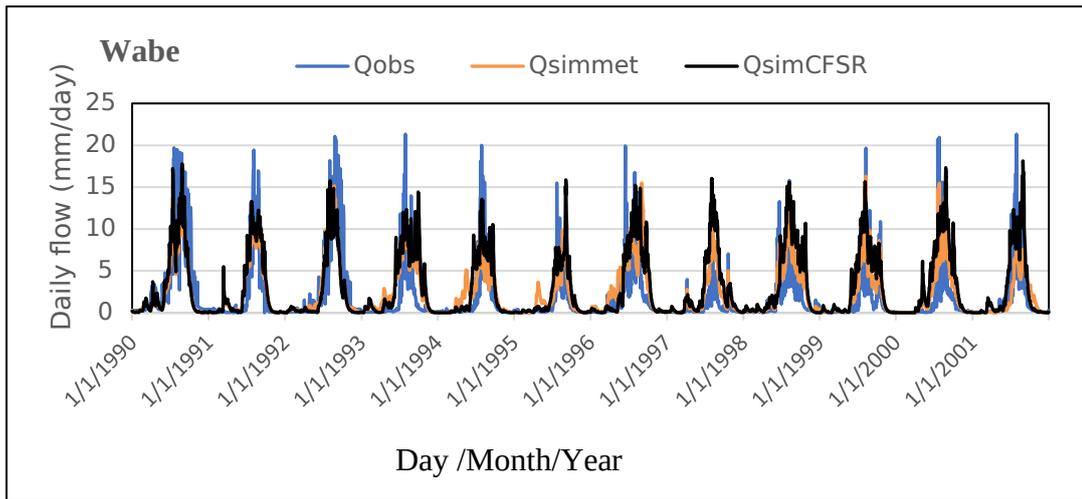




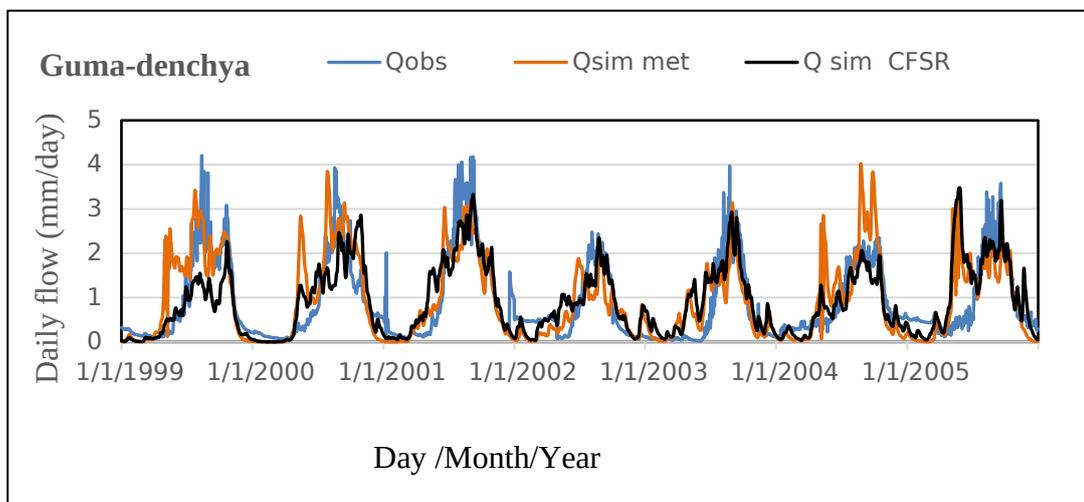
B: Daily calibration (1989-2002)



C: Daily calibration (1990 -2001)



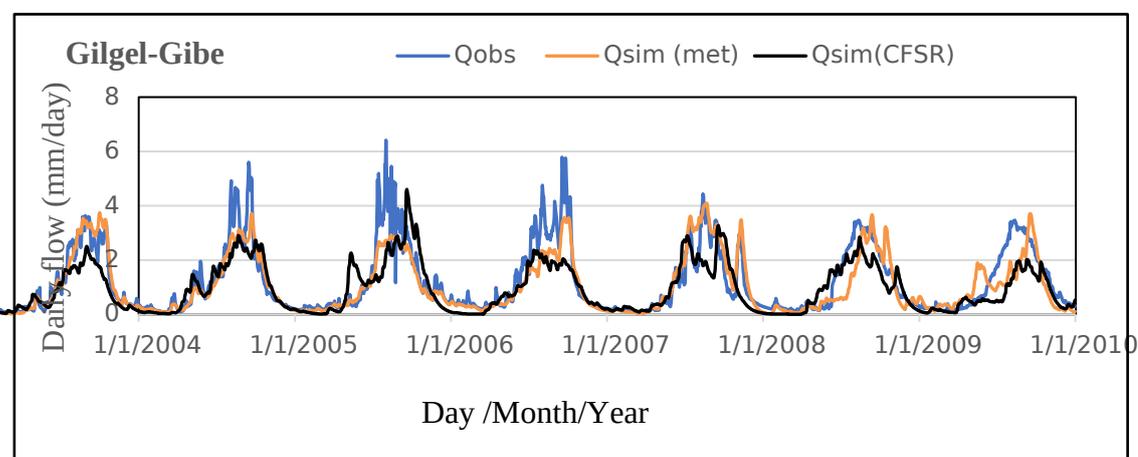
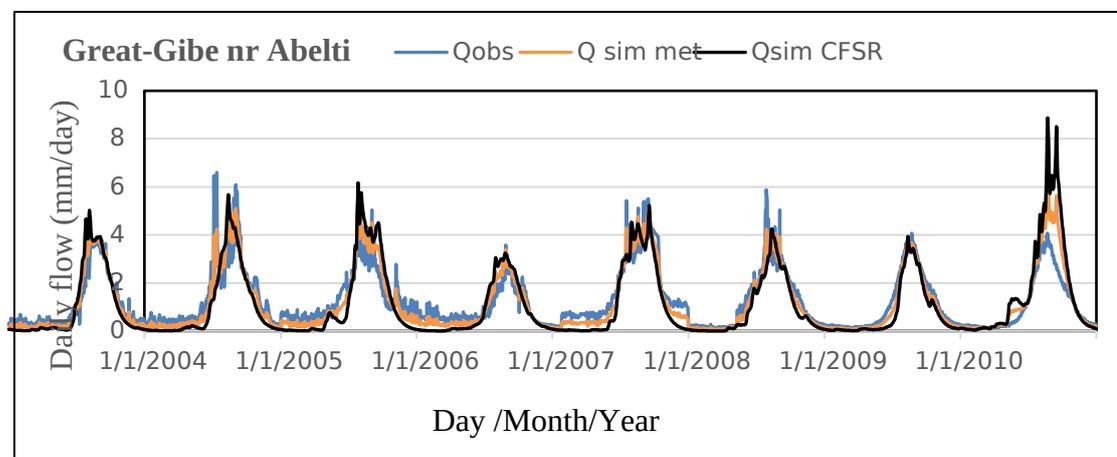
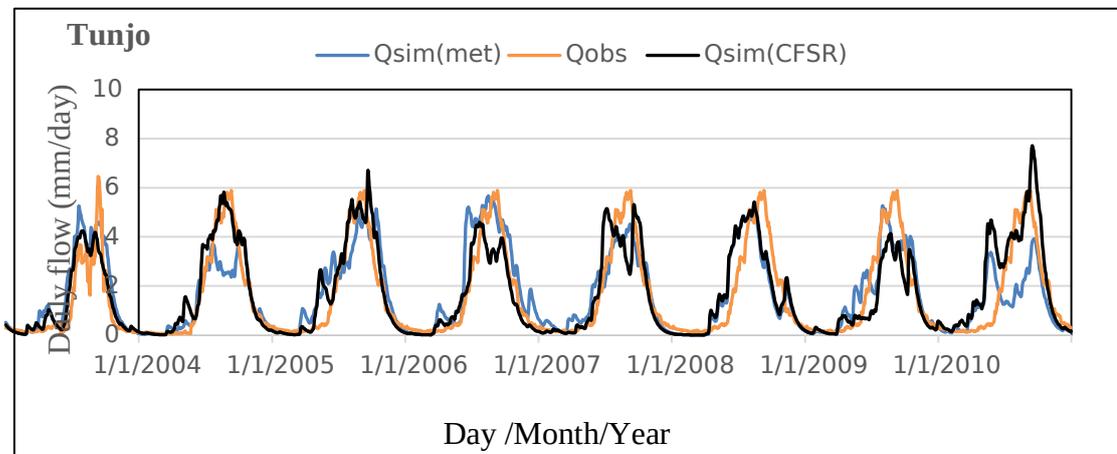
D: Daily calibration (1999 -2005)



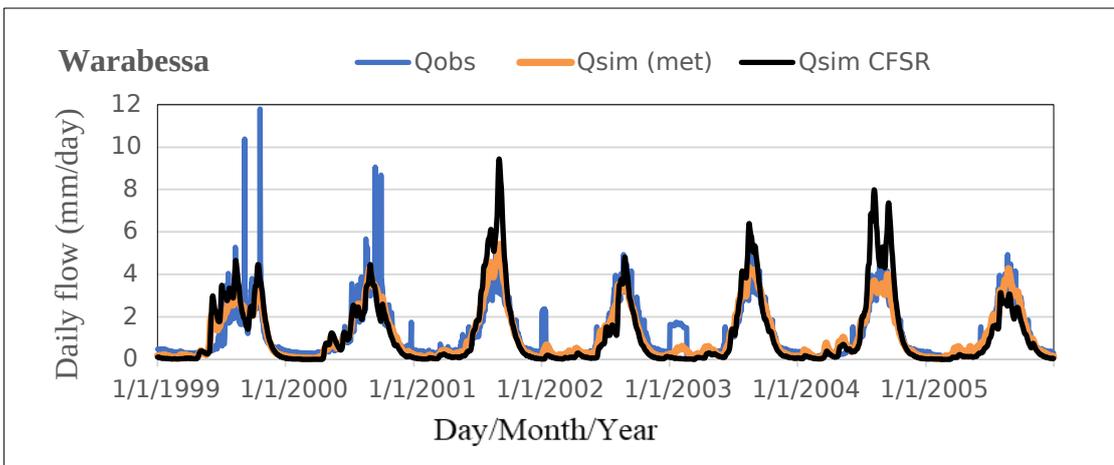
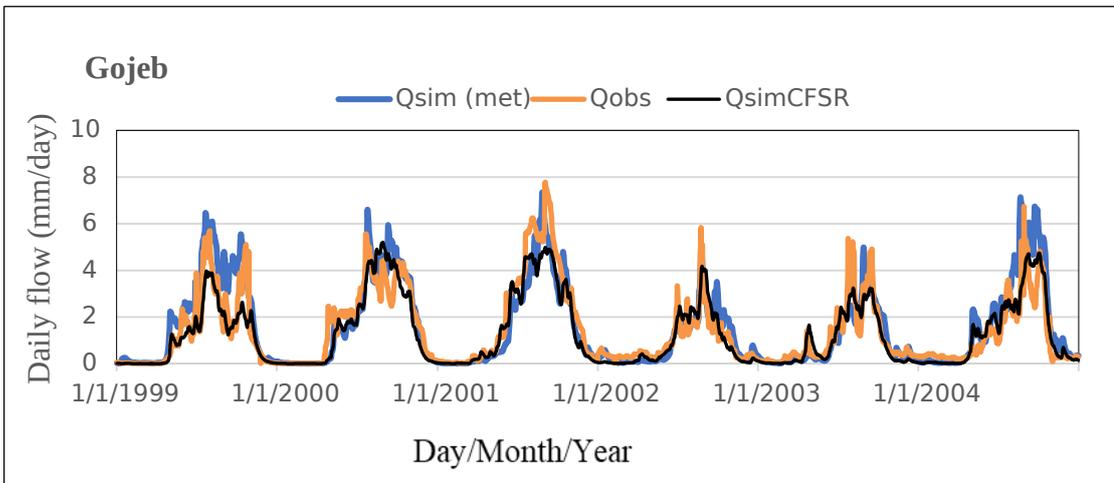
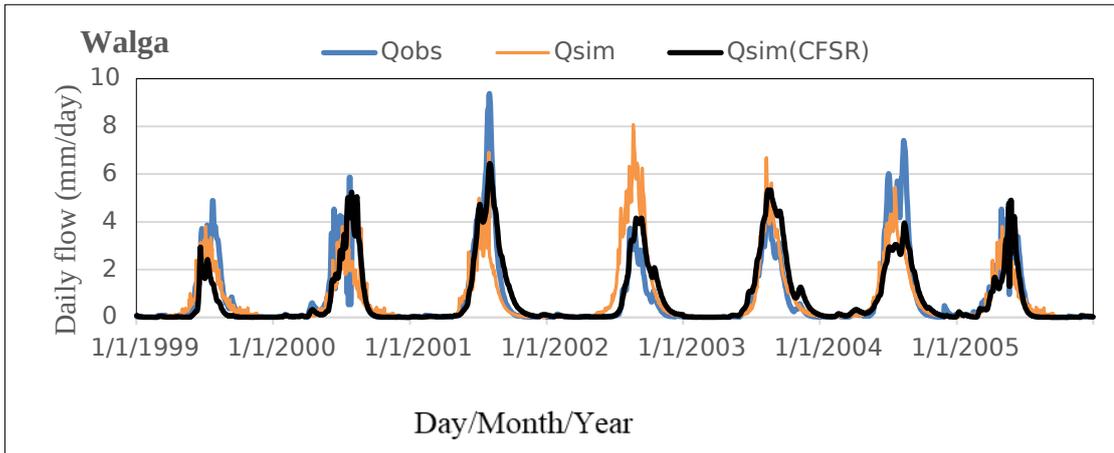
APPENDIX B

Observed and simulated flow hydrographs of HBV-Light model using CFSR and ground observed climate data as an input at the gauged catchments, daily validation period (2004-2010).

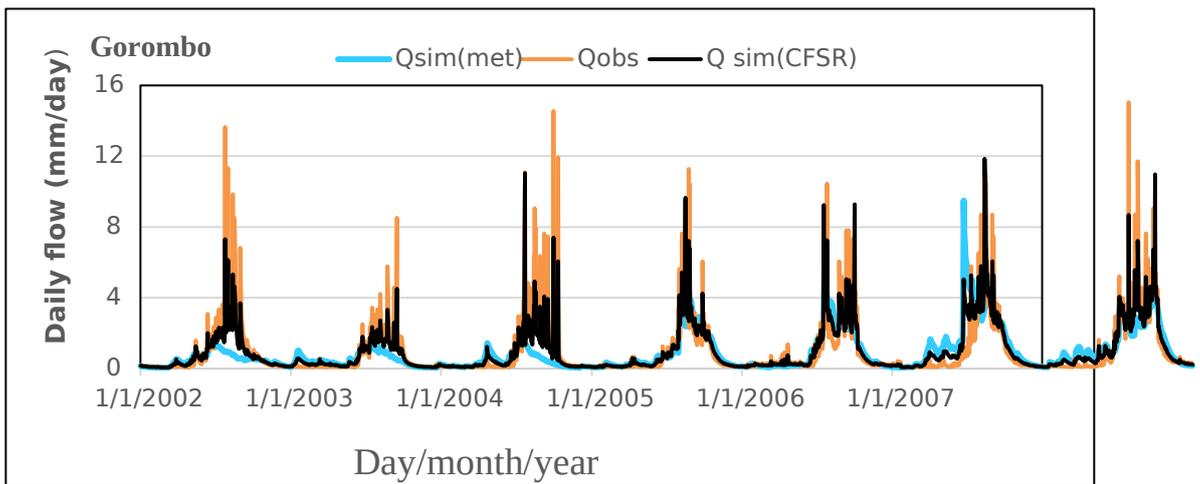
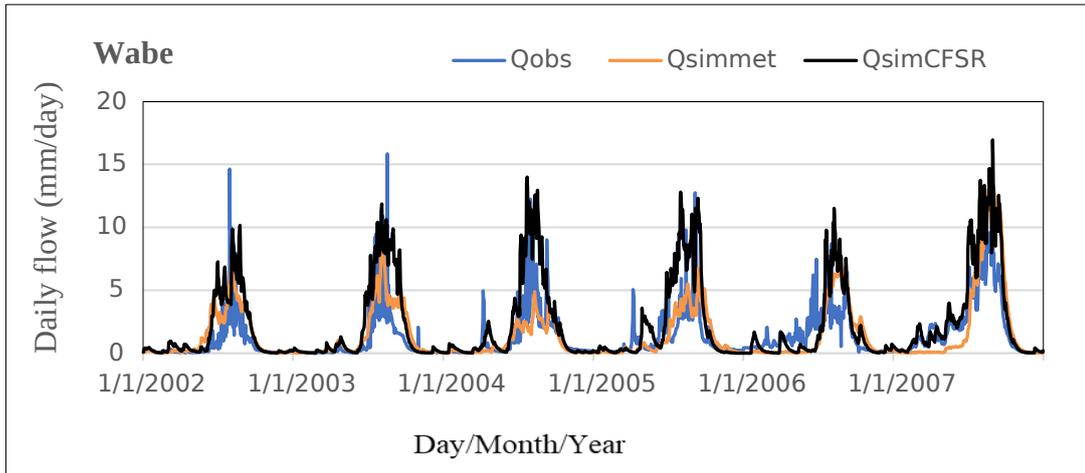
A: Daily validation (2004 -2010)



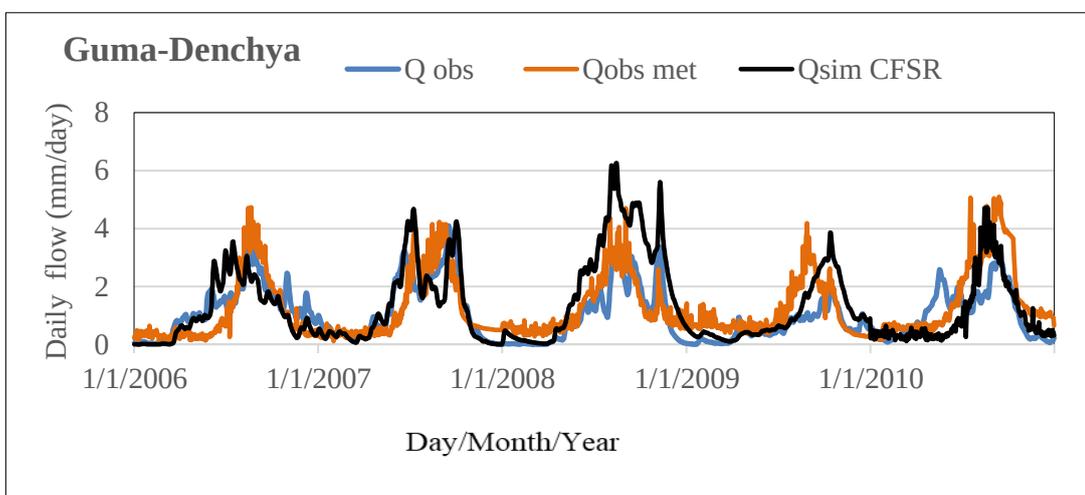
B: Daily validation (1999 -2005)



C: Daily validation (2002 -20007)

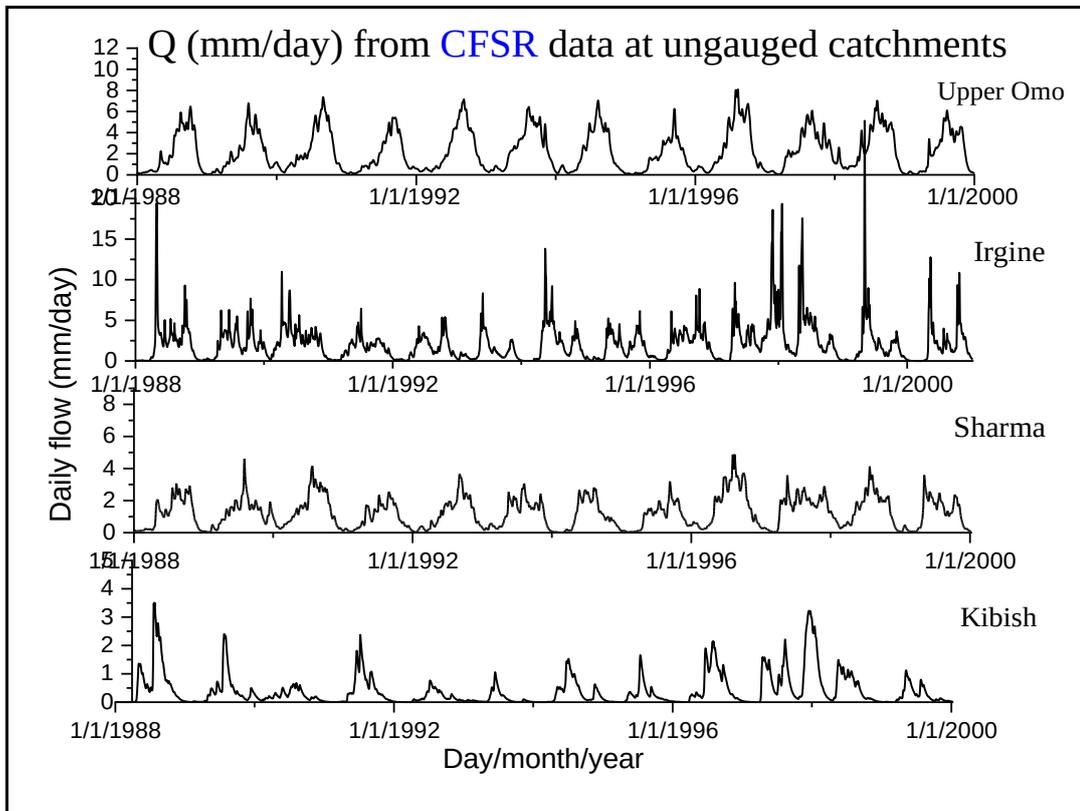
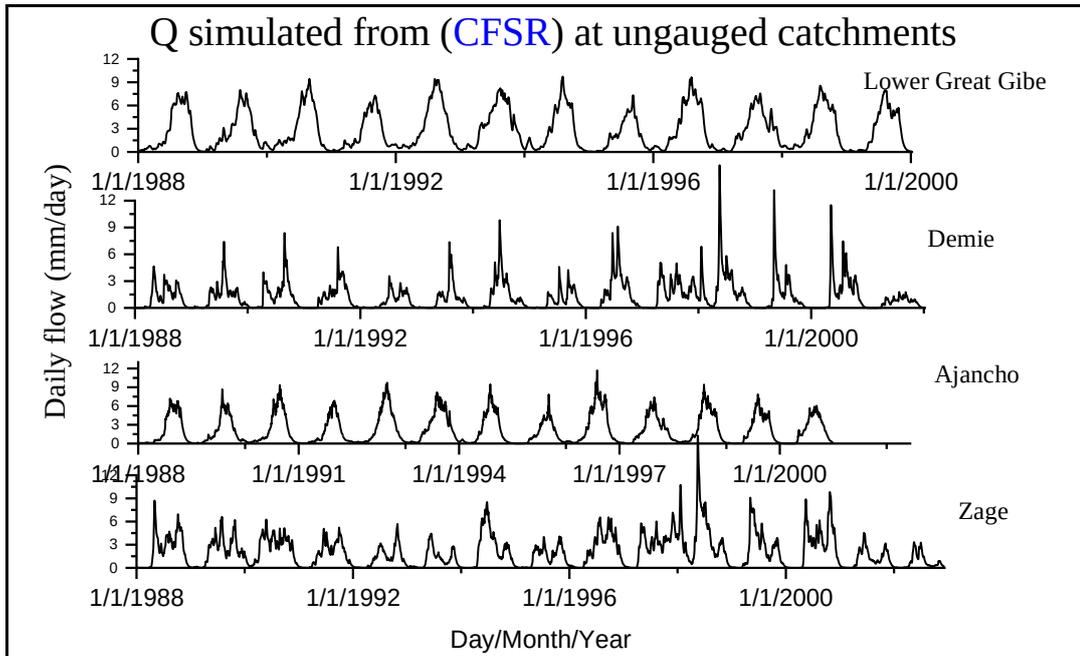


D: Daily validation (2006 -2010)



APPENDIX C

Figure A: Flow generated at ungauged catchments applying the regionalized model using CFSR data as an input data for the period of (1988 to 2000).



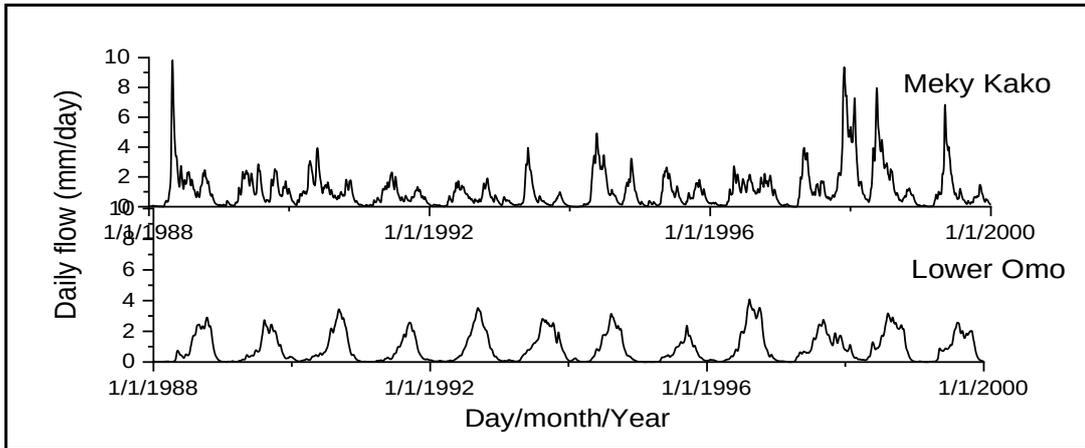
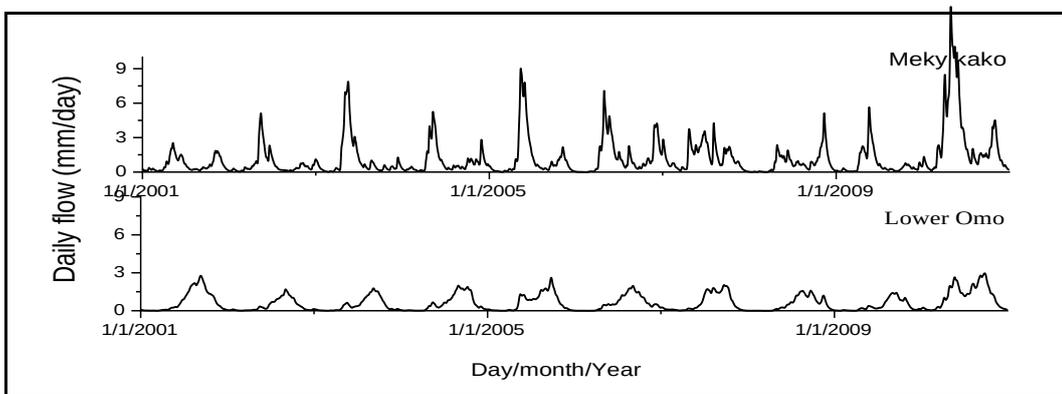
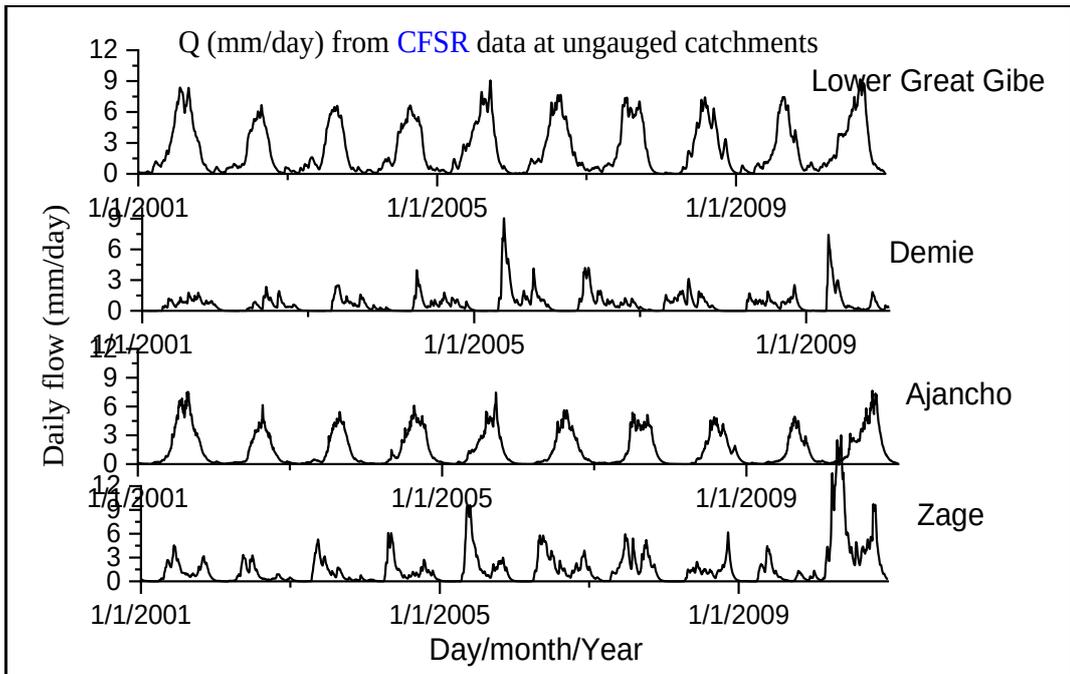
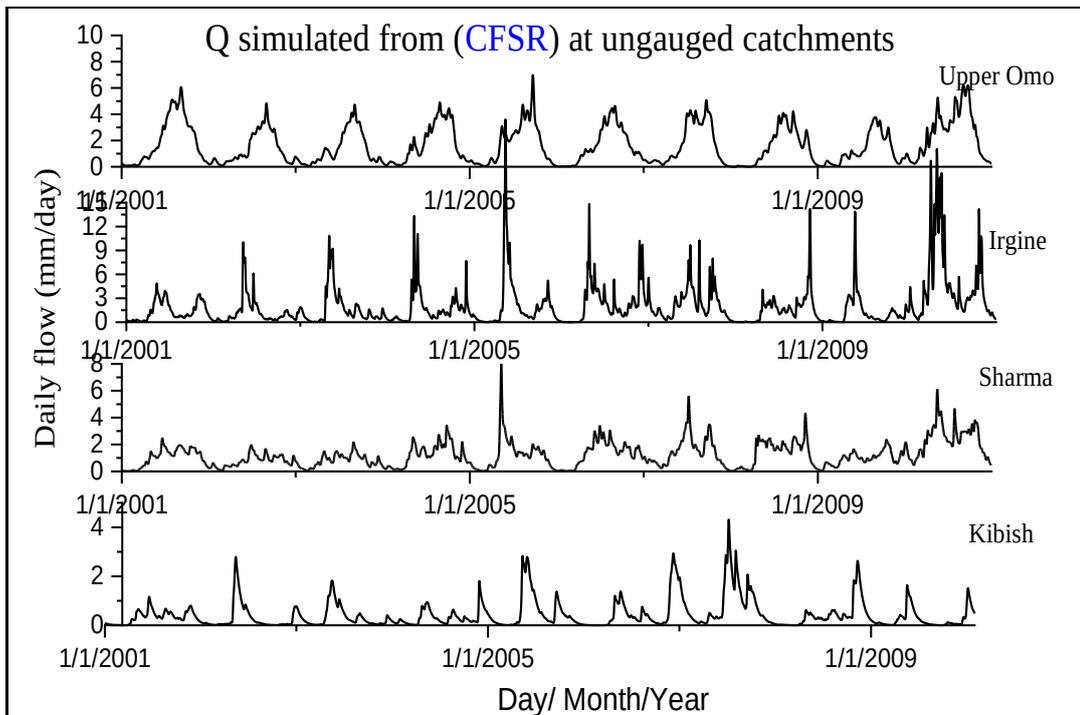


Figure B: Flow generated at ungauged catchments applying the regionalized model using CFSR data as input data for the period of (2001 to 2010).





APPENDIX D

Table 1. Physical Catchment Characteristic indices that were used for nine gauged and ten ungauged catchments, as described in Table 3.3.

Catchments	DA	ME	LLP	TWI	LS	Aspect
Great Gibe	5272	1918	215	7.5	13.76	174
Tunjo	2421	2043	162	8.1	14.81	188
Walga	2829	2039	89	4.8	7.81	199
Wabi	1862	2266	96	7.3	12.17	211
Gorombo	1116	2370	70	7.8	7.8	212
Gumadenchiya	3488	1604	141	7.5	24.79	190
Gilgel Gibe	5136	2048	231	7	12.13	186
Gojeb	6699	1868	210	7.8	20.99	174
Warabessa	1657	1767	30	8.2	13.61	172
Sharma	5212	1211	180	8.01	18.7	178
Middle Omo	6392	690	103	8.7	8.52	180
Meki kako	5574	797	274	8.5	10.2	208
Kibish	4736	1569	10	7.3	13.4	181
Lower Omo	9350	600	262	8.2	8.9	180
Zage	2399	1479	76.5	7.85	21.52	197
Irgene	1266	1768	81	7.15	34.91	192
Deme	1807	1731	83	7.96	18.41	190
Lower- Gibe	4228	1959	114	8.01	16	195
Sokie-Woybo	2968	1647	36	8.03	18.41	188
Upper Omo	5580	1373	177	7.37	27.5	191

Table 2: Physical Catchment Characteristic indices that were used for nine gauged and ten ungauged catchments, as described in Table 3.3.

Catchment	SAT_K	NDVI	% FOREST	%Grass land	% CULTIVATION
Upper-Great Gibe	8.37	0.53	11.7880627	70495	60.50699
Tunjo	8.02	0.59	28.95172	4.91E-05	71.04824
Walga	9.32	0.53	7.1575428	944768	83.89769
Wabi	9.15	0.39	0	8.53734	91.46266
Gorombo	10.13	0.5	0	34.01683	65.98317
Gumadenchiya	9.41	0.69	55.0217340	23944	4.73883
Gilgel Gibe	9.26	0.54	19.8147358	68132	21.50396
Gojeb	8.27	0.55	55.6359137	87704	6.487046
Warabessa	10.33	0.56	15.32258	25	59.67742
Sharma	9.48	0.63	77.59612	20.1226	2.281279
Middle Omo	4.96	0.47	59.228340	53643	0
Meki kako	9.3	0.52	76.2592622	13391	0
Kibish	4.12	0.45	19.4035180	59649	0
Lower omo	8.4	0.4	2.61531644	88549	0
Zage	13.2	0.45	5.6642663	75079	30.58495
Irgene	13.31	0.45	36.0574963	94251	0
Deme	34.8	0.51	22.2545969	31881	30.68119
Lower- gibe	15.42	0.43	8.77803641	11064	50.11132
Ajancho	14.66	0.49	48.9916842	83376	8.174558
Upper Omo	14.5	0.67	17.43546	78.5986	3.965942

Table 3: Physical Catchment Characteristic indices that were used for nine gauged and ten ungauged catchments, as described in Table 3.3.

Catchments	FD	FA	MAR	MAP	AWCSBD
Great Gibe	16	699	1615	1701	0.15 1.300
Tunjo	16	719	1788	1773	0.15 1.300
Walga	8	545	1252	1659	0.14 1.290
Wabi	16	736	1190	1531	0.14 1.290
Gorombo	16	580	1152	1544	0.14 1.300
Gumadenchiya	8	801	1699.5	1735	0.14 1.270
Gilgel gibe	8	713	1376	1715	0.14 1.300
Gojeb	8	1209	1675	1703	0.15 1.290
Warabessa	8	3835	1383	1771	0.13 1.260
Sharma	8	1065	1530	1822	0.14 1.290
Middle Omo	8	1271.84	928.6	1874	0.13 1.440
Meki Kako	16	11925	800	1795	0.11 1.300
Kibish	8	580.07	948	1777.6	0.12 1.500
Lower Omo	8	17760	893	1873	0.13 1.440
Zage	16	518	1192	1688	0.17 1.100
Irgene	16	440	1522	1571	0.1 1.210
Demie	16	544	1265	1717	0.17 1.080
Lower- Gibe	16	5725	1181	1723	0.12 1.200
Ajancho	16	8388	1279	1779	0.12 1.210
Upper Omo	8	11483	1400	1780	0.11 1.200

Table 4: Contribution of the variables (PCCs) to each Principal Component (%)

Variables	PC1	PC2	PC3	PC4	PC5	PC6
Area	0.377	-0.052	0.089	-0.055	-0.112	0.307
elevation	-0.38	0.015	0.175	0.08	0.03	0.128
LLP	0.22	0.024	0.127	-0.432	0.145	0.474
TWI	0.258	-0.076	-0.208	-0.323	-0.097	-0.355
LS	-0.055	0.479	-0.221	0.146	0.212	0.069
Aspect	-0.207	-0.251	-0.221	-0.121	0.411	0.095
FD	-0.235	-0.116	-0.308	-0.347	0.161	-0.179
FA	0.27	-0.137	-0.253	-0.162	0.141	0.406
Prec	-0.15	0.43	0.3	-0.124	0.061	0.078
Pet	0.367	0.071	0.028	-0.107	-0.227	-0.124
AWC	-0.162	0.055	0.159	-0.332	-0.594	-0.06
SBD	0.243	-0.285	0.302	0.278	0.048	-0.115
SAT_K	-0.155	0.2	-0.404	-0.223	-0.295	0.085
NDVI	0.043	0.424	0.263	-0.131	0.15	0.11
% FOREST	0.218	0.264	0.051	-0.178	0.327	-0.478
% Grass land	0.073	0.199	-0.389	0.42	-0.255	0.179
%CULTIVATION	-0.322	-0.251	0.228	-0.165	-0.072	0.102

Table 5: Principal Components used in the regression model

Principal Components at each catchment						
Catchments	PC1	PC2	PC3	PC4	PC5	PC6
Upper Great Gibe	-0.6892	0.0348	1.6851	-1.0096	-0.7236	0.6486
Tunjo	-1.3359	0.2922	2.0959	-1.8672	0.2834	-0.6203
Walga	-2.4924	-1.3750	2.2361	0.8868	0.2063	1.0828
Wabe	-3.4958	-2.7865	0.2810	-0.2498	0.8316	0.0386
Gorormbo	-3.1715	-2.1860	-0.0675	0.0785	0.5734	-0.1938
Guma- Dincha	0.2679	2.7348	1.1859	0.0696	0.8190	-0.1717
Gilgel-Gibe	0.1707	0.3000	1.1448	0.5432	-0.5117	1.3124
Gojeb	1.0128	1.8638	1.6419	-0.2523	-0.2994	0.2784
Warabessa	-0.3968	0.1629	1.0130	0.5223	-0.8136	-0.7672
Sharma	1.9833	2.0169	1.5429	-0.8249	0.3122	-0.9714
Middle Omo	0.4162	-1.1816	0.3850	0.5356	-0.6826	-9.0986
Meki Kako	0.2049	0.0066	-1.3025	-2.1081	1.0152	-0.3310
Kibish	1.4795	1.1020	0.0987	3.5365	-0.7262	-0.8007
Lower Omo	5.0300	9.4960	-0.7842	-0.4740	-1.1471	0.2597
Zage	-1.8923	0.2450	-1.8191	-0.3262	-1.4522	-0.2778
Irgene	-1.6900	1.8120	-2.0462	1.7732	0.1186	-0.3175
Demie	-2.3076	1.6030	-2.7694	-1.3340	-2.5509	-0.2774
Lower- gibe	-0.9241	-1.0230	1.1199	-0.2623	0.1221	0.4579
Ajancho	0.1423	0.4223	-1.6470	-0.0071	0.5621	-0.9700
Upper Omo	0.1081	2.2491	-1.2547	0.7697	0.7679	2.3190

Table 6: Regression equations used for filling and extension of missed hydrological data

Missed River stations(Y)		Nearby River Stations (X)		Correlation (r ²)	Equation	Remark
River Name	Id	River Name	Id			
Ghibe Nr. Baco	91020	Amara	91021	0.73	Y=0.751x+0.039	Filled
Walga	91010	Walga nr tole road	91004	0.76	Y=1.553X+1.228	Filled
Werabessa near tole	91022	Warabessa near selkamba	91010	0.6	Y=1.155X+1.074	filled
Wabi	91004	Megech	91005	0.92	Y=3.127X+1.043	Filled
Gumma	92004	Dincha	92005	0.74	Y=0.9X+2.078	Filled
Gibe near Limugenet	91011	Nr.Assendabo	91008	0.74	Y=1.245X+11.71	Filled
Gogeb Nr. Endeber	91007	Wabi	91004	0.86	Y=0.54X+2.458	Filled
Gibe Nr Seka	91017	G.gibe Nr.Assendabo	91008	0.82	Y=0.334X+0.590	Filled
G.gibe Nr.Assendabo	91008	G.gibe @ Abelti	91001	0.86	Y=0.70X+0.55	Filled
G.gibe @ Abelti	91001	G.gibe Nr.Assendabo	91008	0.86	Y=2.93X+1.031	Filled
Ajancho Nr Areka	92010	Sokie	92009	0.79	Y=0.498X0.958	filled & extended

Appendix E: Correlation matrix of Physical Catchment Characteristics

Correlation matrix of physical Catchment Characteristics																	
	Area	Elevation	LLP	TWI	MS	Aspect	FD	FA	MAR	PET	AWC	SBD	SAT_K	NDVI	% FOREST	% grass land	% CULTIVATION
Area	1.00	-0.78	0.69	0.21	-0.15	-0.38	-0.26	0.19	-0.39	0.74	-0.41	0.38	-0.45	-0.15	0.27	0.12	-0.55
elevation	-0.78	1.00	-0.40	-0.18	0.16	0.35	0.34	-0.25	0.63	-0.75	0.53	-0.23	0.25	0.29	-0.29	-0.16	0.61
LLP	0.69	-0.40	1.00	0.09	-0.01	-0.07	0.12	0.03	0.01	0.39	-0.20	0.07	-0.24	0.12	0.28	-0.19	-0.25
TWI	0.21	-0.18	0.09	1.00	-0.32	-0.28	-0.01	0.45	-0.21	0.09	-0.36	0.35	-0.49	-0.27	-0.28	-0.14	0.10
MS	-0.15	0.16	-0.01	-0.32	1.00	0.01	0.07	-0.12	0.47	-0.02	0.33	-0.57	0.32	0.33	0.18	0.46	-0.35
Aspect	-0.38	0.35	-0.07	-0.28	0.01	1.00	-0.02	-0.62	-0.12	0.07	-0.17	1.00	0.10	-0.19	-0.14	-0.17	0.31
FD	-0.26	0.34	0.12	-0.01	0.07	-0.02	1.00	-0.17	0.47	-0.25	0.02	-0.08	0.28	0.08	-0.03	-0.01	0.06
FA	0.19	-0.25	0.03	0.45	-0.12	-0.62	-0.17	1.00	0.05	0.44	-0.09	0.12	-0.27	0.19	0.05	-0.13	-0.02
Prec	-0.39	0.63	0.01	-0.21	0.47	-0.12	0.47	0.05	1.00	-0.18	0.45	-0.20	0.15	0.71	0.24	-0.17	0.15
pet	0.74	-0.75	0.39	0.09	-0.02	-0.67	-0.25	0.44	-0.18	1.00	-0.15	0.26	-0.36	0.19	0.43	0.08	-0.52
AWC	-0.41	0.53	-0.20	-0.36	0.33	0.07	0.02	-0.09	0.45	-0.15	1.00	-0.28	0.06	0.34	-0.06	-0.12	0.38
SBD	0.38	-0.23	0.07	0.35	-0.57	-0.17	-0.08	0.12	-0.20	0.26	-0.28	1.00	-0.72	-0.04	0.18	-0.21	-0.09
SAT_K	-0.45	0.25	-0.24	-0.49	0.32	0.10	0.28	-0.27	0.15	-0.36	0.06	-0.72	1.00	-0.06	-0.07	0.41	-0.09
NDVI	-0.15	0.29	0.12	-0.27	0.33	-0.19	0.08	0.19	0.71	0.19	0.34	-0.04	-0.06	1.00	0.38	-0.09	0.00
% FOREST	0.27	-0.29	0.28	-0.28	0.18	-0.14	-0.03	0.05	0.24	0.43	-0.06	0.18	-0.07	0.38	1.00	-0.12	-0.62
% grass land	0.12	-0.16	-0.19	-0.14	0.46	-0.17	-0.01	-0.13	-0.17	0.08	-0.12	-0.21	0.41	-0.09	-0.12	1.00	-0.60
% CULT	-0.55	0.61	-0.25	0.10	-0.35	0.31	0.06	-0.02	0.15	-0.52	0.38	-0.09	-0.09	0.00	-0.62	-0.60	1.00