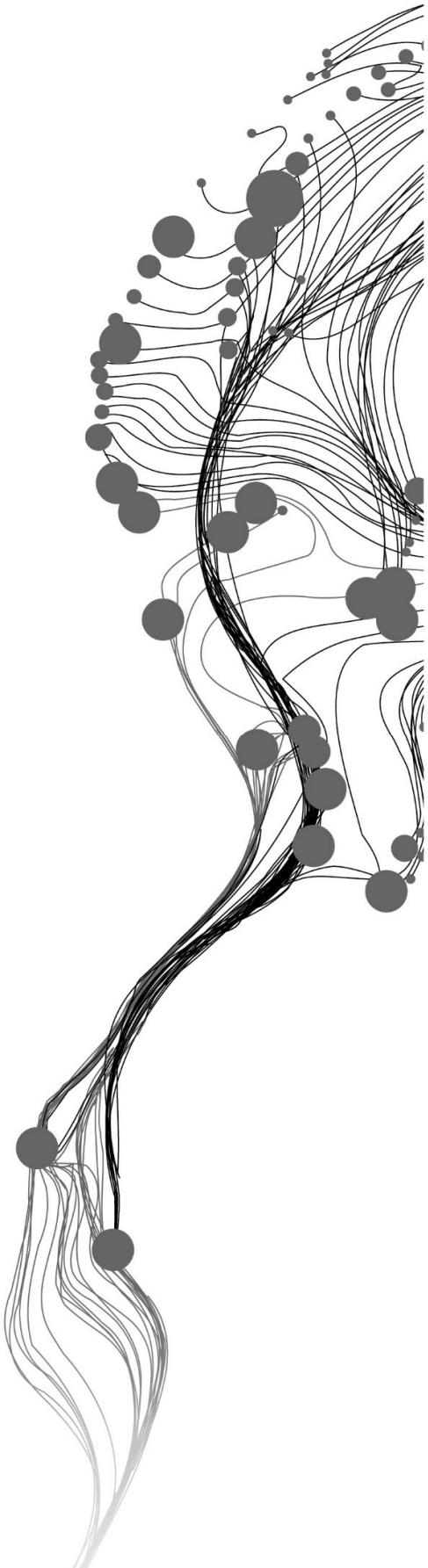
Satellite based rainfall and Potential evaporation for streamflow simulation and water balance assessment: A case study in Wabi watershed, Ethiopia

ASRAT AYELE ALATO
February 2019

SUPERVISORS:
Dr. Ing.T.H.M. Rientjes
Ir. Gabriel Parodi



Satellite based rainfall and potential evaporation for streamflow simulation and water balance assessment: A case study in Wabi watershed, Ethiopia

ASRAT AYELE ALATO

Enschede, The Netherlands, February 2019

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

SUPERVISORS:
Dr. Ing.T.H.M. Rientjes
Ir. Gabriel Parodi

THESIS ASSESSMENT BOARD:
Prof.Dr. Z. Su (Chair)
Prof. Dr. P. Reggiani (External Examiner, University of Siegen, Germany)

ABSTRACT

In this study streamflow simulation and water balance closure was assessed in Wabe watershed located in Ethiopia, Eastern Africa by applying the HBV Light model. Limited ground meteorological measurements restrict water resources planning and management. Such for gauged based rainfall as well as satellite-based rainfall estimates from CMORPH, ARC2, and CHRIPS, and satellite-based potential evaporation estimates from PET-20km was tested.

Satellite-based rainfall estimates was compared with five-gauge stations over the entire time series, wet and dry season (2012-2016). The point to pixel approach was used at daily base and image pixel. The comparison was evaluated by detection indices, scatter plots and frequency-based statistics. The result shows the source of error for a dry was missed rain whereas for wet season was false rain. The result shows that CMORPH outperforms by detecting rainfall depth ~80% in wet season and ~60% in a dry season. Findings reveal that uncorrected CHRIPS matches mean annual rainfall with gauge besides underestimation at the highest elevation. ARC2 underestimates mean annual rainfall followed by CMORPH.

Four bias correction schemes were tested to refine systematic errors in satellite rainfall estimates before being used for the hydrological application. The research findings show that the distribution transformation bias scheme reasonably matches gauge observations with daily accumulated error as low as 5.4mm and coefficient of correlation up to 0.64. However, the prevailing rain rate (<1mm), which accounts for 65%, was reasonably reproduced by space and time variant bias scheme. Furthermore, propagation of errors when comparing and applying bias scheme of SRE's with poor quality gauge measurement is evaluated and verified (Gubire station).

HBV Light model was calibrated following Trial and Error procedure (2012-2016) for gauge rainfall Model efficiency was evaluated by NSE = 0.72, RVE = (-2.56%), Q_{bias}=0.97. Model validation (2009 and 2011) showed NSE = 0.77, RVE = 6.78%. Replacing in-situ ETo with satellite PET resulted in increased peak flows (RVE=2.25). Recalibrating the model with bias-corrected SRE's resulted in minimized bias in streamflow simulation Q_{bias}=0.997 (ARC2) and Q_{bias}=0.994 (CHRIPS) whereas, CMORPH showed deterioration (Q_{bias}=1.017). However, no perfect fit of base flow and peak flow could be simualed by respective SRE's products.

The mean annual water balance closure analysis result shows that water is taken from the system over a five-year period for the respective rainfall and potential evaporation forcing. However, improvement in water balance closure is shown by recalibrating bias corrected SRE's as low as 0.09 (9%).

Keywords: Water balance closure; Streamflow simulation; Distribution transformation; HBV Light; Satellite rainfall estimates; Wabe watershed

ACKNOWLEDGEMENTS

Thank you, almighty GOD. Words cannot explain your support Nigussie.
I would first like to acknowledge my first supervisor Dr.Ing.T.H.M. Rientjes for your valuable and unbounded guidance and advice in every stage of my thesis. Tom, your critical comments and week to week discussions are the backbone of my thesis. Although special thank you goes out to my second supervisor Ir.G.N. Parodi for your excellent guidance to conduct the thesis.

I would like to thanks for the Netherlands Fellowship Programme for providing me a full scholarship. A special thank you goes to my colleagues at WRS department, Faculty ITC for their support and willingness. Mainly, Ir. V. Retsios (Bas) what a kind and helpful person, heartfelt thanks for guidance about ILWIS software throughout the project. C.K. Omondi thank you for providing me part of raw CMORPH satellite-based rainfall data. I am also thankful to the National Meteorological Agency of Ethiopia, and Ministry of Irrigation and Water Resources of Ethiopia, for providing me hydro-meteorological data. Study area visit is not possible without Tadewos Adema and Mr. Teka Moshaga, thank you Tady for your kind support.

I am very thankful to all my families and parents for your sacrifice and financial support.
Thanks to all my WREM classmates. My eighteen months stay in Netherland, ITC was full of happiness with Jing (how to forget you), my classmate, best friend, and everything, thank you. I want to thank Morris my classmate at WRS department, Faculty ITC. You left ITC after nine-month stay but, your support and treatment are throughout my thesis, thank you very much.
Thank you all unmentioned!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ARC2	African Rainfall Climatology Version Two
CMORPH	Climate prediction center MORPHing technique
CHRIPS	Climate Hazards Group Infrared Precipitation with Station
CPC	Climate Prediction Center
CSI	Critical success index
DT	Distribution Transformation
DEM	Digital Elevation Model
ETa	Actual evapotranspiration
EUMETSAT	European Organization for the Exploration of Meteorological Satellites
FAR	False Alarm Ration
FEWSNET	Famine Early Warnings Systems Network
GTS	Global Telecommunication System
HBV	Hydrologiska Byrans Vattenbalansavdelning
IR	InfraRed
ISOD	In-situ and Online Data
ILWIS	Integrated Land and Water Information System
LSA METREF	Land Surface Analysis Reference Evapotranspiration
MoWIR	Ministry of Ethiopian, Water, Irrigation, and Electricity
NMAE	National Meteorological Agency of Ethiopia
ORI_TIRS	Operational Land Imager and Thermal Infrared Sensor
PET	Potential Evaporation
PET-20km	Satellite based- Potential Evaporation with 20km spatial resolution
PMW	Passive Microwave
POD	Probability of Detection
SDRRM	Semi-Distributed Rainfall-runoff Model
SRE	Satellite Rainfall Estimates
SRTM	Shuttle Radar Topographic Mission
USAID	United States Agency for International Development
USGS	United States Geological Survey

1. INTRODUCTION

1.1. Background

Ethiopia has 12 river basins with a total area of approximately 1.104 million km² (99.3% land area and 0.7% covered with water body) (Melesse et al., 2013). The country has an annual runoff volume of 122 billion cubic meters and approximately 2.6-6.5 billion cubic meters of groundwater potential (Awulachew et al., 2007). Omo Gibe river basin with a total area of 79000 km² (Awulachew et al., 2007) is the second largest river basin next to the Nile river basin in Ethiopia. The proposed study area Wabi watershed with an area of 1846 km² is in the northeast part of Omo Gibe river basin.

The hydrologic cycle is the central focus of hydrology (Chow, 1988). Quantification and identifying the interaction of this continuous cycle has been a topic of scientific exploration in the past century, now and in the future. However, the paucity of reference measurement triggered the quantification of catchment runoff response concerning catchment behavior in the globe, regional and or local scale.

Understanding spatiotemporal catchment hydrological behavior is important for water resources planning and management. Hydrological modelling often is practices improving understanding with rainfall and Potential evaporation is main inputs. In the past centuries, in-situ hydro-meteorological measurements provide reliable information to evaluate water balance components and their closure analysis (Rientjes, 2015). However, sparse and inadequate distribution of surface gauge measurements is a challenge to make scientifically sound decisions on water resources and management (Wagner et al., 2009). As argued in a study of Hassan and Jin, (2016) and Dinku et al., (2007) the sparse distribution and limited temporal resolution of ground measurements constrain hydrological modelling in regional and local scale as it requires reliable spatial and temporal time series forcing input data. Concurrently, reliable forcing information in Ethiopia, particularly in Wabi watershed, is hindered by the limitation of surface-based gauge observational networks.

The alternative source for gauge measurement data are satellite rainfall estimates. Nowadays there are several rainfall retrieval satellites in continental and a global scale. Over the past decades and currently, many studies have evaluated the applicability of satellite rainfall estimates. For example, Ashouri et al., (2016), Dinku et al., (2007), Habib et al., (2014), Lyimo, (2015) and Rientjes et al., (2013) was evaluated different SRE's products for streamflow simulation in different regions of the globe, (Hassan and Jin, 2016; Oliveira et al., 2014 and Wagner et al., 2009) was assessed the performance of SRE's products for water balance estimation and Rientjes et al., (2011) evaluated the reliability of SRE's products in regionalization for lake level simulation.

Currently, there are several satellite rainfall products that provide time series of rainfall at spatiotemporal resolution applicable to hydrology. For example the African rainfall climatology version 2 (ARC2; Novella and Thiaw, 2013), the climate prediction center (CPC) morphing technique (CMORPH; Joyce et al., 2004) and the climate hazard group infrared precipitation with station (CHRIPS; Funk et al., 2015), Tropical rainfall measuring mission multi-satellite precipitation (Huffman et al, 2009) etc,. Simultaneously, coarser spatial scale satellite potential evaporation products from USGS Famine Early Warning Systems Network daily global Potential evaporation (FEWSNET; Funk et al., 2015) was used in different regions of the world depending on the time series availability. However, different scholars for example German

and Bolliger, (2006), Artan et al., (2007), Vila et al, 2009, Pan and Wood, 2010 and Nogueira et al, (2018) argued that using satellite precipitation for streamflow simulation, water balance assessment and other hydrological application constrained by systematic errors/bias arise from retrieval algorithm and limitation of sensor.

Hydrological models become essential tools in simulating watershed response and quantification of water balance components. Also, models are crucial for the understanding of hydrological variables and their interaction in a quantitative manner (Seibert and Vis, 2012). There are plenty of hydrological models developed for different purposes for example MIKE SHE, ArcSWAT, HBV, HECRAS, HEC HMS, etc. Seibert and Vis, (2012) argued that Hydrologiska Byrans Vattenbalansavdelning (HBV) rainfall-runoff model was mainly developed for streamflow simulation and quantification of water balance components. Different authors were used HBV rainfall-runoff model for simulating streamflow from satellite meteorological forcing for instance in Ethiopia (Habib et al., 2014; Rientjes et al., 2013; Sendama, 2015; Uhlenbrook et al, 2010) in Chile (Nauditt et al., 2017), in Central Asia (Radchenko et la, 2014), in Tanzania (Lyimo, 2015) and in Rwanda (Sendama, 2015), Mississippi ((Aguirre U. et al., 2013), England, (Rientjes et al, 2010) so and so on.

This study aims to correct and use satellite rainfall and potential evaporation products to simulate and quantify streamflow, and to evaluate the products in modelling. In this study HBV light version (Seibert and Vis, 2012) has been used to address the modelling section. The selection of the model was motivated by its attractive feature of small input data to simulate reasonable result, freely available in lumped and distributed version, its applicability in more than 50 countries and mainly developed for rainfall-runoff simulation.

1.2. Study Relevance

Wabi watershed is intensively agriculture area in Omo Gibe basin. The limited meteorological gauge measurement triggered the determination of runoff response and water balance closure analysis of the watershed. However, nowadays, the development and application of geo-information and earth observation increasingly overcome challenges in limited gauge meteorological measurements. The scientific relevance of this study is to use satellite-based meteorological rainfall and potential evaporation products to data to date, water resources planning and management in Wabe watershed. Also, the study is vital for the catchment community in the sense of producing seasonal variability of rainfall to guide the agricultural production and water resource management in a changing climate. Most studies focus on Blue Nile basin for example (Bhatti et al., 2016; Habib et al., 2014; Haile et al., 2009; Haile et al., 2013), Awash basin (Likasa, 2013) but none has focused in Omo Gibe basin at the catchment scale.

1.3. Problem Statement

Understanding and quantification of the catchment response to rainfall are essential for water resources planning, management, and evaluation. To achieve this availability of reliable and sufficient meteorological rainfall and potential evaporation data is most critical. The Wabe watershed constituted one of a poorly gauged areas in the Omo Gibe basin in the southern part of Ethiopia. Lack of adequate spatio-temporal rainfall and potential evaporation information was a challenge for water resources planning and management in Wabi watershed.

Different scholars for example Rientjes et al., (2013) Artan et al., (2007), Habib et al., (2012), Pan et al., (2010), Wagner et al., (2009), Sendama, (2015) stated that satellite-based rainfall and potential evaporation products serve as an alternative source of data for poorly gauged watersheds. Besides nonexistence of meteorological forcing, hydrological rainfall-runoff models were not applied in the watershed to simulate catchment streamflow and to assess water balance closure. But rainfall-runoff models with satellite rainfall and potential evaporation input, used in different catchments (Abebe et al., 2010; Deckers et al., 2010; Habib et al., 2014; Radchenko et al., 2014; Rientjes et al., 2011; Sendama, 2015) shown reasonable result in simulating catchment streamflow and water balance components. Therefore, motivated by the existing gap and the use of satellite rainfall estimates, this study uses the alternative source of satellite-based rainfall and potential evaporation estimates to address the formulated problem.

1.4. Objective, Research Questions, and Hypothesis

1.4.1. General Objective

The primary objective of this study is to evaluate performance of bias corrected satellite rainfall and potential evaporation products at daily time step to simulate streamflow and assess water balance components applying HBV Light.

1.4.2. Specific Objective

i. To evaluate the performance of CMORPH, ARC2 and CHRIPS satellite rainfall products following the point to pixel approach
ii. To assess the effect of seasonality on CMORPH, ARC2 and CHRIPS rainfall estimates
iii. Apply and evaluate selected bias correction schemes for satellite rainfall products
iv. To assess differences in calibrated HBV light model parameters when gauged and bias-corrected satellite rainfall products serve for model forcing
v. To evaluate the change in streamflow response when in-situ based potential evaporation is replaced by satellite-based potential evaporation
vi. To assess how water balance closure is affected by selected satellite and in-situ based model forcing terms.

1.4.3. Research questions

i. What is the performance of CMORPH, ARC2 and CHRIPS SRE's in capturing point rainfall in Wabi watershed?
ii. Do CMORPH, ARC2 and CHRIPS SRE's capture rainfall distribution concerning gauge?
iii. Which bias correction scheme performs well in Wabe watershed?
iv. To what magnitude do bias corrected CMORPH, ARC2 and CHRIPS SRE's affect the performance of the rainfall-runoff model in simulating streamflow in Wabi watershed?
v. How is water balance closure affected when using SRE's and satellite-based potential evaporation instead of in-situ measurements in Wabi catchment?

This study hypothesizes that bias-corrected African Rainfall Climatology version two results in improved streamflow simulation and water balance closure in Wabi watershed.

2. STUDY AREA AND DATASETS

2.1. Study Area

2.1.1. Geographic location and Topography

Figure 2-1 shows the location map and elevation of Wabe watershed, Omo Giber River basin, and the discharge gauge location at the basin outlet. The study area Wabi watershed with area 1846 km^2 is one of the largest tributaries of Omo-Gibe basin located in the south-western region of Ethiopia. Omo Gibe river basin is the second largest river basin with area 79000 km^2 in Ethiopia next to the Blue Nile. The watershed is located between 8^{0}5'00" to 8^{0}40' 00" latitude and 37^{0}45'00" to 38^{0}40'00" longitude. The Wabe river flows to the Omo-Gibe river that subsequently discharges into Lake Turkana at the Ethiopia-Kenya border. The river originates from Gurage zone mountains. Based on SRTM 30m the elevation of Wabe watershed ranges from 1672 to 3600m above mean sea level with the lowest elevation in the discharge outlet and highest elevation upstream Gurage zone mountains. The catchment is selected for this study, due to most intensively used agricultural in the basin and its complex topography In Figure 2-1 the light green filled is Ethiopia boundary inside Africa continent, Omo Gibe river basin inside Ethiopia boundary (bottom right edge) and in the northeastern part of Omo Gibe river basin boundary, Wabe watershed is located (small red boundary).

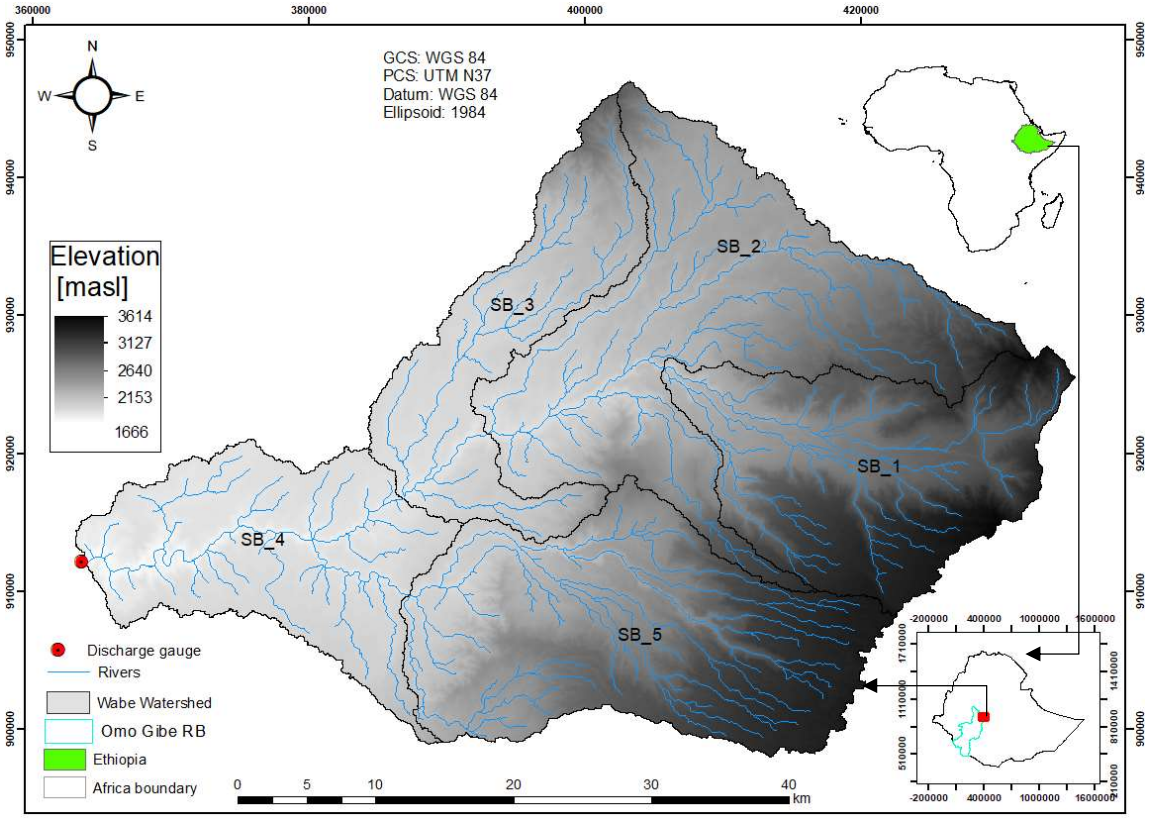


Figure 2-1 Location of Wabi watershed, Ethiopia

2.1.2. Climate, soil, slope and Land cover

Figure 2-2 shows soil type (A) and slope variation (B) in Wabe watershed. Preprocessing of 2007 to 2016 meteorological stations data collected from NMAE inside and around study area shows that the temperature ranges minimum of 8 ^{0}C in the mountainous area during the wet season (August) and a maximum of 37 ^{0}C during the dry season (March) with an average temperature of 18 ^{0}C. The basin receives an average annual rainfall of 1200 to 1300mm from 2007-2016 (NMAE). For the same period, the average yearly potential evaporation is 1631 mm. According to FAO, (1974) soil classification the dominant soil is, plinthic Luvisols and Vertisols whereas Nitosols cover some area in southwestern and northeastern part of the Wabe watershed sea Figure 2-2. According to FAOCLASS1 classification (LPq=Plinthic Luvisols, LVx = Vertic Luvisols, NTu= Nitosols, and VRe= Vertisols). As shown in Figure 2-2 (B) the variation in slope [%] ranges from 0 to 62.1.

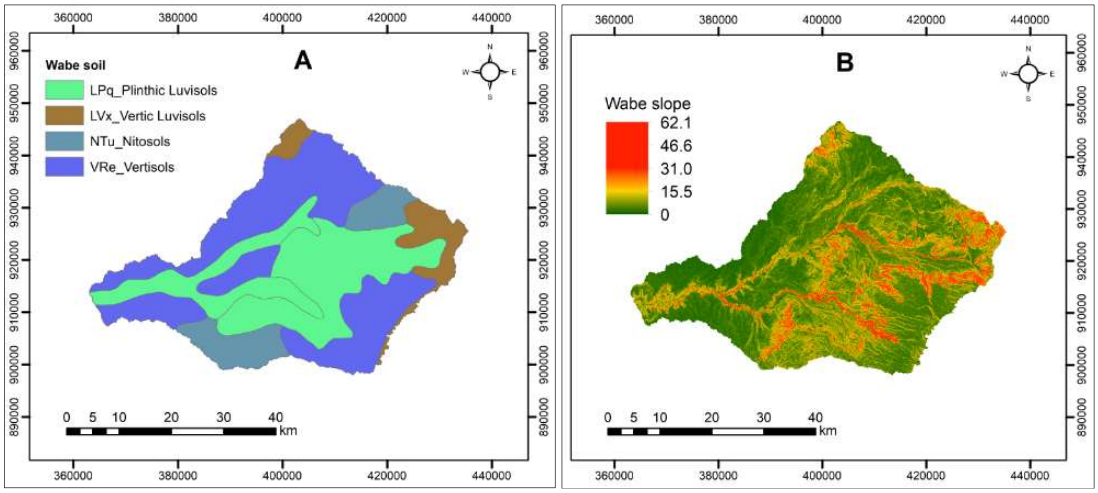


Figure 2-2 Wabe soil type and slope variation

Figure 2-3 Shows the landcover map of Omo Gibe river basin and the study area. The landcover map was collected from MoWIE for Omo Gibe river basin and masked to Wabe watershed. Focusing on the study area, there are five land cover types. As seen in Figure 2-2 (A) the areas covered with the two dominant soil types are covered by cultivation and medium cultivation Figure 2-3, showing the suitability of soil type for agriculture. Nevertheless, the coverage of forest and open water were different from what were visited during fieldwork duration. In addition to that, the northeast concave part which is covered by Wolkite city, but not seen from the map. The other issue is, the open water is shown in southwest, center, north, and northeast (light orange color) which is not currently happening in the study area, besides the occurrence of the wetland during the wet season in the northeast part. Presumably, the mismatch of landcover in the ground and collected from MoWIE office could be, the collected data is too old.

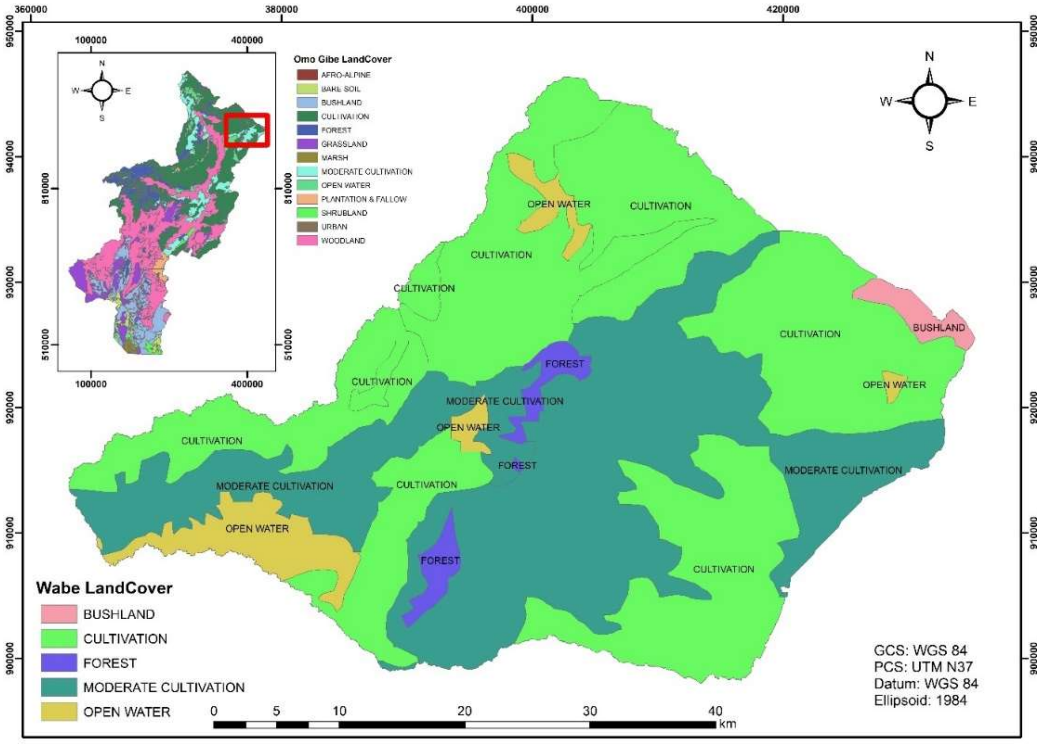


Figure 2-3 Omo Gibe and Wabe watershed land cover map collected from MoWIE

Based on fieldwork visit the landcover of the study area is classified as a cereal crop, enset, chat, built up, grassland, forest, bushland, and eucalyptus (fieldwork visit). However, the catchment is dominated by cereal crop, enset, and forest. Enset is the leading home garden food crop in Wabe watershed. As depicted in

Figure 2-4 it looks like a large, single-stemmed banana plant with an underground corm, a collection of leaf sheaths and large broad leaves. Enset is larger than banana with up to 10m height and 2m width (field work). It has a multipurpose crop with all parts utilized for the different purpose, i.e., human food, construction materials, animal forage, and cultural practices.

Figure 2-4 Enset plant in the garden of local buildings

2.2. In-situ data (collected from offices)

2.2.1. Meteorological measurements

Figure 2-5 shows the location and distribution of meteorological gauge station in and around Wabe watershed collected. As stated in section 1.3 meteorological measurements are poorly distributed for the watershed, and this is evidenced in Figure 2-5. Particularly, there are no gauge measurements inside and northeast high elevated mountainous regions of Wabe. The location and available data for each station are indicated in
Table 2-1 Summary meteorological data collected from NMAE during fieldwork
.

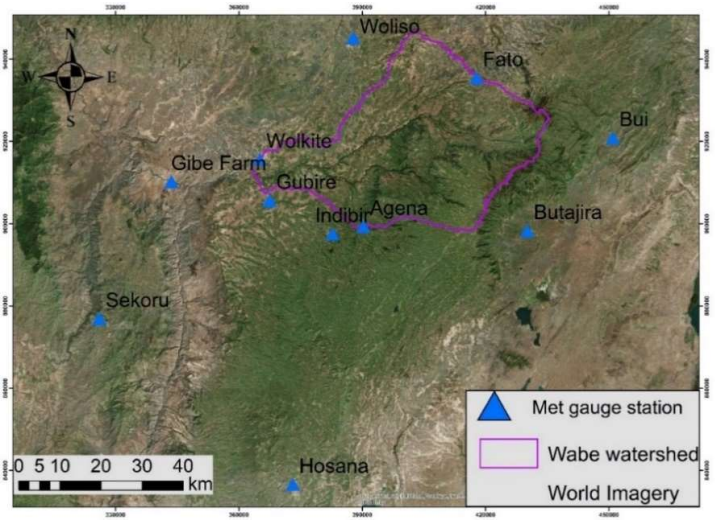


Figure 2-5 Indicates the 11 meteorological stations in and around Wabe watershed.

Based on NMAE classification out of eleven stations five are first class measurements, i.e., Sekoru, Woliso, Butajira, Bui and Hossana. The distribution of gauge location is mostly in the border of watershed even some of them are far away (southwestern part).

Table 2-1 Summary meteorological data collected from NMAE during fieldwork

depicts meteorological data collected during the fieldwork time window. Stations which are written in bold are located inside and nearest to study area.

Table 2-1 Summary meteorological data collected from NMAE during fieldwork

#	Station name	Coordinate of station		Elevation	Type of meteorological data					
		Lat [°S]	Long [°E]	[m.a.s.l]	Rainfall [mm/day]	Tmax [°C]	Tmin [°C]	WS [m/s]	SH [hr]	RH [%]
1	Sekoru	7.93	37.42	1926	A	A	A	A	A	A
2	**Woliso**	8.55	37.98	2158	A	A	A	A	A	A
3	Hossana	7.57	37.85	2307	A	A	A	A	A	A
4	Bui	8.33	38.55	2054	A	A	A	A	A	A
5	**Imdibir**	8.12	37.94	2081	A	A	A	x	x	x
6	**Wolkite**	8.28	37.77	2000	A	A	A	x	x	x
7	**Fato**	8.46	38.25	2520	A	A	A	x	x	x
8	**Agena**	8.13	38.00	2310	A	A	A	x	x	x
9	Butajira	8.13	38.37	2000	A	A	A	x	x	x
10	Gibe farm	8.23	37.58	1092	A	x	x	x	x	x
11	**Gubire**	8.19	37.80	1892	A	x	x	x	x	x

Where: Tmax and Tmin is maximum and minimum temperature respectively, WS is wind speed, SH is sunshine hour, RH is relative humidity, X is not available data, and A is available data

2.2.2. Observed Discharge

Table 2-2 depicts observed discharge data (2007-2016) collected from MoWIE. During the fieldwork duration, the study catchment was visited, and some representatives were selected for an interview. Mr. Teka Moshag (see Figure 2-6 B) in the middle was working as Wabe river water level recorder since 1994 (more than 30 years). As per his work experience indicated the occurrence of runoff mostly depends on the upstream rainfall event. He also pointed out that the river is perennial with having base flow throughout the season. The problem he notices is there is a high volume of sediment driven to downstream particularly during wet season due to the high elevation variation, and upstream agricultural areas border the Wabe river. The water level stage is in a wide and flat cross-section part of the river near the upstream of Wabe old bridge. There is a sediment deposition during the wet season and affects the station stage-discharge curve. This sediment deposition causes the river bed channel to silt up with sediments and thus affects the reliability of the streamflow discharge estimated by water level measurements (Figure 2-6 A). Perhaps this may increase the discharge volume since there was no sediment flushing carried out around water level banks.

Table 2-2 Observed discharge collected from MoWIE during fieldwork

Station name	Elevation	Location		Data availability		Area
Wabe near Wolkite	[m.a.s.l]	Lat	Lon	from	to	[km²]
	1672	8.23	37.98	1/1/2007	31/12/2016	1846

Figure 2-6 shows pictures taken at Wabe outlet location during fieldwork. As shown in Figure 2-6 A and C the color of the water is different. C is taken during a rainy day, and A is taken two day after rainy day.

Figure 2-6 Pictures showing water level location (A), discussion with Mr. Teka Moshag (B) and downstream flow of Wabe river from the top of the new bridge (C)

2.2.3. Satellite rainfall estimation products

Three satellite rainfall products are selected and used in this study. There are common and individual selection criteria. The common criteria are 1) their high spatial and temporal resolution, 2) freely available

time series data for study time domain, and 3) their wide range of application in different regions of the world. Individual selection criteria are discussed in each section.

2.2.3.1. Climate Prediction Center MORPHing rainfall product

As Joyce et al., (2004) climate prediction center MORPHing rainfall is based on the approach where PMW derived precipitation and IR brightness temperature are blended to retrieve global rainfall (Latitude: 60^0N - 60^0S and Longitude: 180^0W-180^0E). For this study Version, two was used. More detail description about input data, algorithm, and methodology for CMORPH is accessible from (Joyce., 2004; Maathuis and Mannaerts, 2013). The selection of product is due to its high spatial and temporal resolution (see Table 2-3). And also, its applicability if evaluated for different regions. For example, Habib et al., (2014, 2012) and Haile et al., (2013) evaluated the performance on streamflow simulation in the Gilgel Abay basin (Ethiopia), Gumindoga et al., (2016) assessed the performance in Zambezi river basin (Zambia). CMORPH was retrieved from freely accessible online source ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/CRT/0.25deg-DLY_00Z/.

2.2.3.2. African Rainfall climatology version 2 (ARC2)

The African rainfall climatology Version 2 (ARC2) of the famine early warning system was launched by the climate prediction center of united states agency for international development (USAID) (Novella and Thiaw, 2013). The two main input sources for ARC2 are three hourly geostationary IR data centered over Africa from EUMETSAT and quality controlled GTS 24-h gauge rainfall accumulations over Africa (Novella and Thiaw, 2013). The selection of this rainfall product is due to, its catchment scale coverage helps to assess the impact of rainfall on water resources management in poorly gauged Wabe watershed, it was not tested in Ethiopia and particularly in Wabi watershed, minimal research effort done for ARC2, but it is with approximately same spatial and temporal resolution with CMORPH (see Table 2-3). Detailed information about ARC2 is found in (Novella and Thiaw, 2013). ARC2 was retrieved from freely accessible from ftp://ftp.cpc.ncep.noaa.gov/fews/fewsdata/africa/arc2/ online climate prediction center's ftp server.

2.2.3.3. CHRIPS rainfall product

As Funk et al., (2014) and Funk et al., (2015) the Climate Hazards group InfraRed Precipitation with Stations(CHRIPS) gets its main input data from CHPClim, quasi-global IR geostationary satellite observations from CPC and NDC, TRMM 3B42 product from NASA and gauge precipitation from different sources. This satellite rainfall product is selected due to its very high spatial and temporal resolution as well as its application in global and local scale. CHRIPS rainfall was retrieved from ftp://ftp.chg.ucsb.edu/pub/org/chg/products/CHIRPS-2.0/africa_daily/tifs/p05/.

Table 2-3 Summary of satellite rainfall products used in this study with data existence, time window, temporal and spatial resolution

Satellite rainfall product	Spatial resolution	Temporal resolution	Spatial and temporal resolution used	Data existence	Time window used	Data provider
CMORPH	0.07°x0.07°	30-min	0.05°x0.05°/Daily	1998-present	2012-2016	NOAA-CPC
ARC2	0.1°x0.1°	Daily	0.05°x0.05°/Daily	1983-present	2012-2016	NOAA-CPC
CHRIPS	0.05°x0.05°	Daily	0.05°x0.05°/Daily	1981-present	2012-2016	CHG, USGS
Reference	(Joyce et al., 2004), (Novella and Thiaw, 2013) and (Funk et al., 2015) for CMORPH, ARC2 and CHRIPS respectively					

2.2.4. Satellite Potential evaporation estimation products

Compared to satellite rainfall products there are only few satellite potential evaporation products. In the last decades, only FEWSNET product is tested in different regions because of the freely available time series of data. Table 2-4 shows a comparison of currently available potential evaporation estimates temporal domain. From the indicated three products, LSA METREF is available from 2016 onwards. Therefore, it is not included in the selection. Different authors validation shows that MODIS16A3 product only suitable for global application(Alvarado and Orozco, 2017) and also applicable to limited temperate and dry regions for example African Savana (Ramoelo et al., 2014), North West China (Zhigang et al., 2007) and Mexico Yucatan Peninsula (Alvarado and Orozco, 2017).

Table 2-4 Available satellite potential evaporation products with time domain.

#	PET Products	Time domain	Reference
1	LSA METREF	2016 to present	(LSA LISA Team, 2016)
2	FEWSNET	2001 to Present	http://earlywarning.usgs.gov/fews/
3	MODIS16A3	2000 to Present	(Steven et al., 2017)

2.2.4.1. FEWSNET potential evaporation product

Therefore, for this study FEWSNET PET is selected due to its time series availability. It calculation is based on the Penman-Monteith equation which was applied in many hydrological studies (Allen et al, 1998).

2.2.4.2. PET-20km potential evaporation product

The PET-20km satellite potential evaporation data currently not freely available on the online database. This data was collected from Dr.ir. C.M.M. Mannaerts (Chris) department of Water resources (ITC-WRS). This data source comes from NASA Global Modelling and Assimilation Office and GEOS-5 Goddard Earth Observation System Model v.5 and DAS data Assimilation System.
PET-20km information available at https://gmao.gsfc.nasa.gov/weather_prediction/.

Table 2-5 Summary of satellite potential evaporation products used in this study with data existence, time window, temporal and spatial resolution

Satellite potential evaporation product	Spatial resolution	Temporal resolution	Spatial and temporal resolution used	Data existence	Time window used	Data provider
FEWSNET	1°x1°	Daily	0.2°x0.2°/Daily	2001-present	2012-2016	NOAA-GDAS
RET/PET-20km	0.2°x0.2°	Daily	0.2°x0.2°/Daily	not free	2012-2016	NASA-GEOS-5
Reference	http://gmao.gsfc.nasa.gov/ forPET-20km and https://earlywarning.usgs.gov/fews/datadownloads/Global/PET for FEWSNET					

2.2.5. Digital Elevation Model (DEM)

The delineation of watershed boundaries and the extraction of topographic information requires digital elevation model data. Khan et al., (2014) stated that DEM is a primary requirement for hydrological modelling. However, Kenward, (2000) and Thomas et al., (2014) show that the horizontal resolution and vertical accuracy of digital elevation model do affect the modelling outcome. According to these studies the decision of DEM to be used for specific study based on two approaches. The first is the study domain area and purpose second, the vertical accuracy which is tested through error statistics such as root mean square error, mean error and standard deviation against ground truth points. Different researchers used 90m SRTM DEM for rainfall-runoff modelling (Lyimo, 2015; Omondi, 2017; Radchenko et al., 2014; Haile et al., 2011; Rientjes et al., 2011) for catchment area ranges between 1655 to 70,000 km^2 whereas

Nauditt et al., (2017) used 30m SRTM for the catchment area of 814 km². The role of DEM in this study is first to delineate and extract watershed boundary, second to partition the entire watershed to sub-basins, third and to redistribute the limited measured meteorological forcing data to the whole subcatchments based on elevation slicing. To achieve this SRTM 30m global elevation data the product of NASA (Farr et al., 2007) is used. It is offered and distributed free of charge by NASA/USGS through earth explorer with geographic coverage of 60⁰N-57⁰S latitude and 180⁰E-180⁰E. The main criteria for selection of 30m DEM is due to its high horizontal resolution, less vertical error ~16m with respect to datum as reported by (Farr et al., 2007), free of charge, availability in different format and applicability for the same study area size and taking consideration the processing time of model to be used. SRTM was retrieved from https://earthexplorer.usgs.gov/.

2.2.6. Landcover satellite imager data

Table 2-6 shows the description of Landsat 8 OLI-TIRS data used for landcover classification. The selection of sensor and date is due to, freely availability cloud-free Landsat image and to evaluate classified image by collected ground truth points respectively. The georeferenced Landsat imagery was downloaded from freely available earth explorer archive https://earthexplorer.usgs.gov/.

Table 2-6 Landsat 8 image data used for land cover class classification

Product	Landsat 8	Sensor_id	OLI_TIRS
File date	2018-11-12	Date acquired	2018-05-12
Format	Geotiff	WRS_PATH and Row	169/54
Spacecraft id	Landsat_8	Scene_center_time	07:45:37.4861350Z
Number of bands	11	Projection/unit	UTM/Meter
Datum	Wgs84	Ellipsoid/ zone	WGS84/37
Corner lat es	9.72935and 9.72139	Corner_lon_we	36.652 and 38.730

Figure 2-7 Shows ground control points collected during fieldwork visit. The global positioning system GPS Garmin E-Trex 30 was used. The ground control land cover points was collected for clear sky day with an accuracy of 3m. In total 258 ground control land cover points were collected. Individually for built up 43 control points, forests 84 control points and cereal crop (including Enset) 131 control points.

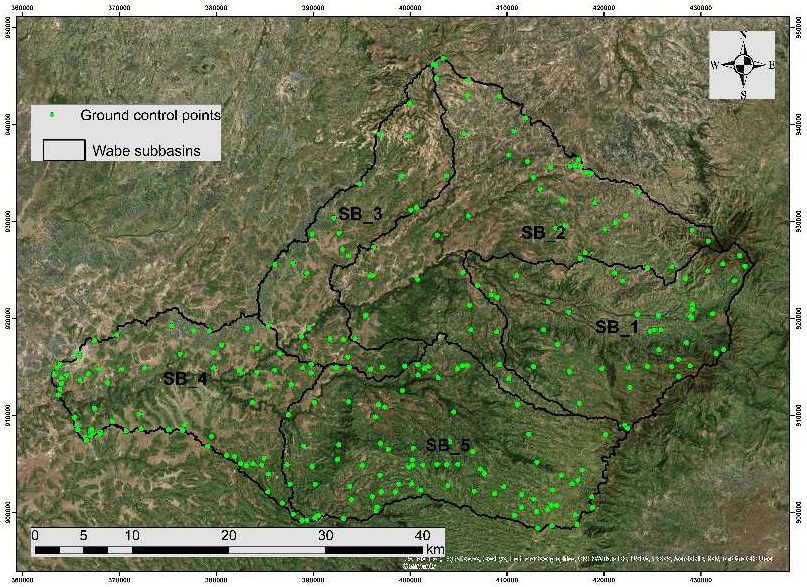


Figure 2-7 A ground truth land cover points

3. RESEARCH METHODOLOGY

3.1. Methodology applied

Figure 3-1 Shows the methodology employed to address the stated objectives systematically. The flowchart begins with two main input data sources of gauge and satellite products. The first is in-situ hydro-meteorological data collected from NMAE and MoWIE whereas; the second is satellite-based rainfall estimate from ARC2, CHRIPS and CMORPH, satellite potential evaporation estimates from FEWSNET and PET-20km, DEM from SRTM DEM and Landsat landcover data from Landsat archive. Following this quality assessment, performance analysis and bias correction (gauge to a pixel in daily scale), watershed delineation and catchment partition, streamflow simulation and water balance assessment shown in the flowchart.

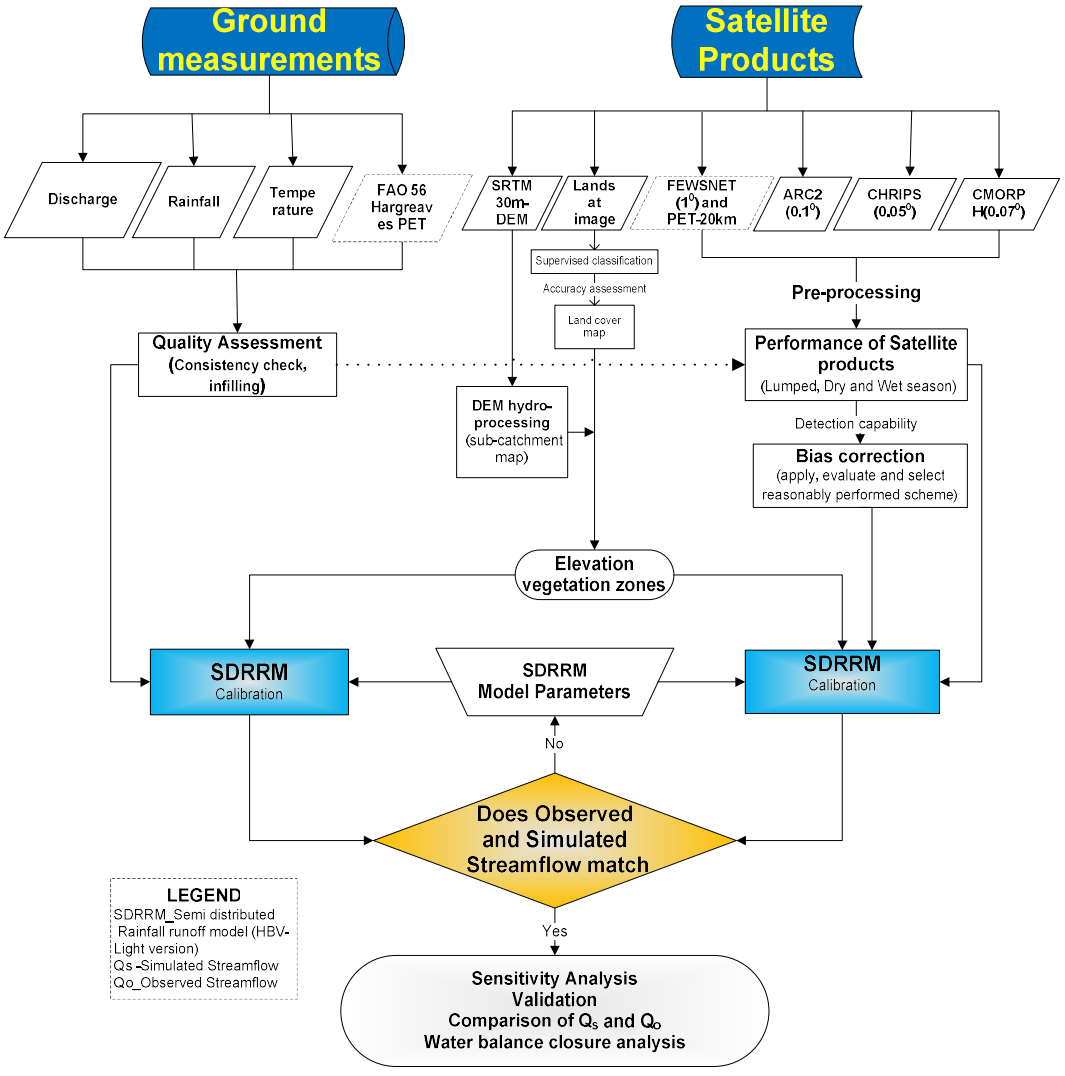


Figure 3-1 Methodology Flowchart

3.2. Hydro-meteorological in-situ measurement pre-processing

In this section the consistency and completeness of the in-situ hydro-meteorological measurements were checked and corrected for further analysis.

3.2.1. Selecting, screening and correcting rainfall measurement

The double mass curve (Searcy and Hardison, 1960) was used to see the consistency between individual stations (see Appendix 1A . Reasonable consistency is seen at Sekoru, Woliso, Imdibir, Fato, and Agena whereas poor consistency is observed at Butajira, and increasing measurement was observed at Gubire station. In addition to DMC, the proximity of gauge station to the basin is considered as it captures catchment condition see section 2.2.1 and Figure 2-5.

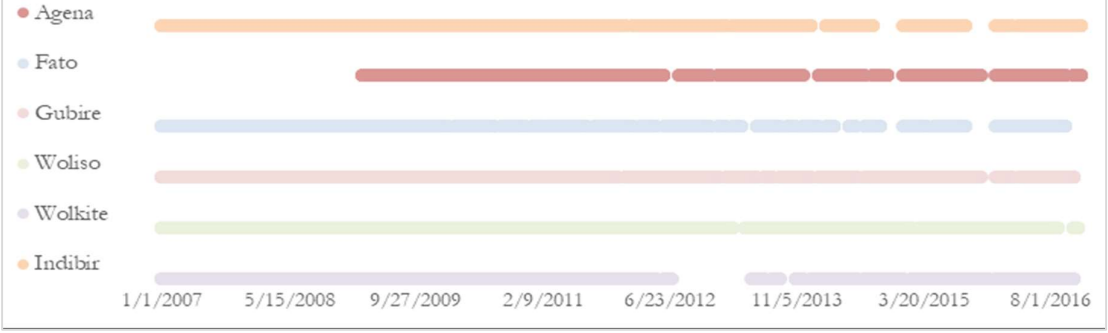


Figure 3-2 Available rainfall data for reference station (2007-2016)

Table 3-1 Analysis of gauge precipitation showing data gap and availability

Station Name	Elevation [m.a.s.l]	Location		Data availability		Missing data	
		Lat	Long	From	To	%	#
Imdibir	2081	8.12	37.94	1/6/2002	31/12/2022	8.01	439
Agena	2310	8.13	38.00	1/3/2009	31/12/2016	21.19	1161
Fato	2520	8.46	38.25	1/2/2002	31/12/2018	15.44	846
Gubire	1892	8.19	37.80	1/3/2002	31/12/2019	6.31	346
Woliso	2158	8.55	37.98	1/4/2002	31/12/2020	3.28	180
Wolkite	2000	8.28	37.77	1/5/2002	31/12/2021	11.81	647

3.2.2. Filling missed rainfall data

The analysis in Table 3-1 shows the met-gauge station measurements was having missed values of highest 21.19% and lowest 11.81% at Agena and Woliso station respectively. On the other hand, these gauge measurements are considered as ground truth to compare and to correct the bias of satellite rainfall products. Therefore, infilling the rainfall measurement should reflect the catchment characteristics. As such multiple linear regression model in equation (3-1) and (3-2) is selected to fill the missed data based on neighboring gauge stations. The method is used after (Michael L, 1996; T. H. M. Rientjes, 2016).

$$P_x = \beta_0 + \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3 + \cdots \beta_n P_n \quad (3\text{-}1)$$

$$P_x = \beta_0 + \sum_{i=0}^{N} \beta_i P_i + e \quad (3\text{-}2)$$

Where: P_x is the dependent variable (rainfall station in question), $P_{1\text{-}n}$ is the independent variables (neighboring rainfall stations), β_0 is the intercept, $\beta_{1\text{-}n}$ is the regression coefficients for N^{th} gauge station, and e is the error term. This plot is an indication to ignore non-correlated station to prevent the issue of redundancy. Appendix 1B shows how the station in question (Gubire) is correlated with neighborhood stations. It has a weak correlation which underestimates the predicted rainfall. This step was done for all other five stations in question. Where the blue color depicts the ground measurement at astation whereas orange is predicted value by the model. In addition to the correlation scatter plots, the goodness of the model is characterized by regression statistics (RS) such that Coefficient

of determination (R^2), multiple regression (R) and adjusted coefficient of determination (R^2_{adj} in all three RS cases value closer to 1 tells the model is reasonably fitted. The model is used to infill the missed in-situ meteorological forcing with RS between 0.7 to 0.87 except for Gubire station.

Figure 3-3 shows a double mass curve (Eris and Agiralioglu, 2012) for selected rainfall gauge stations for further analysis. As shown from DMC analysis the slope keeps constant for five-gauge station except at Gubire. At a Gubire station there is two slopes. This is due to the increased ground measurements from 2012 onwards.

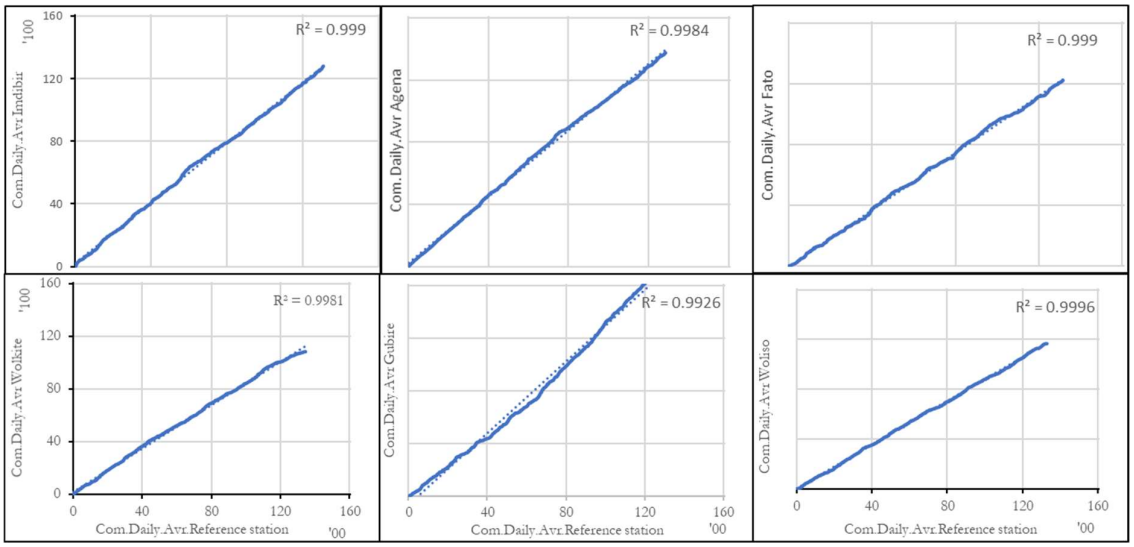


Figure 3-3 Double mass curve showing precipitation consistency check. Test station at abscissa and the cumulative mean of other reference station at ordinate Wabe watershed 2007-2016

Figure 3-4 shows annual rainfall of reference meteorological gauge (2007-2016). The red dotted line between the clustered column separates forcing data used for model warm up, and validation (left side) and satellite rainfall estimates performance and calibration (right side). The analysis demonstrates there is a consistent accumulated rainfall pattern for six stations from 2007 to 2011 with relatively higher rainfall accumulation in 2010. Whereas, increased rainfall accumulation is observed for Gubire station from 2012-2016 and relatively decreased rainfall accumulation was seen at Wolkite station in 2015 and 2016 even if both stations are located at similar elevation range (see
Table 2-1 Summary meteorological data collected from NMAE during fieldwork
).

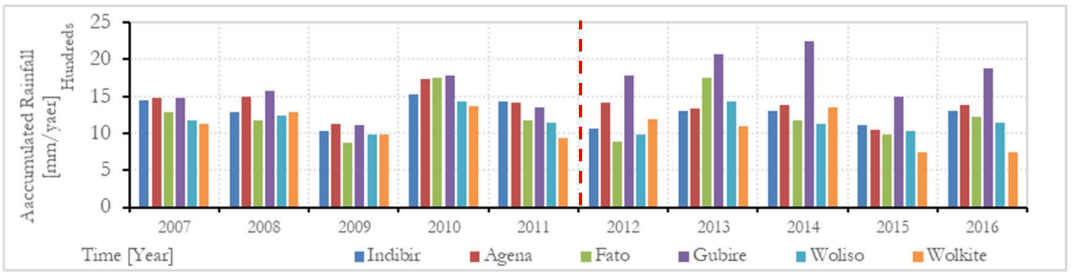


Figure 3-4 Annual rainfall of the reference meteorological station from (2007-2016)

Further analysis of the measured reference rainfall forcing from 6-gauge stations in mean monthly and mean annual bases in terms of standard deviation (STDEV) and coefficient of variation (CV) is shown below. This subsequent mean monthly and yearly based analysis aids to look at the trends of rainfall and to ensure they are free from anomalies for further analysis. Firstly, Figure 3-5 A, B and C depict the mean

monthly based average, standard deviation and coefficient of variation for the reference gauge stations from 2007-2016. From this analysis (see Figure 3-5A) the dry and wet seasons is shown. Accordingly, June, July and August (inside red dotted line) is the major three rainy season showing similar rainfall pattern except for increase at Gubire. September is the end of the rainy season whereas May is the beginning of the rainy season. Furthermore, the rainfall trend for the dry season of November to February shows similar patterns for all stations. According to NMAE major categorization September to May is the dry season. However, April, May and September showing more accumulation compare to other dry seasons.

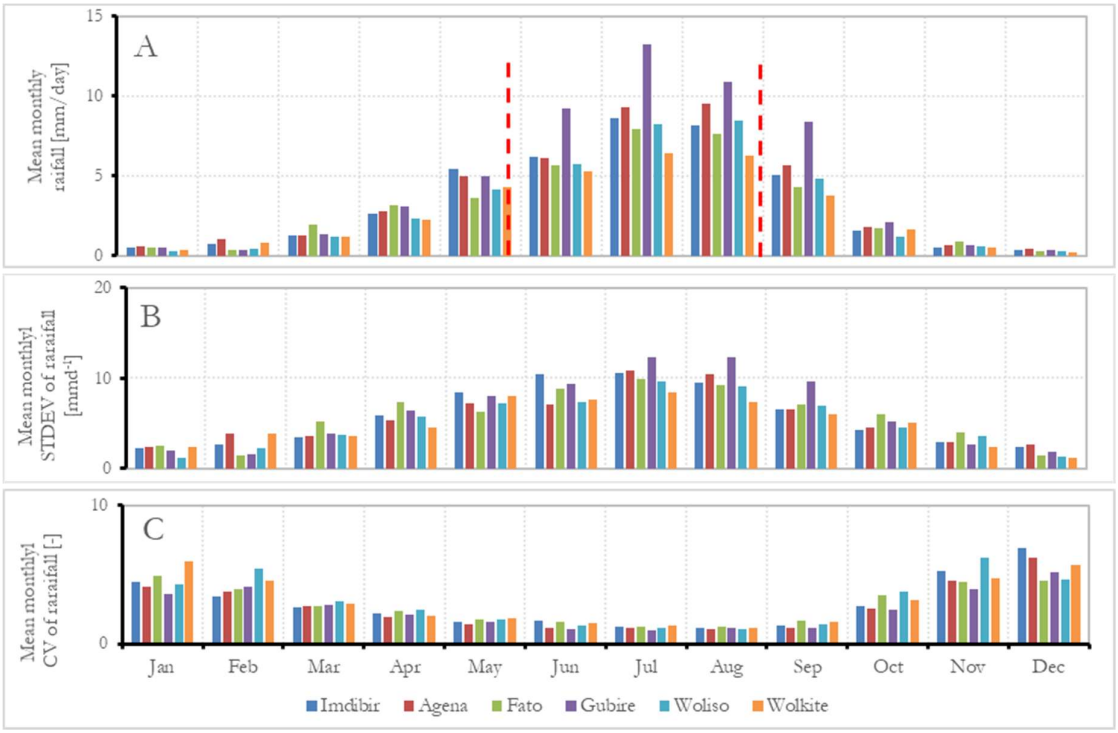


Figure 3-5 Mean monthly rainfall(a), mean monthly standard deviations (B) and mean monthly coefficient of variation(C)of the reference meteorological station from (2007-2016)

The mean monthly rainfall in the wet season is up to 8.5mm and 8mm for Agena and Fato stations respectively located in highest elevation compared to others. The monthly average at Gubire stations reaches 13 mm in July and showing a different trend Figure 3-5A. Figure 2-6 A, B and C shows mean annual average, standard deviation and coefficient of variation for the reference meteorological gauge station from the 2007-2016-time window. Generally, the average rainfall of 2009, 2012 and 2015 ranges ~3mm whereas 2010 ranges ~4.2mm. This trend in all other years shows ~3.8mm besides the suppressed range of 5.7mm at Gubire station in 2013. The mean annual CV is ~2.

Figure 3-6 Mean annual coefficient of variation (A), standard deviation (B) and coefficient of variation (C) of rainfall from 2007-2016

3.2.3. Potential Evapotranspiration

Determination of daily potential evaporation is mainly dependent on available in-situ PET variables (Djaman et al., 2015; Gao et al., 2017; Hargreaves and Allen, 2003). For this study maximum and minimum temperature is collected from five-gauge location Figure 3-7 (A) whereas RH, SH and WS are available only at Woliso station Figure 3-7 (B) with more than 70% of missed data. Also, it is evidenced in Table 3-2 Agena has more missed temperature data (32.9%) and 77% missed sunshine duration at Woliso station. In one hand, only one station (Woliso) have PET variables of RH, SH, WS even with more than 77% missed sunshine hour. On the other hand, even without missed data of all PET variables one station is not representative fo this study area due to a limited network. Therefore, due to this shortcoming, the commonly applied Penman-Monteith method is not applied.

Due to the paucity of in-situ PET variables and to reduce error propagation in modelling phase the empirical radiation-based method, (Hargreaves, 1985; Hargreaves et al., 2003) were used to calculate in-situ potential evaporation. They evaluated the performance of Hargreaves in Haiti, Bangladesh, Australia, and the United States with modified Penman-Monteith concerning lysimeter measurement and concluded Hargreaves matches reasonably with in-situ. However, the method has a limitation with the area when the maximum and minimum temperature are relatively constant but, is not the issue in this study area. Maximum, minimum and average temperature are shown in Appendix 1C .

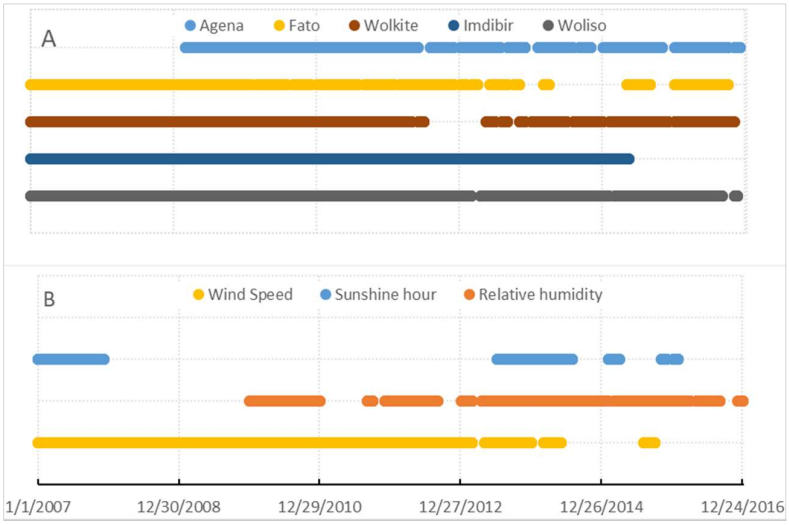


Figure 3-7 Available potential evaporation parameter: (A) temperature and (B) Windspeed, Relative humidity and Sunshine duration (2007-2016)

Table 3-2 Data gap analysis of in-situ potential evaporation variables

Station Name	Elevation [m.a.s.l]	Location		Data gaps [%]			
		Lat	Long	Temperature	Sunshine hour	Relative Humidity	Wind speed
Agena	2310	8.13	38.00	32.9			
Fato	2520	8.46	38.25	28.9			
Wolkite	2000	8.28	37.77	17.4			
Imdibir	2081	8.12	37.94	17.5			
Woliso	2158	8.55	37.98	5.0	77	47	28

Daily potential evaporation was determined using the Hargreaves method (3-3) which is extensively used for limited weather data condition based on maximum and minimum temperature. The data gaps in temperature were assessed and completed using neighbourhood gauge stations. Extra-terrestrial radiation (Ra) for each day of the year and different latitudes can be estimated from solar constant, solar declination and time of the year. Ra is computed in [MJM-2day-1] then converted in to [mm/day] by multiplying a conversion factor the inverse of latent heat of vaporization ($1/\lambda$) is 0.408.

$$ETo = 0.0023\,(T_{mean} + 17.8) * (T_{max} - T_{min})^{0.5} * Ra \qquad 3\text{-}3$$

Where ETo is Potential evaporation in [mm/day], T_{mean} is average temperature, T_{max} and T_{min} are daily maximum, and minimum temperature in [^{0}C] and Ra is extra-terrestrial radiation in [mm/day]. All the necessary equations of Hargreaves ETo are based on FAO 56 documentation. Hargreaves ETo may underpredict or overpredict under high wind speed (U2>3m/s) and high relative humidity respectively (Allen and Smith, 1998).

Figure 3-2 showed potential evaporation parameters at Woliso station for 2013 in Wabe watershed and calculated ETo based on Penman-Monteith (orange) and Hargreaves (blue). To evaluate the applicability of the Hargreaves method in Wabe watershed Woliso station year 2013 with all available PET in-situ parameters are selected. After that, Hargreaves and Penman-Monteith models are prepared in a spreadsheet to determine in-situ potential evaporation. The detailed formulation of the Penman-Monteith method is based on (Allen et al., 1998) and latitude and elevation of Woliso station were used.

3.3. Comparison of FAO-56 Penman-Monteith and Hargreaves ETo methods

Statistics presented in Table 3-3 and scatter plots in
Figure 3-9 were used to compare the accuracy of Penman-Monteith and Hargreaves method. It is noted that in this study, Hargreaves were used as discussed above section 3.2.3. Hargreaves method does not reflect the seasonal variation by producing ETo between 3 to 5 [mm/day] Figure 3-8D. On the other hand, Penman-Monteith reproduces ETo ranging 1.9 (wet season) to 7.2 (dry season) [mm/day] Figure 3-8D. This result is consistent with (Yates and Strzepek, 1994) who evaluates the sensitivity of ETo determination method in 4 basins (Blue Nile, Vistula, East River, and Mulberry) with different spatial scale and climatological behavior.

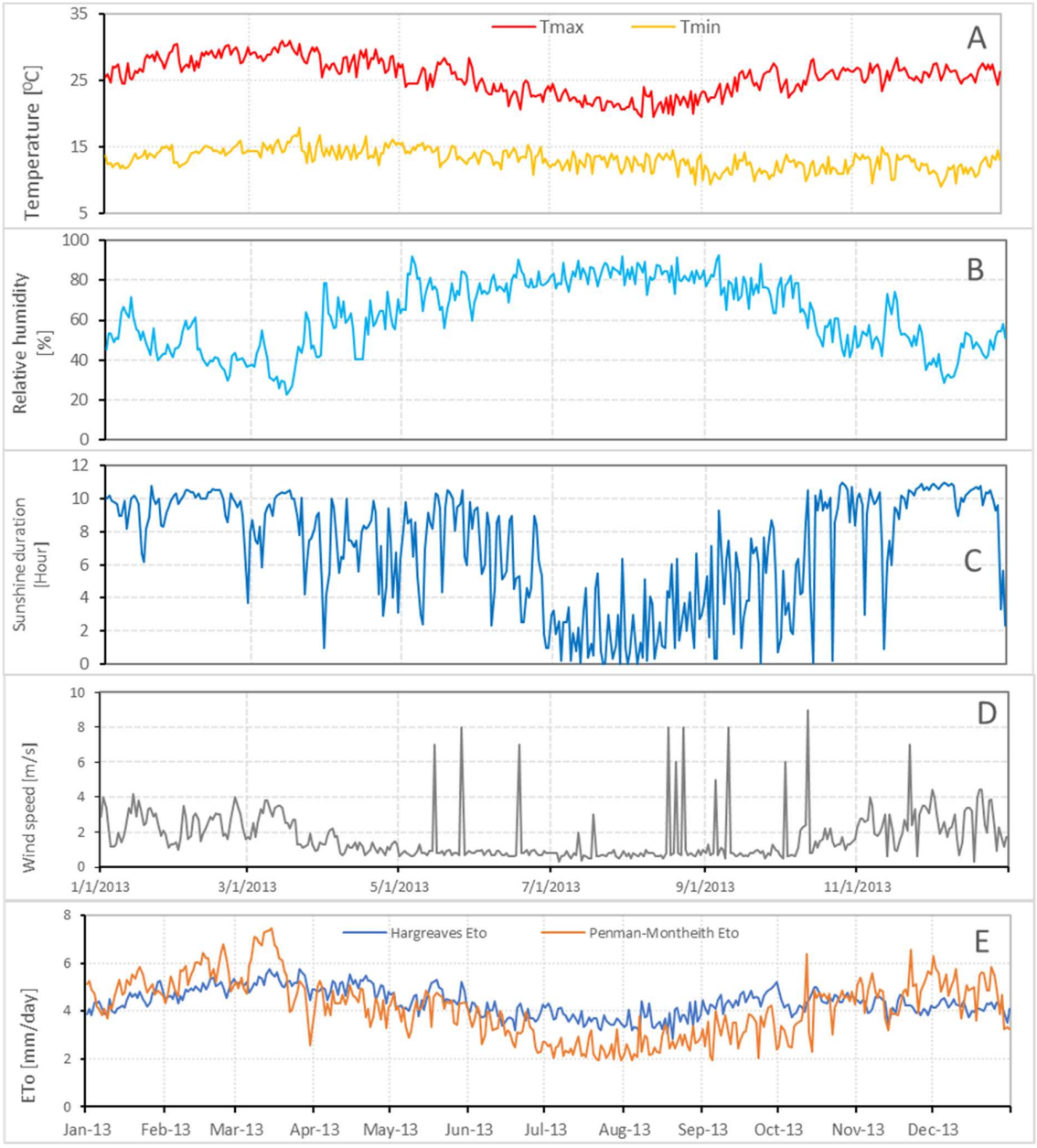


Figure 3-8 Potential evaporation parameters for Woliso station 2013. A is Maximum and minimum temperature, B is Relative humidity, C is Sunshine duration, D is Windspeed and E is potential evaporation from two model

Table 3-3 Statistics to compare Hargreaves and Penman-Monteith ETo method in Wabe watershed

Statistics	Hargreaves ETo	Penman-Monteith ETo
Sum	1596	1508
Max	5.76	7.45
Min	2.82	1.92
Diff/Ratio	88	1.06
Avr	4.37	4.13
STDEV	0.57	1.24
CV	0.13	0.30
CC	0.60	

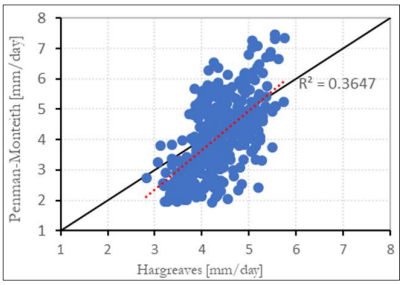


Figure 3-9 Scatter plot comparison of Hargreaves and Penman-Monteith ETo method

The Hargreaves method was used due to the paucity of in situ ETo variables. This potential evaporation is used to evaluate the satellite PET as a benchmark. Qualitative and quantitative assessments were used to select which satellite PET captures well the catchment condition. In this manner, visual inspection, scatter plots, cumulative plots and evaluation metrics of root mean square error (3-4), correlation coefficient (3-5) and mean bias (3-6) are used to select reasonably performed satellite potential evaporation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (P_{s,i} - P_{g,i})^2}{N}} \quad 3\text{-}4$$

$$CC(r) = \frac{\frac{1}{N}\sum_{n-1}^{N} (P_{s,i} - \overline{Ps})(P_{g,i} - \overline{Pg})}{\sigma_s \sigma_g} \quad 3\text{-}5$$

$$ME = \frac{\sum_{i=1}^{N} (P_{s,i} - P_{g,i})}{N} \quad 3\text{-}6$$

Where N is the total number of data elements, P_{si} and P_{gi} are satellite and gauge PET, $\sigma_s\sigma_g$ is the standard deviation of satellite and gauge PET, $\overline{P}_{si}$ and $\overline{P}_{gi}$ are mean of satellite and gauge. These statistics were serving to perceive daily average difference, their distribution and association respectively.

3.3.1. Screening and correcting observed discharge data

Figure 3-10 depicts observed rainfall and discharge time series for 2007-2016 of Wabe watershed at the outlet of the basin. By visual inspection, it is evident that some records are suspicious as indicated in red and yellow doted circles. Observed discharge from MoWIE in the year 2010 and 2014 shows zero measurements (for four days); since there is no neighboring discharge measurement; data from the first and next day of incorrectly measured record were linearly interpolated to fill in the discharge. Although, in 2010 and 2013 the observed discharge shows an outlier, this is not the case in rainfall (indicated in purple dotted circles). Perhaps, this inconsistency between the observed discharge and rainfall may be due to spatial interpolation of rainfall, errors in rain gauge and in stage-discharge rating curve relation. Therefore, further analysis using Double mass curve(Eris and Agiralioglu, 2012; Gao et al., 2017; Searcy and Hardison, 1960) and incremental difference method ((Rientjes et al., 2011) were done to perceive the consistency of rainfall and observed discharge.

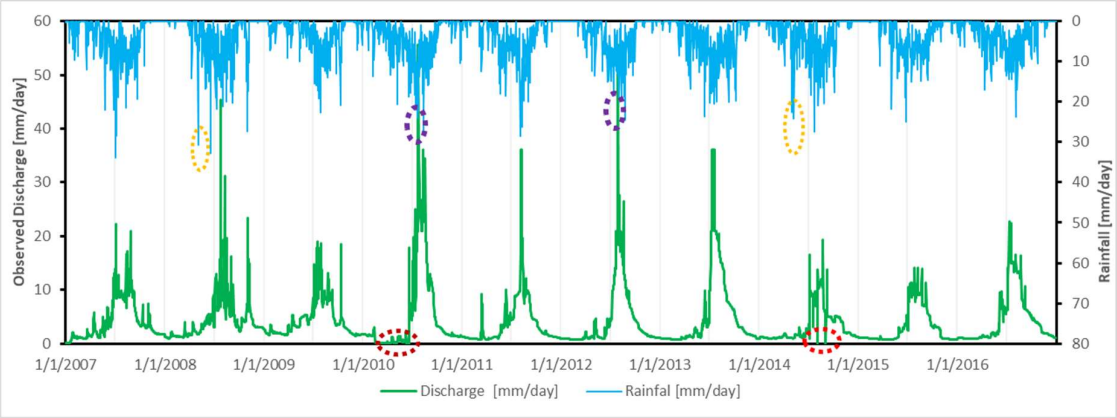


Figure 3-10 Rainfall and observed discharge in daily time serious (2007-2016)

For the hydrological year (start month 6 and ends 5th month) annual runoff coefficient was determined using equation (3-7). As shown in Figure 3-11 (A1) most of the rainfall results in a runoff in the basin. Notably, in the year 2008 and 2010 different sort of runoff is visualized. Extremely very high runoff coefficient (0.9) is shown in 2010.

On the other hand, the lowest runoff coefficient (0.55) shown in 2008. In hydrological perception, this situation is not common as the rainwater infiltrate, evaporate and intercept before going to direct runoff. The catchment groundwater resource evident there is considerable flow during dry and wet seasons this results in a baseflow throughout the year. Overall, the generated runoff depicts the majority of rainfall in the watershed directly converted to flow. As discussed in section 2.2.2 the watershed water level recorder indicated the influence of sediment during the winter season. Mr. Teka Moshag said during rainy season due to sediment load water level increases causes in overestimated daily water level. Perhaps, the runoff coefficient results evidence the sediment effect. Therefore, further investigation is required to identify inconsistently, and outlier observed discharge during rainy season regarding gauge rainfall.

The incremental difference method is adopted (after Rientjes, et al., 2011). The idea in this method is to properly adjust an outlier (mostly during high rainy season) from the observed discharge concerning measured rainfall. Equation (3-8) and (3-9) shows the determination of incremental difference of discharge and precipitation respectively. The procedure is following three steps. First, for the hydrological year increment of precipitation and discharge for each time step were calculated. Second, subsequently, absolute value ΔP and ΔQ are obtained. Then thirdly, the ratio of absolute precipitation difference with observed discharge difference (y-axis) and the reverse are plotted against time domain (x-axis) Figure 3-11 (A and B). The ratio of |ΔPI / ΔQ and the cumulative rainfall against observed discharge Figure 3-11 (A2-green dotted box) clearly shows that there is overestimated discharge in the year 2010. Although of |ΔPI / ΔQ Figure 3-11 (A) and |ΔQI / ΔP Figure 3-11 (B) plot shown As such, this outlier and unreliable measurements were observed, inspected reference to precipitation and appropriately corrected.

$$Runoff\ coefficient = \frac{annual\ measured\ streamflow\ [mm]}{annual\ rainfall\ [mm]} \quad 3\text{-}7$$

$$\Delta Q = Q_t - Q_{t-1} \quad 3\text{-}8$$

$$\Delta P = P_t - P_{t-1} \quad 3\text{-}9$$

Where ΔP is the incremental difference in precipitation, ΔQ is incremental difference in discharge, P_t , P_{t-1} and Q_t, Q_{t-1} is initial and final time step for precipitation and observed discharge respectively.

The incremental difference method analysis result in Figure 3-11A shows that most of the IΔPI/ ΔQ and value lies close to zero. However, some appear largest distance from zero as outlier particularly in 2010. The outliers were properly adjusted by linearly inerpolating based on pervious and next day gauge rainfall measurement.

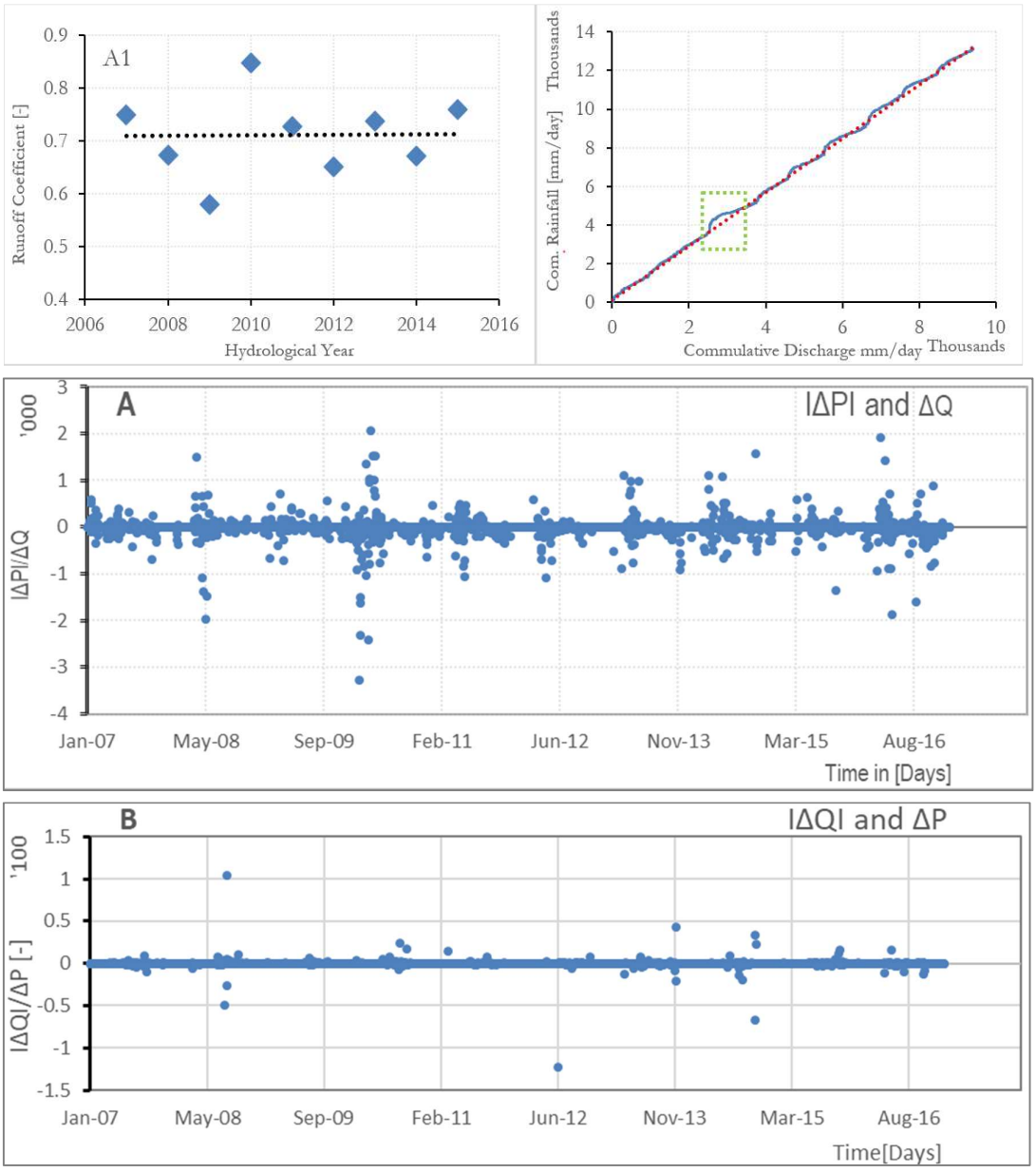


Figure 3-11 Rainfall-discharge relation by runoff coefficient (A1), by double mass curve (A2), the ratio of the incremental difference of IΔPI and ΔQ (A) and Ratio of IΔQI / ΔP (B)

Figure 3-12 shows appropriately adjusted rainfall, discharge, and potential evaporation. The consistent pattern is shown in both wet and dry season. However, some rainy days on the onset of wet (2010,2016)

and end of the wet period (2013,2014) is not observed from observed discharge. Also, it is shown that there is always baseflow indicating a perennial river.

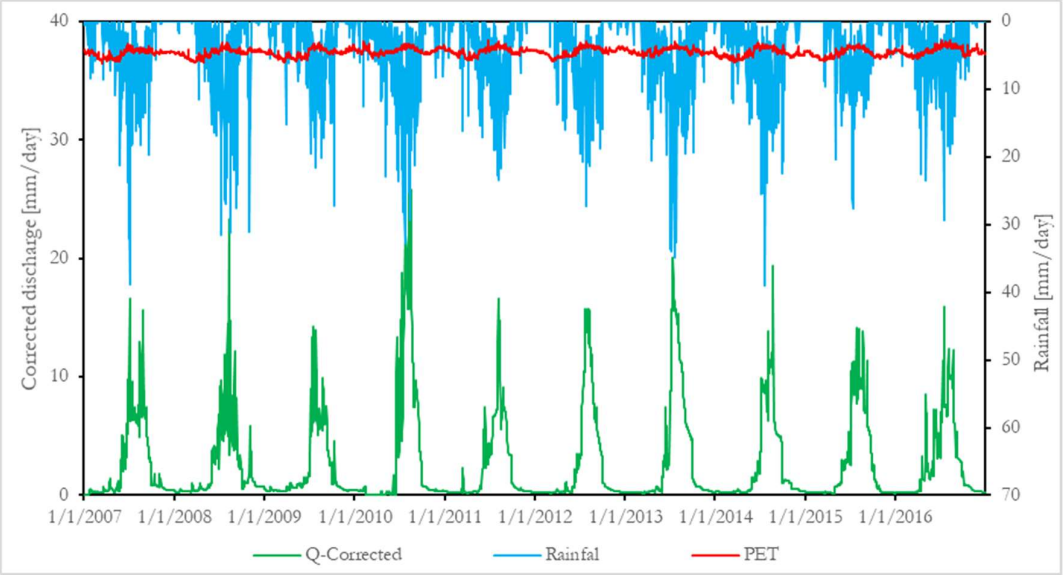


Figure 3-12 Corrected observed streamflow, rainfall and Hargreaves ETo (2007-2016). PET refers to Hargreaves potential evaporation

3.4. Spatial representativeness of the in-situ meteorological data

The conceptual semi-distributed rainfall-runoff model HBV Light version requires a time series of the meteorological forcing for each sub-basin. As shown in Figure 3-13 the distribution of meteorological gauge station is poorly distributed in Wabe basin particularly the mountainous region of north eastern part. Limited gauge measurements were spatially interpolating to sub-basin as well as the whole basin. In hydrology, there are several interpolation techniques which are suitable for different catchment characteristics. Based on studies from (Omondi, 2017) and (De Silva and Ratnasiri, 2007) Thyssen polygon spatial interpolation method (Thyssen, 1911) was adopted in this study ([3-10]). Omondi, (2017) was tested commonly used inverse distance weighting and Thyssen polygon method for the Kabompo river basin in Zambia and selected Thyssen method based on statistics root mean square error, mean absolute error and Pearson's correlation coefficient. The Thiessen method formula and naming are directly used after (De Silva et al., 2007) as follows. The weights of rain gauges are determined by their relative areas, from the Thyssen polygon network. Although, average weighted (global value) elevation of respective gauge station were assigned for precipitation and temperature were used in catchment setting.

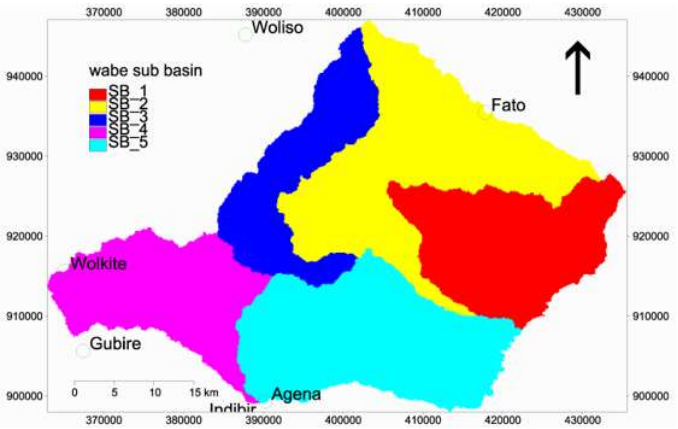


Figure 3-13 Reference met-gauge distribution Wabe watershed

$$P_x = \frac{\sum_{i=1}^{N}[(A_j - A_i)P_i]}{\sum_{i=1}^{N}(A_j - A_i)} \quad 3\text{-}10$$

Where Aj and Ai is Thiessen polygon area when a station with missing value is excluded and included respectively, Pi is precipitation of surrounding station, Px is precipitation to be estimated and $\sum_{i=1}^{N}(A_j - A_i)$ is Thiessen polygon area for the station with missing values.

3.5. Satellite rainfall and potential evaporation products preprocessing

In this study satellite-based rainfall (except CMORPH/ downloaded ½ hourly) and potential evaporation products at the daily base were downloaded from their respective archive via GeoNETCAST ISOD toolbox in ILWIS385x version. The ILWIS software is effective remote sensing and GIS application and freely available from https://www.itc.nl/ilwis. Satellite rainfall products from ARC2, CMORPH and CHRIPS, and potential evaporation products from FEWSNET and PET-20km were downloaded and processed for time window (2012-2016). The time window selection is based on the available rainfall and observed discharge data. The CMORPH raw data for one day contains 48 rain rate maps for each ½ h time step in [mm/hr]. Scripts were written in customized ILWIS software to subsequently get aggregated rainfall in mm/day see Figure 3-14. Three ILWIS script is coded to get a daily aggregated value of ½ hourly rainfall [mm/day]. After that, the products are subset to AOI Wabe watershed, and map lists were created. Finally, the point map of reference gauge location was crossed with the satellite products to retrieve rainfall values for each location then exported to MS Excel worksheet for further analysis.

Figure 3-14 shows processing steps for satellite rainfall and potential evaporation estimates.
The satellite products were exported to MS Excel worksheet in daily temporal resolution and their original spatial resolution. As stated in section 3.4 pixel scale satellite rainfall and PET estimates were interpolated to sub-basin as well as the whole basin using Thiessen method (Thyssen, 1911).

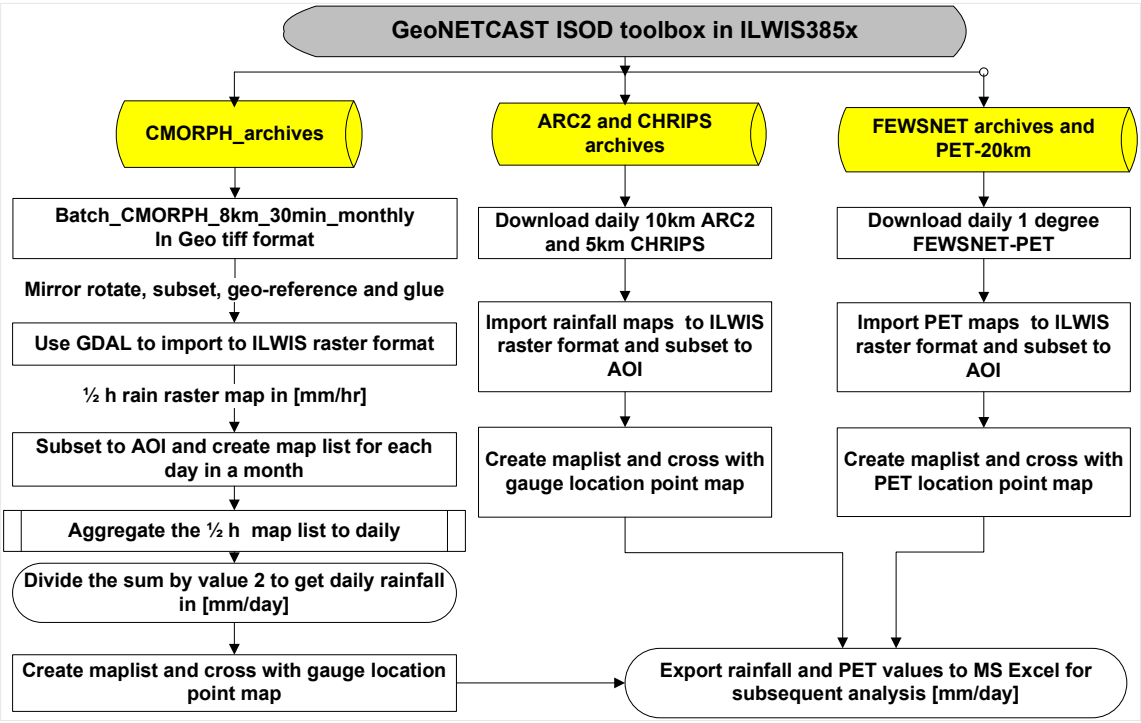


Figure 3-14 Processing sequence for satellite rainfall and potential evaporation estimates

3.6. Performance of SRE from CMORPH, ARC2, and CHRIPS

The inter-comparison is based on point to pixel approach. Reference meteorological gauge stations were used for the performance analysis. The performance of satellite rainfall estimate has been compared using scatter plots, Taylor diagram (Taylor, 2001), double mass curve analysis and statistical criteria. The scatter plots and statistical evaluations are made on a daily basis to see how the uncorrected satellite rainfall estimates perform concerning gauge measurements. The performance of CMORPH, ARC2, and CHRIPS are evaluated by using different statistics as there is no one statistical parameter summarizes the ability of an algorithm with respect to the ground measurements(Ebert, 2007; Jobard and Roca, 2011; Mashingia and Bruen, 2014; Thiemigsss et al, 2012). The standard statistical measures adopted in this study includes root mean square error (3-11), Mean error (3-12), Relative Bias (3-13) and Pearson's correlation coefficient (3-14). In addition to this statistics average, standard deviation, total, maximum and coefficient of variation were analyzed.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (P_{s,i} - P_{g,i})^2}{N}} \quad 3\text{-}11$$

Measures magnitude of accumulated errors and spatial variation. Range: 0 to infinity. Perfect score: 0

$$Bias = \frac{\sum_{i=1}^{N} (P_{s,i} - P_{g,i})}{N} \quad 3\text{-}12$$

Measure a systematic difference. Range minus infinity to infinity. Perfect score:0

$$Rbias = \frac{\sum_{i=1}^{N} (P_{s,i} - P_{g,i})}{\sum_{i=1}^{N} (P_{g,i})} \quad 3\text{-}13$$

Measure a systematic difference. Range minus infinity to infinity. Perfect score:0

$$CC(r) = \frac{\frac{1}{N}\sum_{n-1}^{N} (P_{s,i} - \overline{Ps})(P_{g,i} - \overline{Pg})}{\sigma_s \sigma_g} \quad 3\text{-}14$$

A measure of association (to evaluate agreement between SRE with gauge). Ranges between -1 (negative strong correlation) to +1 (positive strong correlation).

Where N is the total number of data elements, P_{si} and P_{gi} are satellite and gauge rainfall at day i, σ_s and σ_g is a standard deviation of satellite and gauge PET, $\overline{P}_{si}$ and $\overline{P}_{gi}$ are mean of satellite and gauge rainfall values The statistics are commonly applied by authors Haile et al., 2013; Liu et al., 2011; Tian et al., 2009; Zotarelli and Dukes, 2010).

3.7. Bias decomposition and detection capability of satellite rainfall estimates

The bias components and detection capability of CMORPH, ARC2, and CHRIPS were assessed following two steps. First, the total bias in each product is decomposed into three components i.e. hit bias, missed bias and false bias events following methods in (Habib et al., 2012; Wilks, 2006). The three components are calculated based on the equation (3-15) as follows.

$$HE = \sum_{i=1}^{N} (P_s - P_g), if\ (P_s > 0 and P_g > 0) \quad 3\text{-}15$$

$$FE = \sum_{i=1}^{N} P_s, if\ (P_s > 0 and P_g = 0)$$

$$ME = \sum_{i=1}^{N} P_g, if\ (P_s = 0 and P_g > 0)0$$

Where HE =hit events, FE = false events, ME = missed events, P_s and P_g are satellite and gauge rainfall.

Secondly, the ability of each product to detect rainfall is evaluated using categorical measures based on Table 3-4 for the verification of satellite estimates concerning the in-situ, i.e., POD, FAR and CSI see equation (3-16). In both steps, the analysis was done for satellite rainfall estimates overlaying 6-gauge stations in daily bases (2012-2016). As defined by Montero-Martínez et al, (2012) this categorical statistics for satellite rainfall forecasting verification has two probabilities (they are dichotomous) with 0% perfect for FAR or 100% for POD and CSI.

Table 3-4 contingency table used to define categorical measures

	Gauges rain	Gauges no rain
Satellite rain	HE	FE
Satellite no rain	ME	?

$POD = \frac{HE}{HE+ME}$ indicates correctly detected rainfall. Ranges 0 to 1 3-16

$FAR = \frac{FE}{HE+FE}$ indicates the proportion of falsely detected rainfall. Ranges 0 to 1

$CSI = \frac{HE}{HE+ME+FE}$ measures the fraction of correctly identified rainfall. Range 0 to 1

The best score for POD and CSI is 1 whereas for FAR is 0.

3.8. Rainfall distribution and effect of seasonal variability in CMORPH, ARC2, and CHRIPS

The satellite rainfall performance differs from catchment to catchment in detecting rainfall distribution and capturing seasonal dynamics. Classification of rainfall intensity varies with the desired objective, for instance, Chaemiso et al., (2016) uses four class 1) small intensity (<4mm/hr), 2) moderate intensity (4-8mm/hr), 3) high intensity (8-12mm/hr) and 4) very high intensity (>12mm/hr) with the objective of seasonal variation of rainfall characteristics over Peninsular Malaysia. The other unpublished MSc research study was done by Omondi, (2017) uses 6 rain classes to analyze the performance and rainfall distribution detection capacity of TRMM, CMORPH, and CHRIPS in Zambia, Kabompo basin. To evaluate the performance of SRE's products in detecting different rainfall depth, rain rates are partitioned in to 6 rain classes. Similarly, to assess the seasonal dynamics, grouping into wet and dry seasons was done (Table 3-5).

Table 3-5 Two seasons and rainfall distribution clustering in Wabe watershed

Rainfall distribution	0-1, 1-2.5, 2.5-5, 5-10, 10-20 and >20		(Gumindoga et al., 2016; Omondi, 2017)
Seasonal dynamics	Wet	June, July, and August	(Awulachew et al., 2007; Chaemiso et al., 2016; Habib et al., 2014a)
	Dry	January, February, March, April, May, September-December	

3.9. Satellite rainfall estimates bias correction

Satellite-based precipitation measurements served as a secondary source of data for many hydrological applications. However, as an indirect source of measurements, they are exposed to errors. As argued by authors Tian et al., (2009), Vila et al., (2009), Habib et al., (2012) and Gumindoga et al., (2016) these errors need to be understood, indicated, partitioned and corrected. Studies of Habib et al., (2014) and Artan et al., (2007) shows that models could increase or suppress precipitation systematic errors to larger or smaller simulated discharge depending on the response mode of the hydrological model.

In this study, four bias correction methods are evaluated. For each method, the bias correction factor was calculated, and the raw satellite rainfall is multiplied by the bias factor to get the corrected SRE. The best performed bias scheme was selected and used for streamflow simulation and water balance assessments.

3.9.1. Distribution transformation bias scheme

This method is modified and applied after (Gumindoga et al., 2016). The method is based on the formulation of two bias factors which corrects for both mean and variation. The first bias correction factor for the mean (3-17) and a second correction factor for variance (3-18) is established. Here five-day window is used for determination of both mean and variation bias factor. Different days 3,4, 5, 6 are tested for the DT scheme in matching with gauge measurements. The selected evaluation scheme result shows five days analysis reasonably matches the bias corrected SRE's with gauge measurements. Then subsequently, the two formulated bias factors were used to correct the systematic error in satellite-based rainfall (3-19).

$$DT_{\mu} = \frac{D\mu}{S\mu} \quad 3\text{-}17$$

Where DTμ is the bias correction factor for the mean, Sμ and Dμ are 5 days mean value of individual satellite pixel and gauge location rainfall estimate

$$DT_{\tau} = \frac{D\tau}{S\tau} \quad 3\text{-}18$$

Where: DTτ is the bias correction factor for the variation which is calculated from the quotient of the standard deviation, Sτ, and Dτ, for satellite and gauge.

$$S_{DT} = DT\tau\left(S_{i,t} - S_{\mu}\right) + DT_{\mu}S_{\tau} \quad 3\text{-}19$$

Where SD_T is corrected satellite rainfall estimate

3.9.2. Spatio-temporal bias correction

3.9.2.1. Space-time variant (T_VS_V)

The first approach in the method is determining the bias factor (3-20). The bias is corrected for individual satellite pixels that overlay rain gauge stations. Correction is at daily time step. The multiplicative BF for a specific satellite pixel at a selected day (d) and gauge (g) with respect to reference gauge formulated in (3-18). The formulation of bias factors considers three assumptions

- Minimum accumulated gauge rainfall depth for the specified time window
- Minimum number of rainy days for the specified time window
- If one or both conditions are not true what should be the bias factor

This bias factor formulation assumption differs from catchment to catchment as stated by different authors, for instance, Bhatti et al., (2016) and Habib et al., (2014) used 7-day sequential window and Gumindoga et al., (2016) used a 10-day sequential window. In this study the bias factor is calculated based on 6-day's sequential window, minimum of 5mm rainfall accumulation and at least three rainy days within the selected rain-day window is recorded. The adequacy of six-day window in removing bias from uncorrected SRE's products is evaluated by RMSE, CC and accumulated rainfall amount. This analysis is consistent with (Bhatti et al., 2016) who evaluated window size ranges between 3 to 35 at Gilgel Abay catchment (Ethiopia) and stated systematic errors in SRE,s do not increase window size above 7.

$$BF_{TVSV} = \frac{\sum_{t=d}^{t=d-l} G\ (i,t)}{\sum_{t=d}^{t=d-l} S\ (i,t)} \quad 3\text{-}20$$

$$BF_{TFSF} = \frac{\sum_{t=1}^{t=T}\sum_{i=1}^{i=n} G\ (i,t)}{\sum_{t=1}^{t=T}\sum_{i=1}^{i=n} S\ (i,t)} \quad 3\text{-}22$$

Where: S and G are daily satellite and gauge rainfall estimates respectively, l stands for the length of the time window for bias estimation, d refer to certain day, i and t refer to gauge location and Julian day number respectively.

3.9.2.2. Space-fixed-time variant (T_VS_F)

This study follows the approach of Habib et al., (2014) who tested T_VS_F (3-21) and T_FS_F (3-22) to "assess implication for accounting and ignoring of the variability of bias." The second spatio-temporal bias scheme is called time variant space fixed which corrects at daily time step for every pixel.

$$BF_{TVSF} = \frac{\sum_{t=d}^{t=d-1}\sum_{i=1}^{i=n} G\ (i,t)}{\sum_{t=d}^{t=d-1}\sum_{i=1}^{i=n} S\ (i,t)} \quad 3\text{-}21$$

3.9.2.3. Time and space fixed (T_FS_F)

In this approach, the bias factor is obtained for the entire domain and over the total duration of the sample. The bias factor formulation is based on equation (3-21). The uncorrected satellite estimate is multiplied by BF_{TFSF} to get a bias corrected estimate in spatially and temporally lumped scheme. This means one bias factor applies for entire time series.

3.10. Catchment Partitioning

The main reason to partition the Wabe watershed is because of its significant elevation variation (see section 2.2.5 and Figure 2-1); variation in soil types, due to the large catchment area (1846 km$^{2)}$ and, to use the spatially distributed data from SRE's products. Rientjes, (2015) indicates the partitions are made when the catchment is geographically or climatologically heterogeneous. There are different approaches to partition a watershed into sub-basins like the availability of gauge station outlet in each sub-basin, similar land cover, and similar soil type etc. For this case; stream discharge outlets with no gauging stations except the Wabe main outlet collected during fieldwork from MoWIE of Ethiopia were used for sub-basin partition. The main inputs for catchment partitioning are shuttle radar topographic model (SRTM 30m DEM), watershed outlet and sub-basin outlets. All the three inputs are used to delineate and extract Wabe watershed boundary, partition the entire Wabe basin to sub-basins and extract the entire basin and sub-basin area.

Table 3-6 shows data and tools used in DEM hydro-processing. The procedure is based on (Merwade and Rajib, 2014). The ArcSWAT software requires input DEM in UTM projection. Therefore, the DEM tiles were projected in UTM zone 37N with WGS84, 1984 datum and ellipsoid respectively. As shown in Figure 3-15 the downloaded elevation data are mosaiced, masked to the area of interest after that fill the sink to remove depression, flow direction, and flow accumulation to find the drainage pattern was done. Following this step, the main outlet of the whole basin used to delineate the watershed. Subsequently, sub-basin outlets where used to partition the whole watershed into five sub-basins.

Table 3-6 Data and tools used for watershed delineation and sub-catchment partition

#	Input	Format/#	Result	Tools used
1	SRTM DEM 30m	Tiff/1	Analysis of catchment	ArcSWAT 2012.10_5.21 and ArcGIS_10.5.1
2	Catchment outlet	Point/1	Wabe basin	
3	Sub-basin outlet	Point/4	Wabe sub-Bains	
The software is freely accessible from		https://swat.tamu.edu/software/arcswat/		

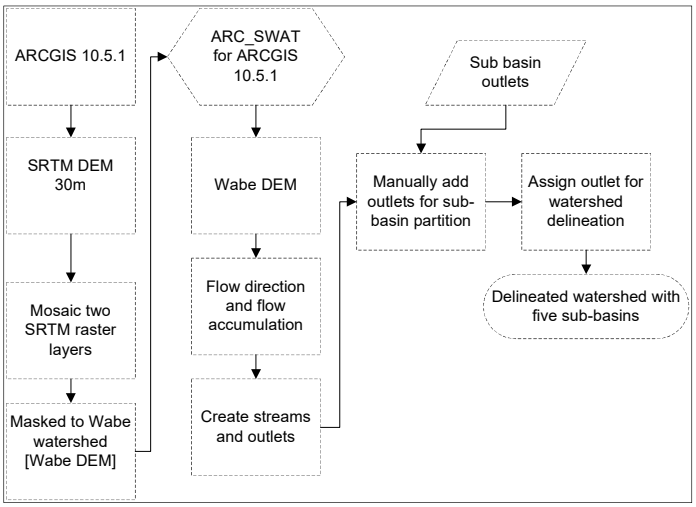


Figure 3-15 Steps followed in ArcSWAT for catchment partitioning

Figure 3-16 shows Wabe watershed boundary, outlets used for partitioning sub-basins (green boxes), five partitioned sub-basin and area for each sub-basin in m² Table 3-7. Numbers from 1 to 5 indicates the outlet of each sub-basin. Among this outlet points, the whole basin water level is located at outlet 4 (observed discharge of the whole watershed is measured) and 1 through 5 is planned to install water level measurement by MoWIE of Ethiopia.

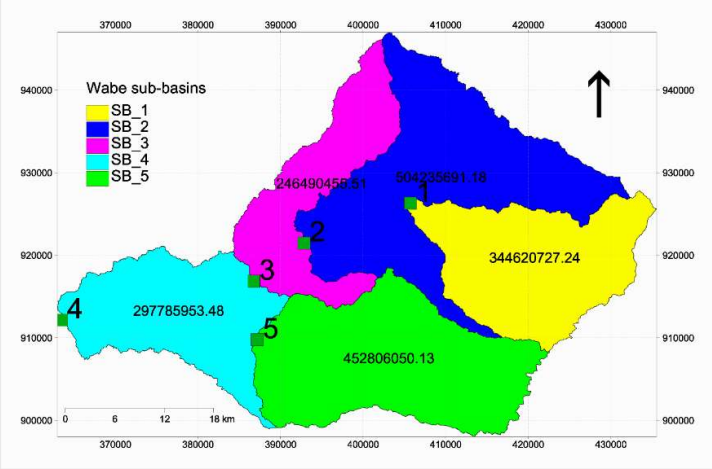


Figure 3-16 Outlet points used to partition sub-basins (green filled small boxes) with numbers 1 to 5 indicating sub-basin number Partitioned sub-basins, and area for each sub-basin.

Table 3-7 Wabe outlets location and area of sub-basins

Outlets ID	Subbasin	Location		Area of sub-basin	
		X	Y	Area [m²]	Area [km²]
1	SB_1	405755.3	926299.9	344620727.24	345
2	SB_2	392860.6	921464.8	504235691.18	504
3	SB_3	386791.7	916836.2	246490456.51	246
4	SB_4	363537.4	912147.7	297785953.48	298
5	SB_5	387214.1	909758.6	452806995.13	453
The total area of Wabe watershed				1845939822	1846

3.11. Hydrological Modelling HBV Light version approach

In this study a modified version of Hydrologiska Byråns Vattenbalansavdelning (HBV) light model by (Seibert and Vis, 2012) was used for streamflow simulation. The selection of the model is due to: originally developed for runoff simulation, simple structure to understand, and apply, is less complexity, and it has been extensively evaluated in different regions of the globe (Seibert and Vis, 2012). The main principle of the model is based on the general water balance equation to simulate the hydrological process on surface and subsurface zones.

The HBV Light model has four routines; snow and rain routine, soil moisture routine, response function routine, and routing routine. But in the Wabi watershed, the temperature is higher than the threshold value as such, the snow routine was ignored, and precipitation is only in the form of rainfall. The model has different structures. In depth details and descriptions about the model routines, model structures, model variants and model parameter overview is refered to (Seibert and Vis, 2012). In this study the standard model structure with basic model variant was used. The main inputs to the model are meteorological forcing of rainfall and potential evaporation [mm/day], temperature [0^c], elevation and vegetation zones, and catchment file. The main outputs are actual evapotranspiration [mm/day], simulated streamflow [mm/day], amount of water in soil box [mm], storage in upper and lower groundwater box in [mm].

Figure 3-17 shows the input data and processing of HBV Light version rainfall-runoff modelling.

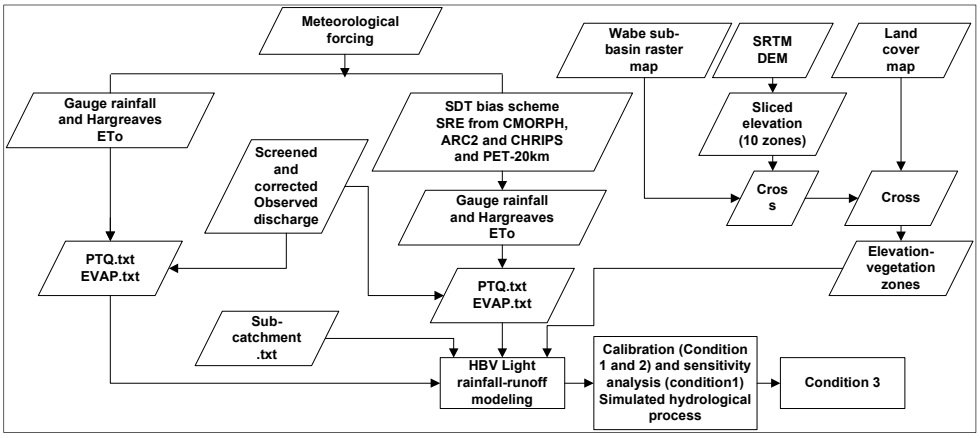


Figure 3-17 Flowchart showing HBV Light rainfall-runoff model input parameters and processing

HBV Light model catchment Setting

The catchment setting needs more attention and should follow real catchment characteristics (Rientjes, 2015; Seibert, 1997; Seibert and Vis, 2012). In this step, three main catchment conditions were set up. 1) a number of elevation zones, 2) a number of vegetation zones and 3) height incremental variables of rainfall and temperature. For the first condition the model allows up to 20 elevation zones (Seibert and Vis, 2012) but for Wabe watershed 10 elevation zones were used as elevation varies from ~1700m to 3600m above sea level. In this manner, the elevation was partitioned within 200m interval. And mean elevation (e.g. for range from (700-800) 600m is characteristic elevation was assigned for each zone.

Figure 3-18 depicts the sliced elevation zones of Wabe watershed with sub-basins. ILWIS software is used for slicing. First domain for each elevation zone is created and subsequently, ILWIS image processing slicing operation was used to portioning the elevation.

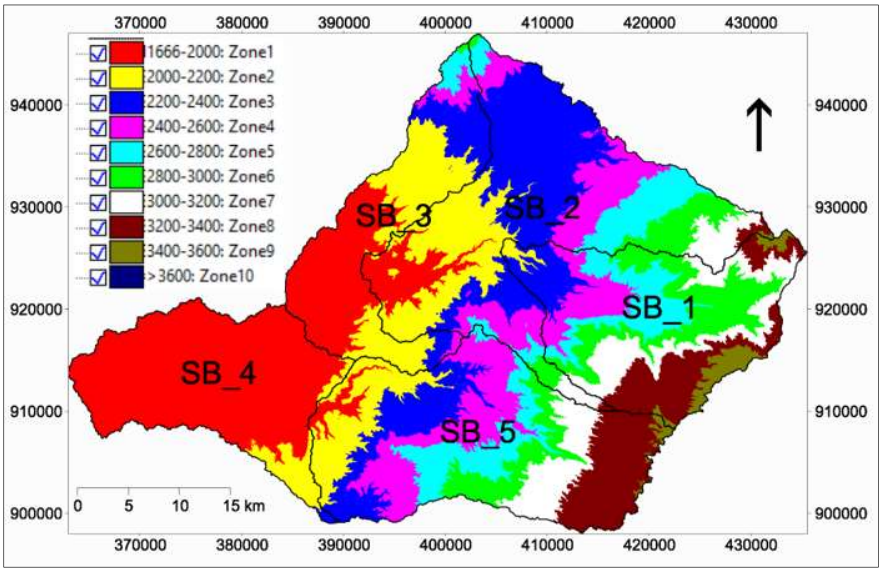


Figure 3-18 Sliced elevation zones of Wabe watershed with a 200m range based on SRTM DEM 30m

Vegetation Zones

Assumption: The change in land cover for study time window is insignificant.

For this study landcover map collected from MoWIE is not used. The reason is the data is too old and could not represent the current catchment characteristics as discussed in section 2.1.2 and Figure 2-3. ERDAS IMAGINE 2016 version software is used for classification. The procedure is based on (Hall et al, 2004; Islam et al, 2018) method Figure 3-19. The downloaded map is imported to the software, stacked and sub-setted to an area of interest. Subsequently; the signature editor is used to collect samples to classify the map. Thereinafter, the maximum likelihood parametric rule is used to classify the map. The advantage of this method is; first, it considers variability within a cluster, and secondly consider the shape, the size, and orientation of the clusters. There are more than three land cover classes in the study area evidenced from fieldwork duration. However, the HBV model allows only 3 landcover class and lake property. Therefore, to incorporate with the rainfall-runoff model the landcover classes are classified into three dominant classes of cereal crop, forest and built up as there is no visible lake body. To quantify the accuracy of land cover classification commonly applied confusion matrix and kappa coefficient(Congalton and Green, 2009) were used. Kappa statistic reflects the difference between the actual agreement and the agreement expected by chance.

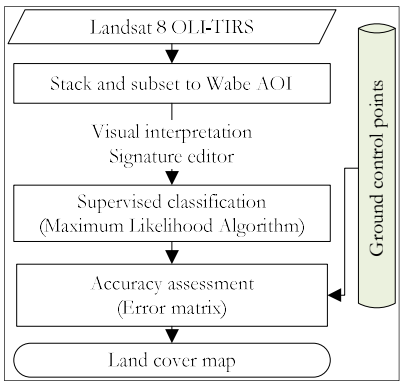


Figure 3-19 Flowchart showing landcover classification

Height increment variables

This section defines how precipitation and elevation values should be corrected with elevation. It can be specified, either a global value for the whole basin or each of the individual sub-catchments. The initialization of the model was very sensitive when using height increment variables (gradients) of precipitation per sub-catchment based on gauge contribution. Therefore, further investigation of precipitation lapse rate is as follows.

Inventory on change of precipitation with the lapse rate

In the study area, elevation of meteorological gauge station varies from minimum at Wolkite 2000 m.a.s.l, to maximum at Fato 2520 m.a.s.l and average elevation of 2214 m.a.s.l (see

Table 2-1 Summary meteorological data collected from NMAE during fieldwork

). Whereas, the variation of elevation in the Wabe watershed is 1700 in the outlet of the whole basin to ~3600m in the upstream with a difference of 1900m see **Error! Reference source not found.**. This indicates there is a significant variation of elevation in the study area. As such further investigation carried out to perceive a change in rainfall with elevation using gauge and two SRE products (CHRIPS and ARC2). The analysis was done for the rainy months of June 2014 and 2015.

3.12. Model calibration, sensitivity analysis, and validation

Any rainfall-runoff model must be proven for its reliability, accuracy and predictive ability as it affected by all the entered input data (modified after Rientjes, 2016). The model was initialized, calibrated and validated using in-situ measurements (2007-2016). The input data were split into three periods warm up (2007 -2008), calibration (2012-2016) and validation (2009 and 2011). The simulation is started at the end of the dry season since the catchment is in steady-state mode with low but stable streamflow discharge.

The calibration is through Trial and Error parameter optimization where model sensitive parameters are optimized following a selected optimization procedure. Firstly, the baseflow recession was fitted. Secondly, the peak flows of observed and simulated discharge are matched. Subsequently, rising and falling limb, and timing of the peak flows were matched. After that, the overall volume and shape of simulated and observed hydrograph were closely matched. The model has been validated for independent data set using the optimized parameter (2012-2016).

Selection of Objective Functions

The selection of objectives functions depends on the desired purpose, for instance, Madsen, (2000) argued that mainly four objectives are considered to evaluate the performance of the calibrated rainfall-runoff model. These objectives are water balance volume agreement, overall agreement of the hydrograph shape, good agreement of high and low flows. Table 3-8 presents a summary of studies which used different objective functions to evaluate HBV model performance. Monte Carlo refers to an automated parameter optimization method in HBV model. The performance of the model was evaluated qualitatively by visual inspection and quantitatively by NSE (Nash and Sutcliffe, 1970), RVE (Janssen and Heuberger, 1995) and Q_{bias} (Habib et al., 2014a) following equation (3-23), (3-24) and (3-25) respectively.

Table 3-8 Summary of objective functions used in previous studies for HBV calibration process

#	Objective function	Method	Reference	This study
t	NSE, Q_{bias}	Monte Carlo	(Habib et al., 2014a)	Based on these references for this study NSE, RVE, and Y were used
2	NSE, RVE	Monte Carlo	(T. H. M. Rientjes, Perera, et al., 2011a)	
3	NSE, NSE_{log}	Monte Carlo	(Radchenko et al., 2014)	
4	NSE, RVE, NSE_H, NSE_L	Monte Carlo	(Deckers et al., 2010)	
5	NSE, RMSE, Bias	Monte Carlo	(Abebe et al., 2010a)	
6	NSE, LnNSE	Monte Carlo	(Nauditt et al., 2017)	
7	NSE, MeanDiff, RVE	Manual Trial and Error	(Sendama, 2015)	
8	RVE, NSE	Manual Trial and Error	(Likasa, 2013)	
9	NSE, RVE	Manual Trial and Error	(Lyimo, 2015)	

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (Q_{sim(i)} - Q_{obs(i)})^2}{\sum_{i=1}^{n} (Q_{obs(i)} - \overline{Q_{obs(i)}})^2} \quad 3\text{-}23$$

Range: minus infinity to 1. Best = 1, Accept >0.6

$$RVE = \left[\frac{\sum_{i=1}^{n} Q_{sim(i)} - \sum_{i=1}^{n} Q_{obs(i)}}{\sum_{i=1}^{n} Q_{obs(i)}}\right] * 100\% \quad 3\text{-}24$$

Range: minus infinity to infinity. Best = 0%, Accept between ±5%

$$Q_{bias} = \frac{\sum_{i=1}^{n} Q_{sim}}{\sum_{i=1}^{n} Q_{obs}} \quad 3\text{-}25$$

Q_{bias} value of 1 indicates bias-free simulated streamflow, a value larger than indicates overestimation of streamflow whereas a value less than 1 reflects underestimation of simulated streamflow.
Where i is time step, n is a total number of the time step. Q_{obs} is observed discharge, Q_{sim} is simulated streamflow, Q_{obs} with top overbar stands for average observed discharge, NSE is Nash and Sutcliffe efficiency and RVE is a relative volumetric error.

Sensitivity analysis

To quantify on the uncertainty of the calibrated model and to see how the model is sensitive to certain changes in input data sensitivity analysis is a crucial step in modelling (modified after Rientjes, 2015). The sensitivity of the model parameters in affecting the simulated streamflow was carried out. The sensitivity analysis was done by changing one parameter at a time while keeping others fixed. In this procedure, the first focus was on the match between the simulated and observed baseflow. Then, fitting the peak flows, rising and falling limb, and the time gap between the simulated and observed flow were analyzed. In this manner, the effect of parameters on simulated streamflow was checked by changing its value ± (10-60%). The model was run for each parameter, and subsequently, objective functions are calculated to understand and analyze the most sensitive parameter in the watershed. Finally, the objective functions (NSE and RVE) for each result were calculated and graphed for assessment.

3.13. Water balance closure analysis

According to Rientjes, (2015) water balance closure for a closed basin implies that rainfall input will be equal to streamflow and ETa over a multi annual time period. This means the difference is equal to zero or change in storage over the assessment period is zero. Water balance closure analysis is done based on the main principle (3-26).

$$P - ETa - Q = \frac{d}{d(t)}S \quad 3\text{-}26$$

Where: P is precipitation [mm/day], ETa is actual evapotranspiration [mm/day], Q is a streamflow [mm/day] and d/d(t) S is change in water storage [mm/day]

Comparison of streamflow and water balance closure

The modelling result of this study is partitioned into two conditions.
Condition 1: First, the model is calibrated and validated based on gauge rainfall and observed streamflow. Second, to assess how the systematic errors in satellite rainfall estimates propagate in streamflow simulation results, the optimized set of parameters by the use of in-situ rainfall to force the model is used to simulate uncorrected and corrected satellite rainfall forcing. Third, the influence of satellite potential evaporation in simulating streamflow is evaluated by replacing in-situ Hargreaves ETo with satellite-based potential evaporation. Then after, the simulated streamflow-based satellite PET, uncorrected and corrected SRE products are compared with observed discharge.
Condition 2: The optimized parameter set by means of in-situ based rainfall are recalibrated using the bias corrected SRE products from CMORPH, ARC2, and CHRIPS. The optimized parameters and model performance are presented in tabular form. Simulated streamflow from the recalibrated bias corrected satellite rainfall was compared with observed and simulated streamflow with in-situ based gauged rainfall (condition1).

4. RESULTS AND DISCUSSION

This chapter describes and discussed the results obtained in daily bases.

4.1. Landcover map

Figure 4-1 shows Wabe landcover map with three classes. Cereal crop is the dominant 1491 km² (80.76%) of a watershed, the second Forest with an area of 348 km² (18.84%) and the last is built up 7km² covering 0.4% of the Wabe watershed. Note that, cereal crop is merged from different signatures such that Inset, Maize, Barley, Teff, Sorghum and Beans whereas, Forest is merged with Chat and deciduous and evergreen forests.

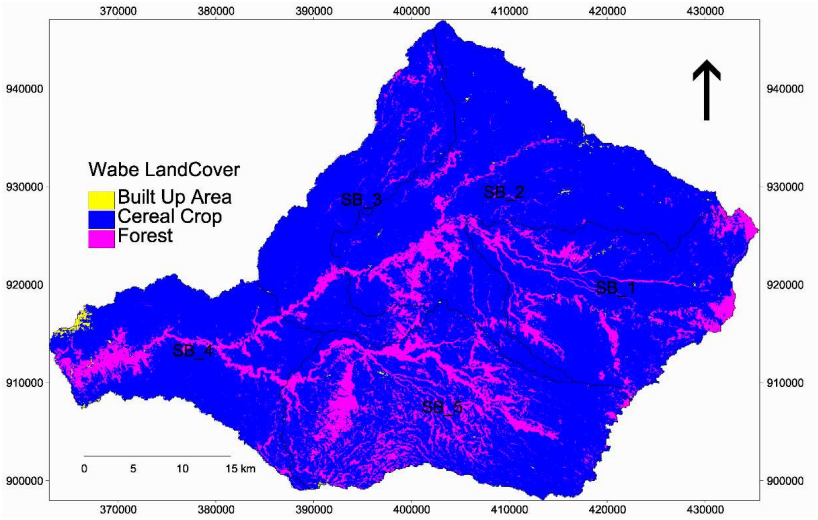


Figure 4-1 Wabe classified land cover map

Table 4-1 depicts the accuracy assessment result of land cover classification. The results of landcover for period 2018 in Table 4-1 showed 74.46 and 0.63 of overall classification accuracy and Kappa statistics respectively which is in substantial range (Rwanga and Ndambuki, 2017). Based on the accuracy result for each class built up and forest pixels were effectively sampled into the ground landcover classes with 1 and 0.96 Kappa statistics respectively. The kappa statistics for cereal crop is 0.46. Perhaps, this lower statistic is due to positional errors in reference (accuracy of GPS) data and errors in a classified map. For example, in Table 4-1, 32 Built up, and 19 Forest pixels are erroneously included (commission error) to cereal crop evidenced with producer's accuracy (omission error) of 25.38 and 77.38 respectively. Similarly, some classes not spectrally separable inset (cereal crop) and forest, traditional built-up area and Forest see section 2.1.2 Figure 2-4.

Table 4-1 Accuracy assessment result for landcover classification

	Reference data			Accuracy totals			Accuracy [%]			Kappa statistics [K^]
	Cereal crop	Forest	Built up	Reference	Classified	Number Correct	Producers	Users	Overall	Overall
Cereal crop	129	19	32	131	180	129	98.47	72	80	0.63
Forest	2	65	0	84	67	65	77.38	97		
Built up	0	0	11	43	11	11	25.38	100		
Total	131	84	43	258	258	205				

4.2. In-situ and Satellite potential evaporation

Figure 4-2 shows the comparison between in-situ, FEWSNET (1^0) and PET-20km in daily temporal scale for Wabe watershed 2012-2016. As discussed in the methodology section 3.2.3 and Figure 3-7 (A, B) based on the availability of meteorological potential evaporation parameters Hargreaves model were used to calculate in-situ ETo. It is shown from time serious plots in Figure 4-2, the Hargreaves potential evaporation ranges 3mmd^{-1} to 5.9mmd^{-1}, FEWSNET PET ranges 0.16 mmd^{-1} to 5.6 mmd^{-1} and the new product PET-20km ranges 1.5mmd^{-1} to 7.5mmd^{-1}. Both satellite estimates show underestimation during the wet season. Overestimation (black dotted rectangle) is shown during the dry period by the product PET-20km.

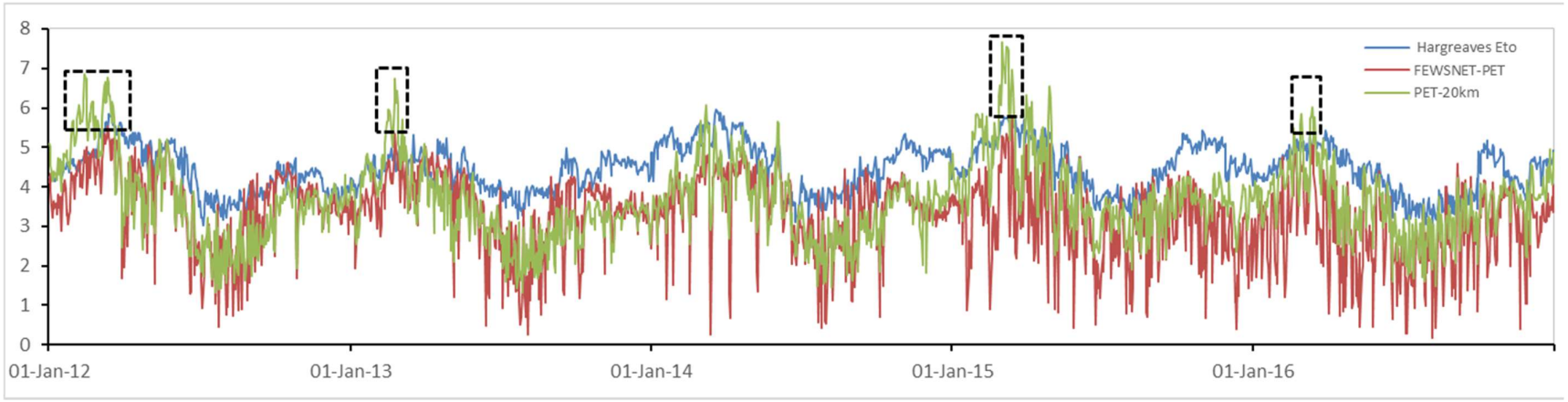


Figure 4-2 Time series showing a daily variation of Hargreaves, PET-20km and FEWSNET estimates at the Wabe basin (2012-2016)

The scatter plots and commutative plots in Figure 4-3 also indicate an overall underestimation of FEWSNET evidenced with reduced cumulative PET depth difference of 2162 whereas for PET-20km is 1481 mm (see Table 4-2). As shown, the PET-20km products match best with Hargreaves with a correlation of 0.7 and RMSE of 1.08. Based on this results in addition to reference ETo PET-20km is selected to see the effect of satellite potential evaporation on streamflow simulation. The overestimation of PET-20km products during low flow condition is adjusted based on Hargreaves ETo for further analysis.

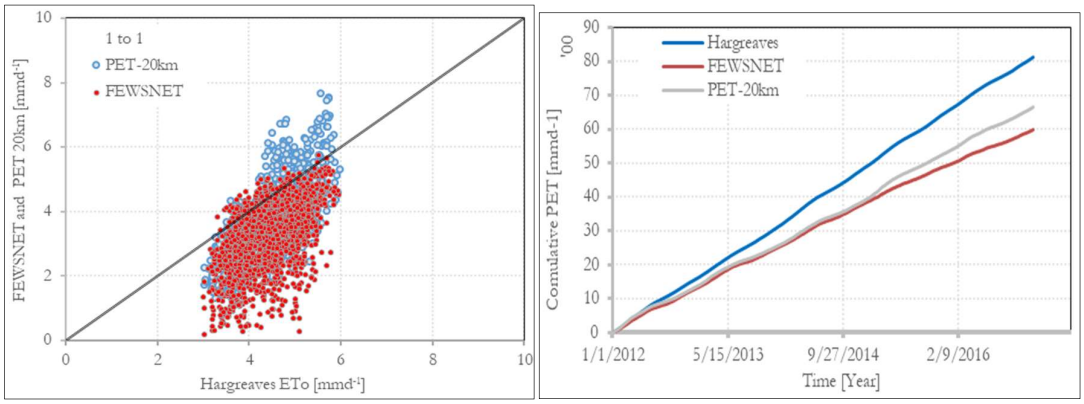
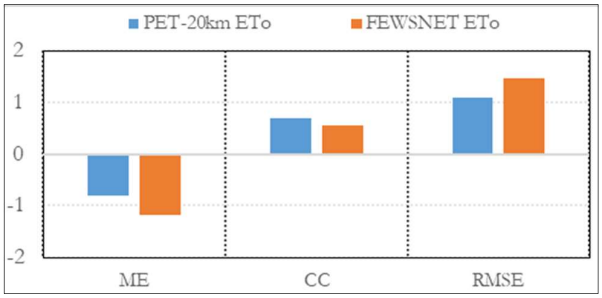


Figure 4-3 Scatter plots for daily PET-20km, FEWSNET PET and Hargreaves PET (left side) and cumulative plot (right side) (2012-2016)

Table 4-2 Evaluation of PET-20km and FEWSNET PET with reference to Hargreaves (2012-2016)

Performance	FAO-56 Hargreaves ETo	PET-20km ETo	FEWSNET ETo
Sum		6665	5983
Diff		1481	2162
ME	8146	-0.81	-1.18
CC		0.70	0.55
RMSE		1.09	1.46



4.3. Performance of satellite rainfall estimates from CMORPH, ARC2, and CHRIPS

The point to pixel comparison of uncorrected SRE products from CMORPH, ARC2, and CHRIPS is shown by scatter plot (Figure 4-4), Taylor diagram (Figure 4-5), accumulated rainfall (Figure 4-6) and mean annual rainfall Figure 4-7.

The performance comparison of all six stations is shown in Figure 4-4. The legends for six-gauge stations is the same as in first station (Imdibir). The three SRE products represent spatial rainfall distribution reasonably well with some differences among this product at a station. CHRIPS shows rainfall depth up-to 60mm for all station while the lower value is shown at the gauge location. There is no substantial difference between ARC2 and CHRIPS at Gubire and Woliso station. However, all products show rainfall estimates far away from 1 to 1 line. Here satellite rainfall detection is visualized from the scatter plots. At ordinate satellite, retrieval shows > 0 but gauge =0 whereas at abscissa rain gauge >0 but satellite rate is 0 and this is happening at 6 stations.

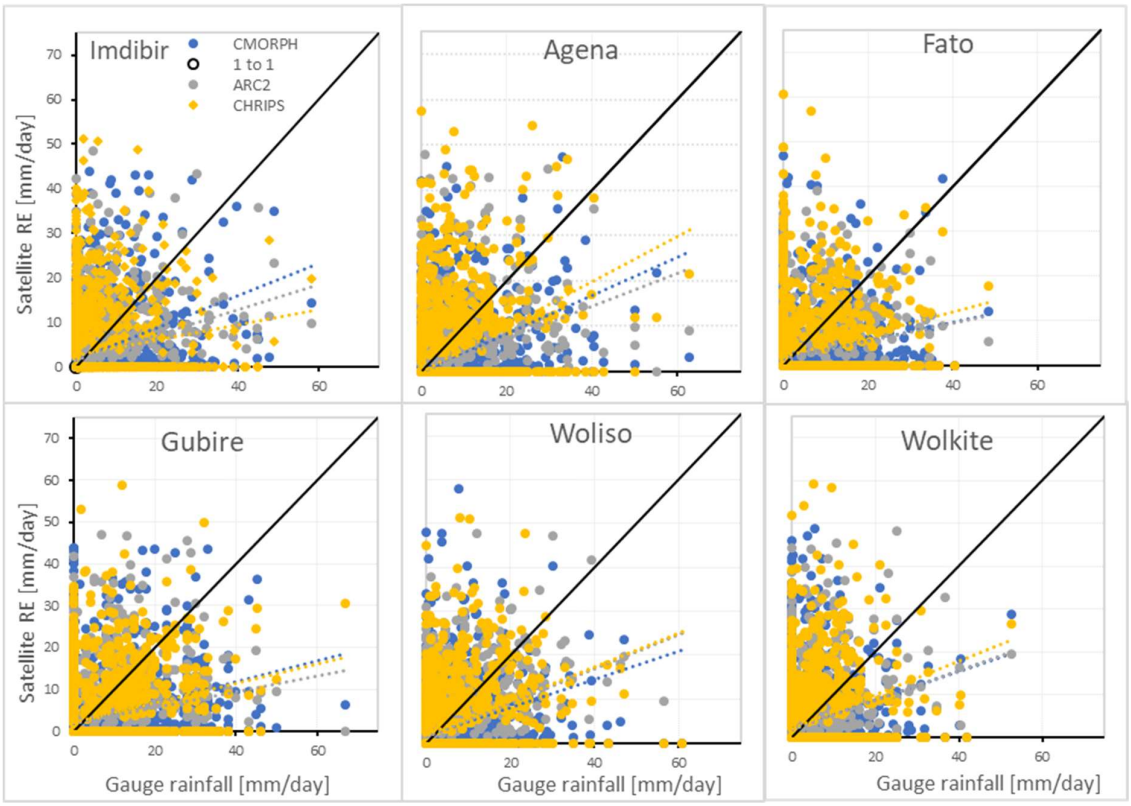


Figure 4-4 Scatter plots at abscissa rain gauge and at ordinate satellite showing the performance of uncorrected satellite rainfall estimates from CMORPH, ARC2, and CHRIPS in Wabe watershed for the 6-gauge station in daily bases (2012-2016). The solid lines are linear fits to the data; the 3 small dotted lines depict each product.

Figure 4-5 shows Taylor diagram showing a concise statistical summary of the performance of raw CMORPH, ARC2 and CHRIPS with gauge rainfall in terms of standard deviation, a coefficient of correlation and root mean square error. In the diagram A =gauge, B=CMORPH, C= ARC2, and D = CHRIPS for the study period 2012-2016. The position of CBD represents how the SRE products match with the reference in-terms of standard deviation (drawn at abscissa and ordinate), correlation coefficient (in the azimuthal line) and RMSE (the cantered curved line). As shown, the correlation coefficient lies between 0.2 to 0.43 indicating weak agreement between SRE products against the reference. ARC2 and CMORPH show a lower correlation at higher elevation but, CHRIPS shows poor CC (Fato station) and good correlation at lower elevation (Woliso station). This result is consistent with (Yang and Luo, 2014) who evaluates performance of CMORPH (in China) found that, SRE are affected by terrain factors. However, they Yang and Luo, (2014) that CMORPH overestimates rainfall depth which contrast with this study. CMORPH and ARC2 have a consistent standard deviation for all gauge locations except at Gubire. The mismatch at Gubire station presumably is due to poor gauging with error of the ground observation data (see section 3.2.1). Although, at all gauge station CHRIPS has the largest distance by standard deviation from the reference. ARC2 has the lowest STDEV followed by CMORPH. To summarize, ARC2 outperforms CMORPH and CHRIPS at higher elevation (Fato) with CC =0.3, STDEV=5.2 and RMSE =7.4 (also see
Appendix 2A).

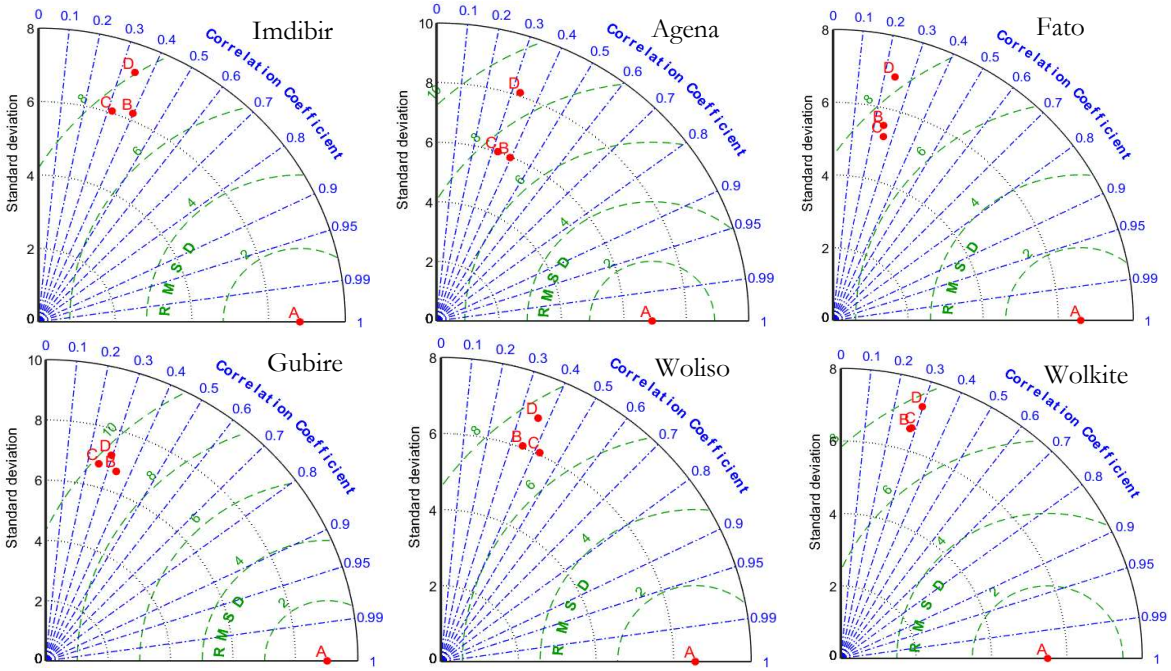


Figure 4-5 Taylor diagram showing a statistical comparison of reference gauge against CMORPH, ARC2 and CHRIPS from (2012-2016)

The Taylor diagram was drawn using the MATLAB code (https://stackoverflow.com/questions/24999338/r-taylor-diagram-plotting). As such, the range for standard deviation (STDEV), correlation coefficient (CC), and root mean square error/difference (RMSE/RMSD) is automatically assigned. The STDEV and RMSE ranges for Imdibir, Fato, Woliso, and Wolkite are 8. While station Gubire and Agena reaches up to 10. Overall, in terms of the accumulated error and CC, CMORPH performs well in all stations except Woliso (medium elevation) and Fato (high elevation). CHRIPS and ARC2 follow.

Figure 4-6 shows the double mass curve accumulated rainfall comparison CMORPH, ARC2, and CHRIPS for Wabe watershed 2012-2016. As shown in Figure 4-6 the uncorrected rainfall accumulation matches with in-situ at gauge location Wolkite (5122mm) with relatively slight overestimation by CHRIPS (5958mm), ARC2(5142mm) and CMORPH (5532mm) located in lower elevation. Probably, the overestimation at Wolkite could be a quality issue of the gauge observation (slightly smaller annual rainfall compared Woliso at nearest location) see 3.2.1 At higher elevation (Fato station) the most underestimation is observed by ARC2 (3713mm) followed by CMORPH (4105mm) and CHRIPS (5059mm). This result is consistent with the study of (Dinku et al., 2018; Funk et al., 2014; Gumindoga et al., 2016) who concludes similar underestimation by SRE's at higher elevation but for different basins. As discussed in the methodology section 3.2.2 there is high mean annual rainfall at Gubire station (see Figure 3-4). That issue is depicted from this comparison for Gubire station. ARC2 and CMORPH show the overall underestimation for stations Imdibir, Agena, Fato, Gubire, and Woliso respectively. CHRIPS shows good match in terms of accumulated rainfall at Imdibir, Agena and Woliso.

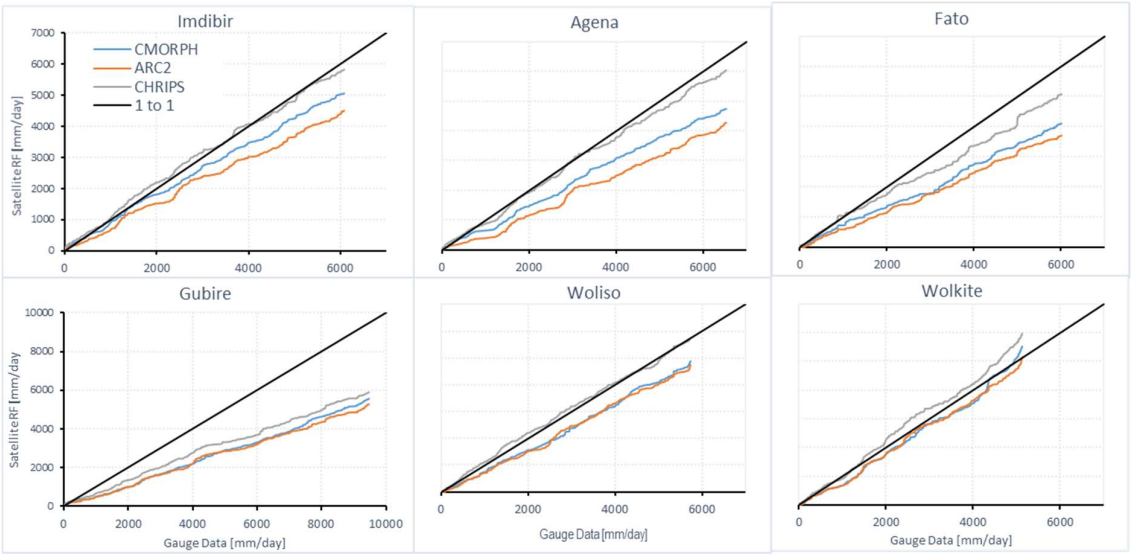


Figure 4-6 Double mass curve showing the accumulated amount of rainfall of the gauge against uncorrected satellite rainfall estimates from CMORPH, ARC2 and CHRIPS in Wabe basin in daily bases (2012-2016)

Figure 4-7 shows the mean annual (2012-2016) rainfall of CMORPH, ARC2, CHRIPS, gauge and the elevation of 6-gauge stations. The result shows that three SRE products underestimate mean annual rainfall at six stations. The most underestimation is seen at the two-gauge stations (Fato and Agena). However, compare to gauge CHRIPS matches annual rainfall depth followed by CMORPH and ARC2.

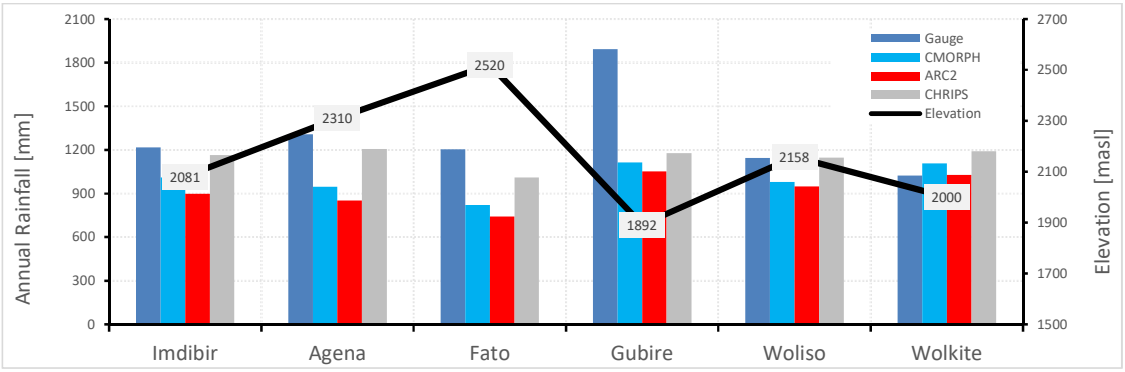


Figure 4-7 Mean annual rainfall of gauge, uncorrected CMORPH, ARC2 and CHRIPS with gauge elevation (2012-2016)

4.4. Evaluation of seasonality effect and bias decomposition

Lumped means the analysis is done for entire time series (2012-2016), wet season means the analysis is done for three most rainy months (June, July, and August) and dry season all other months except wet period.

Figure 4-8 shows the total and average bias of the three satellite rainfall products in the entire time series, wet and dry season (2012-2016). The total and average bias value in Y-axis is multiplied by 100 (as shown '00). The accumulated bias means the sum of hits bias, missed bias, and false bias. For the lumped analysis CHRIPS indicates the highest accumulated bias for all gauge locations which is followed by ARC2 and CMORPH. To understand the satellite products in retrieving rainfall depth in wet and dry seasons the same procedure is applied. In the wet season the total bias in CMORPH increases up-to 2500mm (at Woliso) compared with 1500mm in dry season (at Fato). In average for lumped, wet and dry season the three SRE's follow the same pattern with the lowest bias of CMORPH followed by ARC2 and CHRIPS.

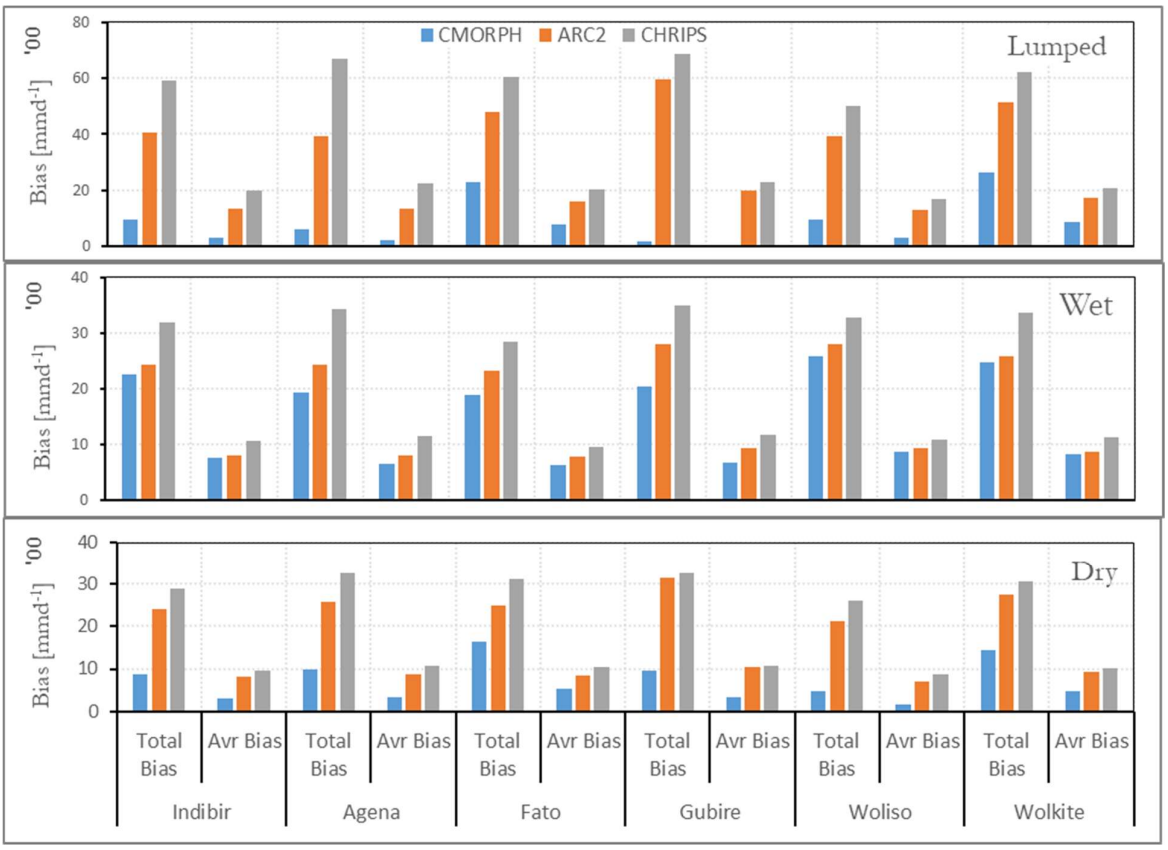


Figure 4-8 Total and average bias of uncorrected CMORPH, ARC2, and CHRIPS satellite rainfall products in Wabe watershed 2012-2016. Top (lumped), middle (wet season) and bottom (dry season), Avr refers to average

The total bias analysis does not give a clear indication for a source of error in satellite rainfall estimates. In other words, there is no clear indication of seasonal variability. Therefore, the total bias is decomposed to three bias as shown below.

Figure 4-9 shows the total bias decomposed to hit bias, missed bias and false bias for daily rain depth. The calculation is based on section 3.7. The top clustered column shows the decomposition considering the entire time series. As shown in this assessment the main source of error for CHRIPS and ARC2 is missed rainfall except at Gubire (red box) showing slight hit bias. On the other hand, there is no one clear source of error shown for CMORPH. The founding reveals that the bias in each of the three estimates depends on the season and elevation. This is consistent with (Romilly and Gebremichael, 2011) who evaluates CMORPH, PERSIAN and TMPA in Ethiopia for different climatic basins.

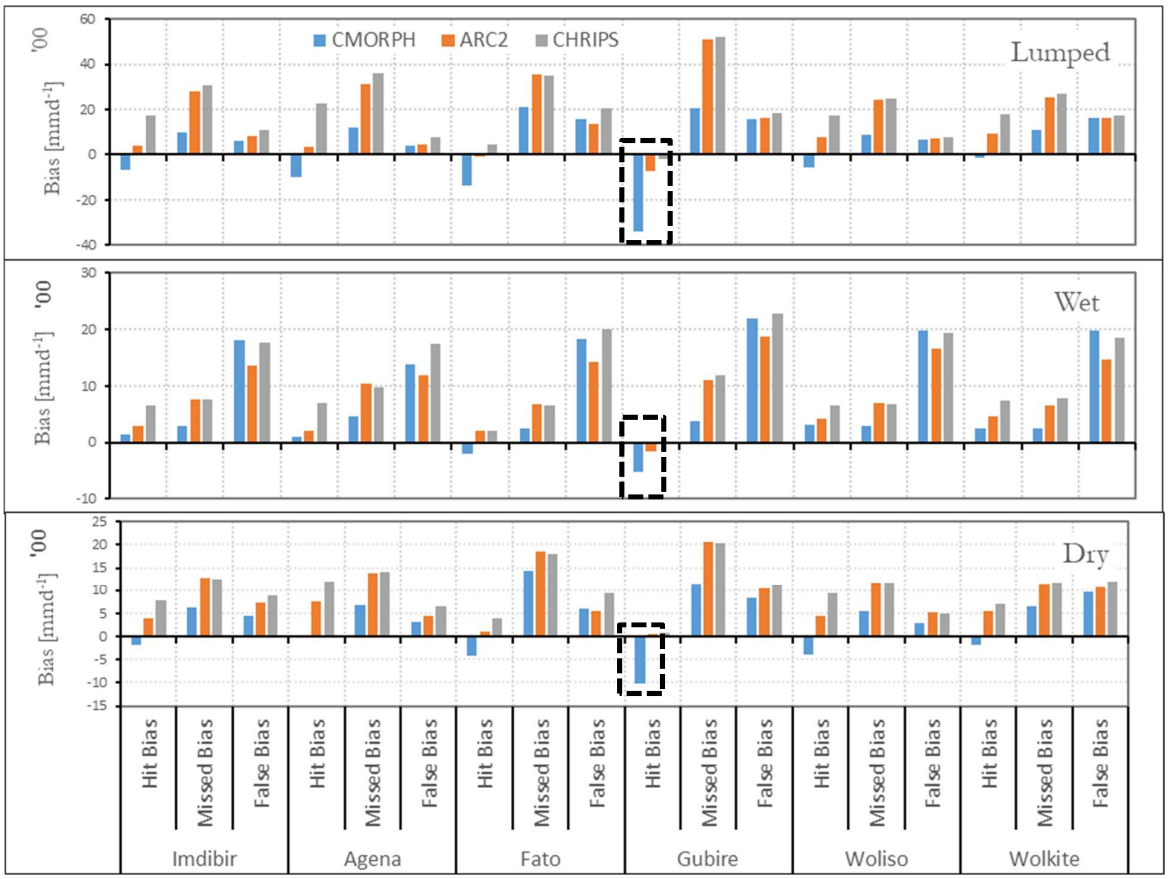


Figure 4-9 Bias decomposition of CMORPH, ARC2 and CHRIPS rainfall estimates at 6-gauge stations for the entire study period on a daily basis (2012-2016)

The clear indication of the influence of seasonality on satellite rainfall estimate is shown from Figure 4-9. False bias predominates during the wet season for all stations whereas missed rain prevails during a dry period. CMORPH has a negative hit bias during the dry period for all six gauges. The issue of hit bias with (out) considering seasonal variation is more visible in highest elevated station Fato.

4.5. Evaluation of detection capability of satellite rainfall estimates

Figure 4-10 shows the categorical measures of selected satellite rainfall estimates from CMORPH, ARC2 and CHRIPS in entire time series (top), wet (middle) and dry (bottom) season against gauge observations (2012-2016). As shown the highest detection capability is seen during wet season compared to dry season by three SRE's. The probability of detection indicates how the product is reliable in retrieving the rainfall depth. As shown in Figure 4-8 for the lumped analysis the fraction of rain events which were correctly captured by CMORPH is within a range of 50 (Fato) to 78% (Woliso). ARC2 and CHRIPS have a consistent pattern of POD with a range of 30 (Fato) to 42% (Woliso) with a slight increase by CHRIPS. Note that Fato is located in highest elevation whereas Woliso is in the lowest elevation. This result conforms with the study of Gumindoga et al., (2016) who evaluate performance of CMORPH in Zambezi river basin. Also, it is consistent with Ward et al, (2011) who evaluates the performance of TRMM and PERSIAN in complex mountainous terrain of Baker river basin in Argentina Patagonia. And also conforms with Scheel et al, (2011) who assesses the performance of TRMM in the central Andes region. This three authors concludes satellite rainfall estimates underestimate in the mountainous region and slight match to overestimation at lower elevation.

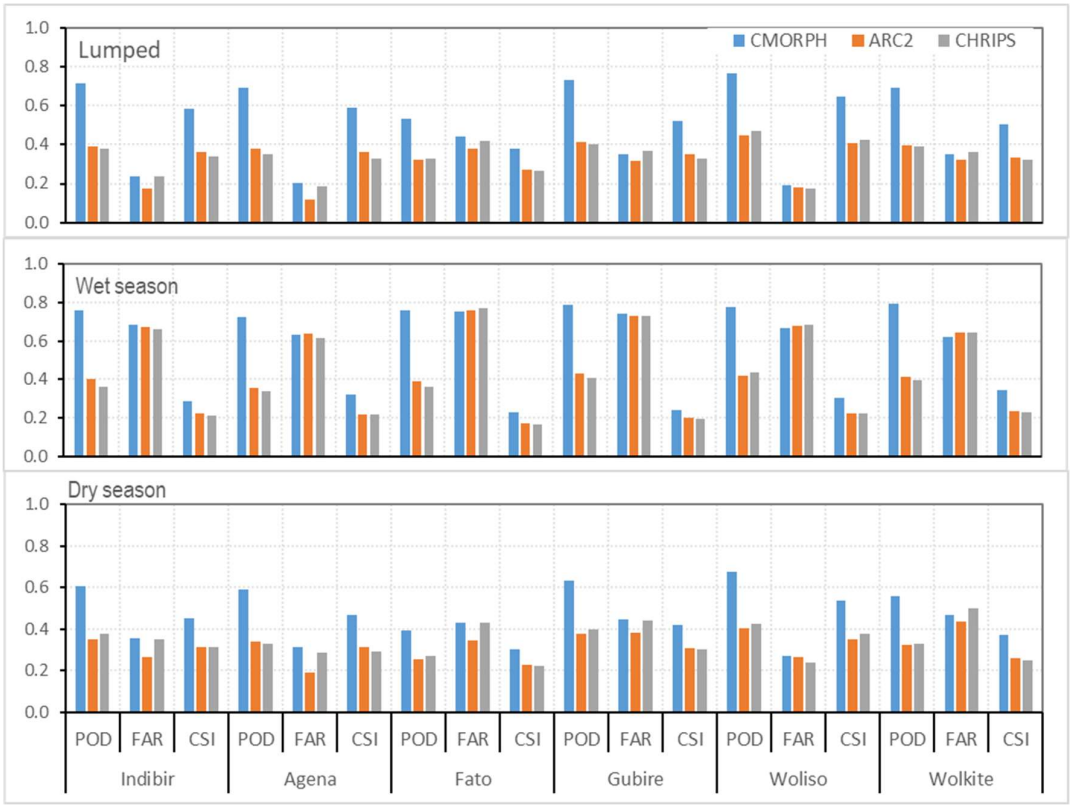


Figure 4-10 Detection skill score of CMORPH, ARC2 and CHRIPS rainfall estimates at 6 gauge stations for lumped (top), wet season (middle) and dry season) on a daily bases (2012-2016)

However, the POD of all satellite performs well during the wet season. The falsely detected rainfall in wet and dry season varies between 63-79% (Wolkite-Fato) and 18-45% (Agena-Wolkite) respectively. This result corresponds to the seasonal dependency of satellite error source by Ward et al., (2011) and Scheel, (2011). CMORPH detects the fraction of correctly identified rainfall in both seasons.

In summary, a POD of CMORPH performs with (out) season variation in all station, FAR of CHRIPS be more than in all station except Woliso (three of them are similar FAR%) and highest performance of CMORPH in-terms of CSI was seen in Wabe watershed. Compared to ARC2 and CHRIPS, CMORPH performance well in-terms of CSI and POD presumably, due to its high spatial (8km) and temporal resolution of 30 minute. This, results with reasonable detection score and fraction of correctly identified rainfall depth.

In summary, the uncorrected SRE's from CMORPH and ARC2 underestimates mean, maximum and accumulated rainfall at the five-gauge location. CHRIPS underestimates mean and accumulated rainfall at five locations but detects the maximum rainfall slightly higher than Insitu at Agena, Fato and Woliso station. However, the three SRE's overestimates maximum and mean rainfall at Wolkite station (see Appendix 2A . The reason for overestimation at Wolkite station could be quality of gauge measured data for this station since the mean annual gauge rainfall is smaller compared to other stations (see section 3.2.1 and Figure 3-4).

4.6. Evaluation of satellite rainfall bias correction

Figure 4-11 shows the evaluation statistics of RMSE, CV and CC for gauge, uncorrected and bias corrected SRE products (2012-2016). Appendix 1A shows accumulated sum, mean annual average, ratio, and the standard deviation. The accumulated error in uncorrected ARC2 ranges between 6.8 to 7.6 [mm/day], CMORPH ranges between 7.2 to 7.6 [mm/day] and in CHRIPS ranges between 7.6 to 8.8 [mm/day] besides Gubire station. Initially, uncorrected estimates from ARC2, CMORPH, and CHRIPS respectively perform well in-terms of RMSE. Afterward, from four applied bias correction scheme distribution transformation shows outstanding performance in successfully reducing the accumulated error for MORPH, ARC2, and CHRIPS at all gauge station.

This result conforms with (Fang et al, 2015; Gumindoga et al., 2016; Teng et al., 2015) who concludes the distribution transformation based bias correction scheme reduces the accumulated error. As shown in Figure 4-11 (A) the introduction of accumulated error after spatio-temporal bias scheme particularly with time and space variable is observed. As discussed in section 3.9 subsection 3.9.2.1 there is no clear common rule for the selection of time window, accumulated rain depth and number of rainy days to determine the bias factor. Consequently, this bias scheme suppresses the accumulated error compared to the uncorrected estimate and SDT method.
Appendix 2A shows the overall statistical findings of bias correction scheme. As noted, from the table there is no clear pattern of bias schemes in correcting the mean value of satellite rainfall estimates. The spatio-temporal bias scheme (TVSV, TVSF, and TFSF) shows poor in matching the standard deviation with gauge observation as well as uncorrected products. TVSV shows well performance in matching the accumulated rainfall. There is an additional introduction of error by TVSF bias correction scheme.

As shown in Figure 4-11(B) distribution transformation bias scheme shows well performance with a coefficient of variation range between 1.8 to 2.0 with respect to gauge (ranges between 1.7 to 2.0). On the other hand, the spatio-temporal bias correction schemes lie far away from the gauge in-terms of a coefficient of variation for instance, TVSV ranges 2.4 to 2.7 at the lowest and highest elevation. The CV of TFSF, TVSF, and TVSV respectively shows reasonable performance next to SDT. In summation, at high elevated gauge location of Fato SDT reasonably performs whereas the three bias schemes purely match with gauge.

Additionally, Figure 4-11 (C) shows the CC of gauge, uncorrected and corrected SRE's products. The uncorrected satellite rainfall estimates CC ranges between 0.2 to 0.4. Both CMORPH and CHRIPS have low correlation coefficient of 0.2 at Fato station compared to ARC2 having a CC of 0.3. As shown in the same figure SDT effectively improves a positive relationship with gauge up to 0.64. Next to SDT, TVSF bias scheme has an improved value of CC.

The distribution transformation which reasonably performs in all stations in-terms of CC and RMSE diminishes its performance at Fato station with an increased accumulated error up to ~20% and decreased correlation by ~20%. The maximum daily rainfall depth in the watershed are underestimated by both uncorrected CMORPH and ARC2 in all stations. On the other hand, the spatio-temporal bias correction scheme suppresses the maximum rainfall depth with reference to gauge measurement (at Fato gauge=48.5, TVSV = 99.3, TVSF = 121.3) (see Appendix 2A).

Further assessment of the accumulated difference of rainfall depth from gauge, uncorrected and bias corrected satellite rainfall for entire time series, wet and dry season (2012-2016) is shown in Appendix 2B.

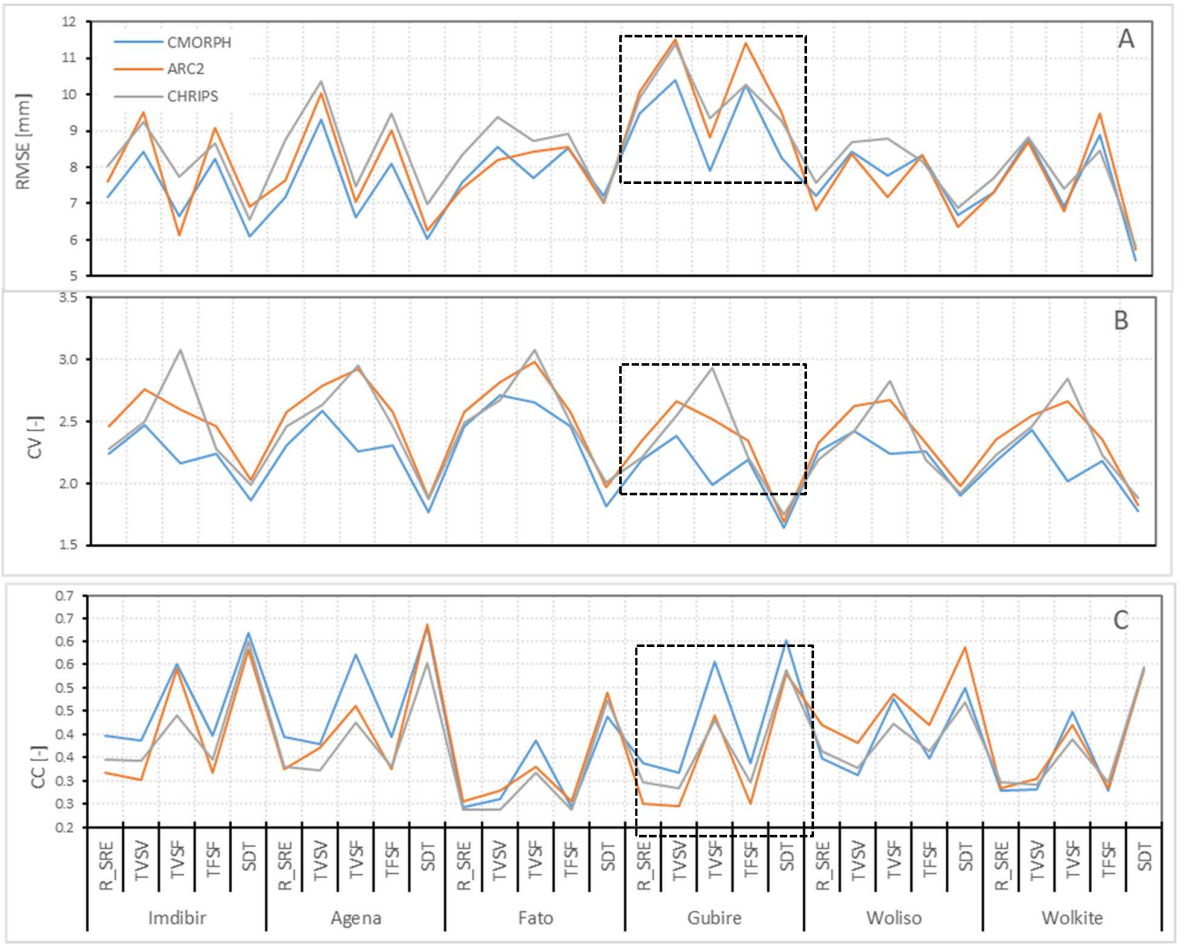


Figure 4-11 Evaluation statistics for uncorrected and bias corrected CMORPH, ARC2 and CHRIPS. RMSE (A), CV (B) and CC (C)

Figure 4-12 shows the bias and relative bias of uncorrected and four bias correction schemes for Wabe watershed. The calculation of bias and relative bias is based on performance analysis section 3.6. The result shows at Gubire station the largest negative bias and relative bias. This is due to the quality issue in the gauge observation. As shown, the most bias is shown by uncorrected ARC2 (-1.3 at Imdibir and Fato station -1.2) and CMORPH 1.0 in the same station. Except for the high elevated station of Fato TVSV shows decreased bias in all station with slight under and overestimation. The distribution transformation bias scheme performs next to the TVSV scheme. SDT has not underestimated in all gauge location for three satellite rainfall retrievals. Except for Wolkite station the relative bias of uncorrected SRE ranges from -0.1 to -0.3. Although, TVSV and SDT show similar pattern in-terms of relative bias.

To summarize, TVSF has the most negative bias and relative bias like uncorrected SRE's. Except for DT all other has a negative bias and relative bias at high elevated Fato station. There is no consistent pattern with the TFSF bias correction scheme. TVSV shows lowest bias (~0.1) followed by DT (0.15).

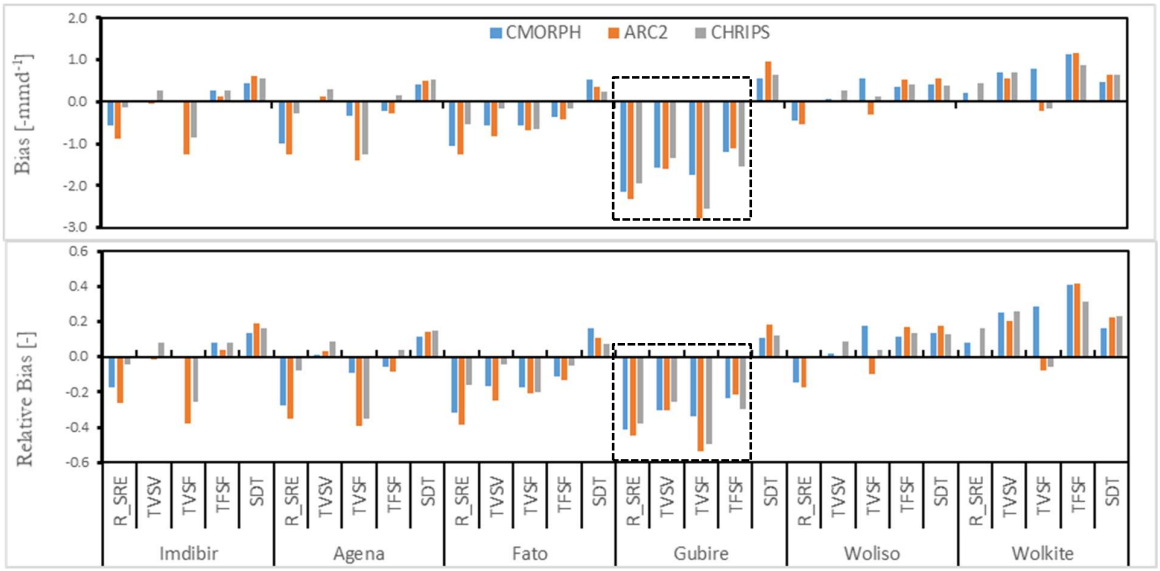


Figure 4-12 Measures of systematic differences bias (top) and relative bias (bottom) for uncorrected, bias corrected and gauge rainfall (2012-2016)

Error! Not a valid bookmark self-reference. shows the percentage of days belonging to six rainfall intensities for Wabe watershed for gauge, uncorrected and four bias corrected satellite rainfall estimates.

SRE		CMORPH						ARC2						CHRIPS					
RF-Intensity		0-1	1_2.5	2.5_5	5_10	10_20	>25	0-1	1_2.5	2.5_5	5_10	10_20	>25	0-1	1_2.5	2.5_5	5_10	10_20	>25
Imdibir	Raw	*68*	8.8	6.6	7.7	5.9	3.3	*79*	0.6	3.1	7.9	6.6	3.0	*78*	0.1	1.9	6.8	9.1	*4.4*
	Gauge	**64**	**6.7**	**9.0**	**8.5**	**7.5**	**4.3**	**64**	**6.7**	**9.0**	**8.5**	**7.5**	**4.3**	**64**	**6.7**	**9.0**	**8.5**	**7.5**	**4.3**
	BF_{TVSV}	67	7.9	6.8	7.2	6.1	*4.5*	79	0.4	3.0	6.6	6.3	4.9	78	0.1	2.1	6.2	8.6	5.4
	BF_{TVSF}	65	5.4	**7.0**	10.8	8.4	3.2	80	1.6	3.0	6.7	6.5	1.9	81	1.6	2.6	6.0	5.4	3.6
	BF_{TFSF}	66	**7.2**	6.6	7.9	6.9	5.0	79	0.1	1.8	5.7	7.9	5.7	78	0.1	1.0	5.5	9.8	6.0
	S_{DT}	55	10.7	10.8	11.5	8.2	3.7	61	*6.4*	8.8	9.7	8.5	5.1	62	*5.5*	7.8	10.3	9.5	4.5
Agena	Raw	*70*	7.4	7.4	6.8	5.7	3.2	*80*	0.6	3.2	7.8	6.1	2.7	*80*	0.1	1.3	5.5	8.2	5.5
	Gauge	**60**	**8.9**	**8.0**	**10.6**	**8.4**	**4.4**	**60**	**8.9**	**8.0**	**10.6**	**8.4**	**4.4**	**60**	**8.9**	**8.0**	**10.6**	**8.4**	**4.4**
	BF_{TVSV}	69	6.9	6.4	6.4	6.0	5.5	80	0.4	2.4	5.0	6.3	6.2	80	0.1	1.4	5.5	6.3	7.2
	BF_{TVSF}	68	5.9	6.1	9.7	6.6	*4.2*	81	1.4	2.7	6.6	5.5	2.5	83	1.1	2.0	5.3	5.7	3.3
	BF_{TFSF}	68	6.8	6.6	7.6	6.2	4.9	80	0.2	1.4	5.9	7.6	5.3	80	0.0	0.8	5.0	8.0	6.6
	S_{DT}	54	*10.0*	11.7	10.8	10.1	3.7	60	*6.7*	7.8	11.4	9.4	4.9	61	*5.0*	7.3	11.3	10.6	*4.4*
Fato	Raw	*74*	6.2	6.1	6.1	5.5	2.4	*81*	0.8	2.8	6.8	6.1	2.2	*80*	0.1	2.9	5.7	7.4	4.1
	Gauge	**70**	**3.8**	**4.7**	**7.1**	**10.9**	**3.9**	**70**	**3.8**	**4.7**	**7.1**	**10.9**	**3.9**	**70**	**3.8**	**4.7**	**7.1**	**10.9**	**3.9**
	BF_{TVSV}	73	6.0	5.8	6.1	5.1	*3.8*	81	0.7	2.4	6.1	5.8	3.6	80	0.1	2.6	4.9	7.9	4.8
	BF_{TVSF}	72	*4.9*	5.6	8.5	5.7	3.3	82	1.1	2.0	4.3	6.2	3.9	82	1.2	2.3	5.3	4.6	4.3
	BF_{TFSF}	73	5.0	6.1	5.7	6.3	4.2	81	0.4	1.5	4.3	7.9	4.5	80	0.1	1.9	5.5	7.5	5.4
	S_{DT}	59	7.3	9.3	10.9	9.1	4.5	66	*4.4*	6.7	8.2	9.4	4.9	68	*4.3*	5.6	9.2	8.5	*4.8*
Gubire	Raw	*67*	7.9	6.6	8.7	6.6	3.6	*77*	0.9	2.6	7.2	8.1	3.7	*76*	0.2	2.2	8.0	8.4	4.9
	Gauge	**64**	**1.8**	**2.9**	**8.0**	**14.6**	**8.3**	**64**	**1.8**	**2.9**	**8.0**	**14.6**	**8.3**	**64**	**1.8**	**2.9**	**8.0**	**14.6**	**8.3**
	BF_{TVSV}	66	7.2	6.8	8.0	6.7	4.9	77	0.8	2.7	6.3	7.4	5.4	76	0.2	2.2	7.3	8.0	6.0
	BF_{TVSF}	65	*5.1*	7.4	10.5	8.0	4.3	79	*1.6*	3.5	6.2	6.8	2.8	80	1.5	3.3	6.7	5.2	3.7
	BF_{TFSF}	65	6.8	6.7	8.4	6.9	6.1	77	0.3	1.4	4.8	8.7	7.3	76	0.2	1.5	7.0	8.8	*6.3*
	S_{DT}	51	7.0	8.4	12.3	12.9	*8.1*	57	3.9	5.8	10.0	12.8	10.6	59	*4.1*	5.4	9.6	11.6	10.5
Woliso	Raw	*69*	*7.0*	7.0	7.4	6.7	2.8	*78*	0.9	3.3	7.3	7.6	3.3	*77*	0.1	2.0	7.4	10.0	3.9
	Gauge	**65**	**7.1**	**7.2**	**9.8**	**7.0**	**3.7**	**65**	**7.1**	**7.2**	**9.8**	**7.0**	**3.7**	**65**	**7.1**	**7.2**	**9.8**	**7.0**	**3.7**
	BF_{TVSV}	69	6.6	6.2	7.4	6.6	4.2	78	0.8	2.7	6.6	7.4	4.7	77	0.1	2.3	7.7	8.5	4.8
	BF_{TVSF}	67	5.7	5.3	8.9	7.8	5.1	80	1.7	2.3	5.1	7.0	4.3	79	1.5	2.7	5.8	6.4	4.8
	BF_{TFSF}	68	5.9	6.5	7.3	7.9	4.6	78	0.3	1.8	5.5	8.2	6.6	77	0.0	1.4	6.5	10.0	5.5
	S_{DT}	58	9.7	8.5	12.9	7.7	*3.4*	63	*6.0*	7.6	10.6	8.8	4.5	62	*6.1*	7.4	11.2	9.3	*3.7*
Wolkite	Raw	*67*	7.6	6.6	8.2	7.0	3.7	*77*	0.7	3.6	7.1	8.0	3.4	*76*	0.2	2.0	8.4	9.1	4.4
	Gauge	**66**	**3.7**	**9.6**	**11.0**	**8.0**	**2.1**	**66**	**3.7**	**9.6**	**11.0**	**8.0**	**2.1**	**66**	**3.7**	**9.6**	**11.0**	**8.0**	**2.1**
	BF_{TVSV}	**67**	7.2	7.1	7.2	6.8	4.8	77	0.7	3.2	6.6	7.3	5.1	76	0.3	2.2	7.8	9.0	4.8
	BF_{TVSF}	64	6.1	6.6	10.3	8.6	4.4	79	1.5	3.4	6.6	6.6	3.2	79	1.3	2.9	7.0	6.0	3.4
	BF_{TFSF}	66	*6.0*	6.6	7.8	7.9	5.7	77	0.5	1.6	5.4	8.4	7.1	76	0.0	1.4	7.1	9.7	5.8
	S_{DT}	55	10.2	12.5	12.0	7.5	*2.4*	59	*5.5*	11.4	12.5	8.6	2.7	59	*7.8*	8.6	12.9	8.4	*3.0*

The uncorrected and corrected pixel rain rates from three satellite products are compared in terms of capturing the percentage of rain rate with reference to six-gauge rainfall station. The uncorrected SRE from CMORPH and ARC2 overestimates the dominant light showers and underestimating (except at Wolkite station) the maximum rain rates. Perhaps, this underestimation of the dominant light showers at

Wolkite station could be consistent with the lowest observed rainfall from the gauge. The CHRIPS product overestimates the maximum rain rates (>25mm) in a five-gauge location except for Gubire. The dominant rain rate (<1mm) from ARC2 and CHRIPS is effectively captured with distribution transformation bias scheme with reference to gauge measurement. This is consistent with the study of (Gumindoga et al., 2016) who use five rainfall intensity classes with light showers <2.5mm and he concludes SDT is the best bias correction scheme in the Zambezi basin.

Figure 4-13 shows percentage of rain rates based on gauge measurements. Based on gauge observations the average percentage of rain rates in the watershed are 65, 6, 8, 9, 8 and 4 for 0-1, 1-2.5, 2.5-5, 5-10, 10-20 and >25 respectively. This evidence, the dominant rainfall in the watershed is light showers (<1mm) and the lowest is heavy rainfall (>25mm). The rainfall intensity analysis result for Gubire gauge station (bold) shows differences compared to 5 other gauge stations even it is nearer to Wolkite station. Particularly, Gubire station showing lower and the highest percentage of measured rainfall of 1 to 5mm and 5 to >25mm respectively. This probably is due to the quality of ground observation (see section 3.2.1).

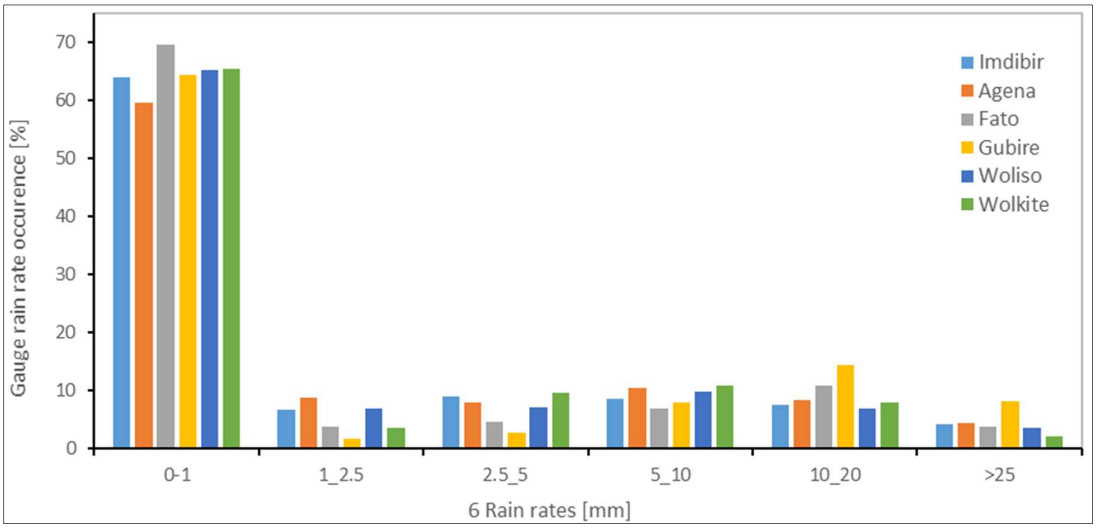


Figure 4-13 Percentage of rain rates in the study area based on gauge measurement

All in all, the bias correction result in Wabe basin shows the performance of the selected scheme varies with respect to evaluation metrics. Also, the result reveals that there is no perfect bias correction scheme which matches the gauge observation. However, there is some definite improvement of correcting satellite rainfall estimate. In this manner, the distribution transformation bias scheme outstands the spatio-temporal in-terms of the reasonable match between daily maximum rainfall depth, standard deviation, a coefficient of variation, a coefficient of correlation and root mean square error. The reasonably successful performance of SDT bias scheme is shown using Taylor diagram (see Appendix **2C**).

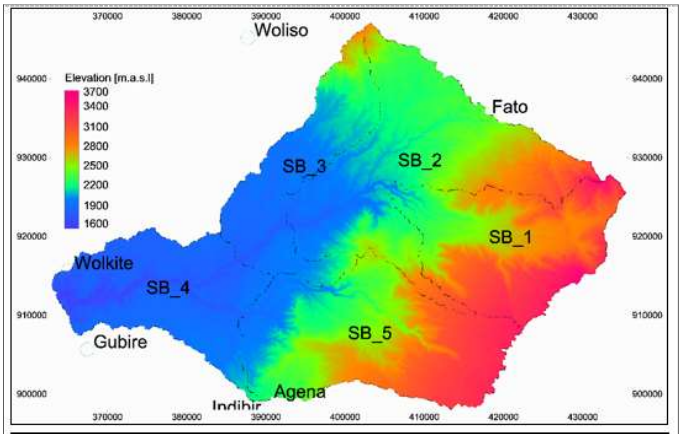
In terms of matching the mean, ratio and accumulated rainfall the time and space-variant spatio-temporal bias scheme shows good performance. From this study, the two first and second bias correction schemes are SDT and TVSV respectively followed by TVSF and TFSF. Therefore, SDT corrected satellite rainfall estimate are used for the streamflow simulation and water balance assessment.

Appendix **2C**shows double mass curve of the accumulated amount of rainfall of the gauge against distribution transformation bias scheme in Wabe basin in daily base 2012-2016. The consistency of corrected satellite rainfall estimates follows with slightly increased rainfall depth by SDT.

4.7. HBV-Light rainfall runoff modelling result

4.7.1. Elevation influence on HBV Light model simulation

Figure 4-14 table (middle) and clustered column(bottom) and Figure 4-15 depicts the inventory result for change in precipitation with elevation for the period June 2014 and 2015. In the figure (top) location and elevation variation for met-gauged station is shown (also see Table 2-1



Station/Year		Indibir	Agena	Fato	Woliso	Wolkite
2014	min	0.0	0.0	0.0	0.0	0.0
	Max	36.5	40.4	34.3	60.9	22.7
	Avr	8.8	10.5	9.7	8.5	8.7
	STDEV	9.8	12.5	9.2	13.3	5.8
2015	min	0.0	0.0	0.0	0.0	0.0
	Max	39.2	29.7	30.0	32.9	23.0
	Avr	8.0	5.7	7.6	8.7	5.6
	STDEV	10.7	7.7	9.7	8.6	5.6

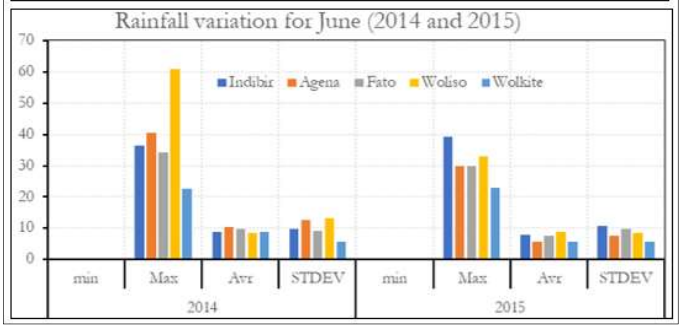


Figure 4-14 Inventory of rainfall variation with elevation

Table 2-1 Summary meteorological data collected from NMAE during fieldwork). The selection of month June and Year 2014, 2015 out of study time window (2012-2016) is due to rainy month, good quality and non-missed gauge observation. In mountainous areas elevation significantly affects precipitation rate. The result from both gauge Figure 4-14 and satellite rainfall products Figure 4-15, it is evident that there is no clear consistent increase in precipitation with elevation in the Wabe watershed. For instance, gauge inventory in 2014 reveals good indication with respect to average rainfall in high elevated Fato station but this is not the case with maximum rainfall and average rainfall in both years. The same analysis from ARC2 and CHRIPS signals shows there is no significant consistent pattern of rainfall increase with elevation in the basin. Similar sort of analysis but for longer window and different rainfall spatial scale were made in China by Xu, (2018) and Chen, (2015) shows there is no significant change of precipitation with elevation. Presumably, small temporal and spatial scale is tested was not shown the lapse rate pattern. As stated by (Seibert and Vis, 2012) but for different catchment and preliminary analysis from this study catchment setting was adopted.

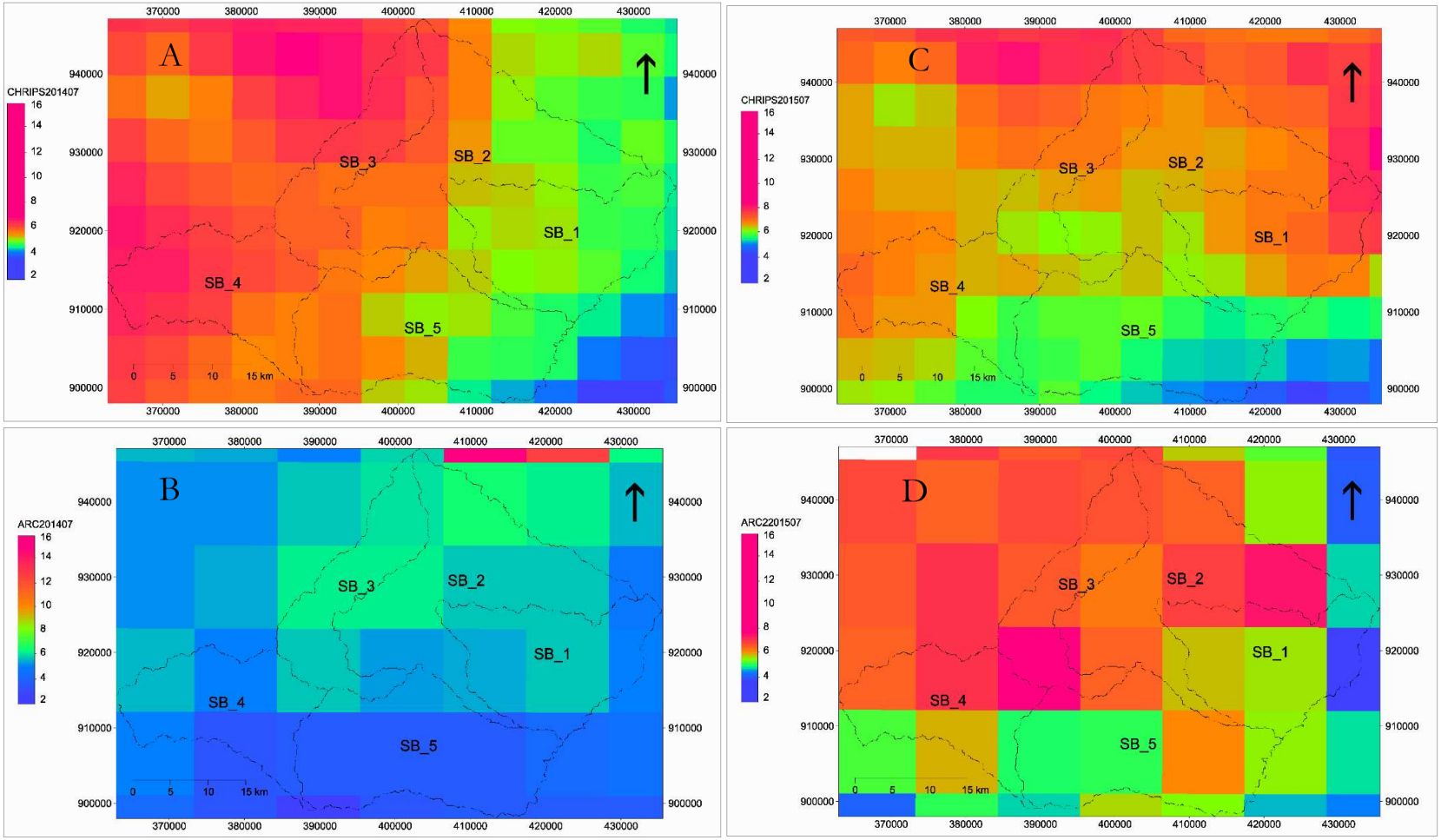


Figure 4-15 Analysis of change of precipitation with the lapse rate. Where: A and C (CHRIPS 2014 and 2015) and B and D (ARC2 2014 and 2015).
The lapse rate analysis could not reveal reasonable result. This is because 1) there are few stations with narrow elevation range (500m), 2) an estimation at high elevation is not available as an indication how rain depth is affected by elevation, and 3) weak performance of SRE's products in mountainous area may not show lapse rate pattern.

4.8. Model sensitivity analysis, calibration and validation result

4.8.1. Model initialization and sensitivity analysis

Table 4-3 Shows parameter values used in the initialization of the model (2007-2008). Virtually, the model well-simulated baseflow at the end of wet season while there is mismatch of baseflow at onset of wet period. The timing of simulated peak flows, rising and the model well captures recession limbs.

Table 4-3 Parameter value used during initialization of the model

Para meter	FC [mm]	LP [-]	BETA [-]	PERC [mm Δt^{-1}]	UZL [mm]	K0 [1/Δt]	K1 [1/Δt]	K2 [1/Δt]	MAXBAS [Δt]
Value	(145,135,45)	(0.75,0.65,0.55)	(0.6,0.35,0.45)	0.45	65	0.18	0.00087	0.0067	1.23

Figure 4-18 shows the sensitivity analysis of the model to changes in nine catchment and vegetation parameters. The effect of each parameter in simulated streamflow is virtually visualized and evaluated with NSE and RVE. The result shows that the most sensitive parameters in Wabe watershed are LP, PERC, BETA and K0 are four most sensitive parameters followed by FC and K1. The remaining three parameters those controls groundwater contribution (K2), the threshold parameter (UZL) and the transformation weighting parameter (MAXBAS) are less sensitive but has their effect in the simulated streamflow. This result is consistent with Mohamed et al., (2010) who applied HBV in Gilgel Abay Ethiopia, and revealed that the most sensitive parameters are FC, PERC, BETA, K1 and K2. Abebe et al., (2010b) assess a parameter sensitivity of HBV Light in United States catchment and found that MAXBAS, PERC, FC, BETA, LP and K1 are sensitive parameters. Furthermore, Lidén and Harlin, (2000) evaluated HBV-96 model in four different climate zones (Turkey, Zimbabwe, Tanzania and Bolivia) and found out that MAXBAS, K0 and K1 are sensitive while FC and PERC are insensitive.

Discussion on the effect of model parameters in simulated streamflow

FC: The storage of soil zone (FC) mainly affects actual evapotranspiration and amount of water in soil box. Decreasing FC from the optimized value sharply decreases the amount of water in soil box and actual evapotranspiration by increasing simulated streamflow with RVE =3. While, increasing FC decreases the simulated streamflow by RVE= -8.4. The decrease in streamflow by increased FC is due to its inverse relation to the recharge with UGB.

LP: The soil moisture value above which ETa reaches PET parameter affects the simulated streamflow as shown in Figure 4-18. The smaller value of LP simulates too little streamflow with RVE = -26.3. Also, it is observed that ETa and amount of water in soil box increases during the wet season following rainfall pattern, and during peak rainfall events ETa reaches PET. Increase in LP decreases model simulated ETa.

BETA: A parameter that determines the relative contribution of runoff from rain. Lower and larger BETA value increases and decreases simulated streamflow respectively. For instance, BETA = 0.1 reduces AET, an abrupt change in rising limbs and peak flows, and less effect on LGB. The decrease in simulated streamflow by an increase in BETA indicates soil moisture is less than the maximum water holding capacity FC.

PERC: Percolation rate occurring when water is available. It mainly affects the base flow, upper groundwater box (UGB) and lower groundwater box (LGB). The effect of PERC is not captured by the performance analysis of RVE and NSE. For instance, at PERC = 0.1 and 0.6 RVE is (-2.65, -2.54) whereas NSE is (0.7, 0.71). Results of further analysis of the effect of PERC on baseflow, UGB and LGB is shown in Figure 4-17 and Figure 4-16 respectively. Lowering PERC increases storage in UGB by reducing in LGB and vice versa. This is because PERC controls flow from UGB to LGB. In summary, increased PERC, results in an increase in delayed runoff and decreased peak flows.

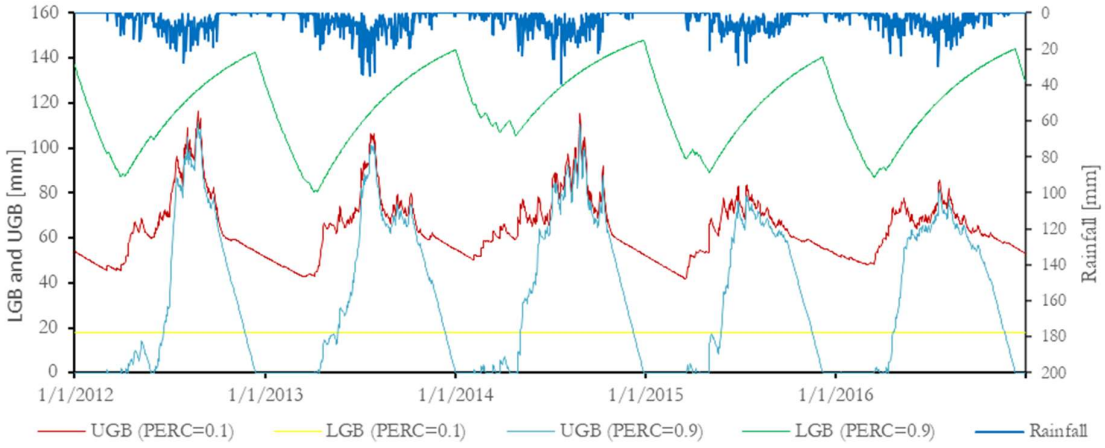


Figure 4-16 Effect of PERC parameter on upper and lower groundwater box

Figure 4-17 shows the effect of PERC in simulated baseflow. Ten PERC values are selected between 0.1 and 6. There is no best fit with observed streamflow for selected value range. The PERC value above 0.4 increases the simulated baseflow while the value below 0.4 deviates the baseflow with reference to observed streamflow. This is shown by the zoomed part of Figure 4-17 (B).

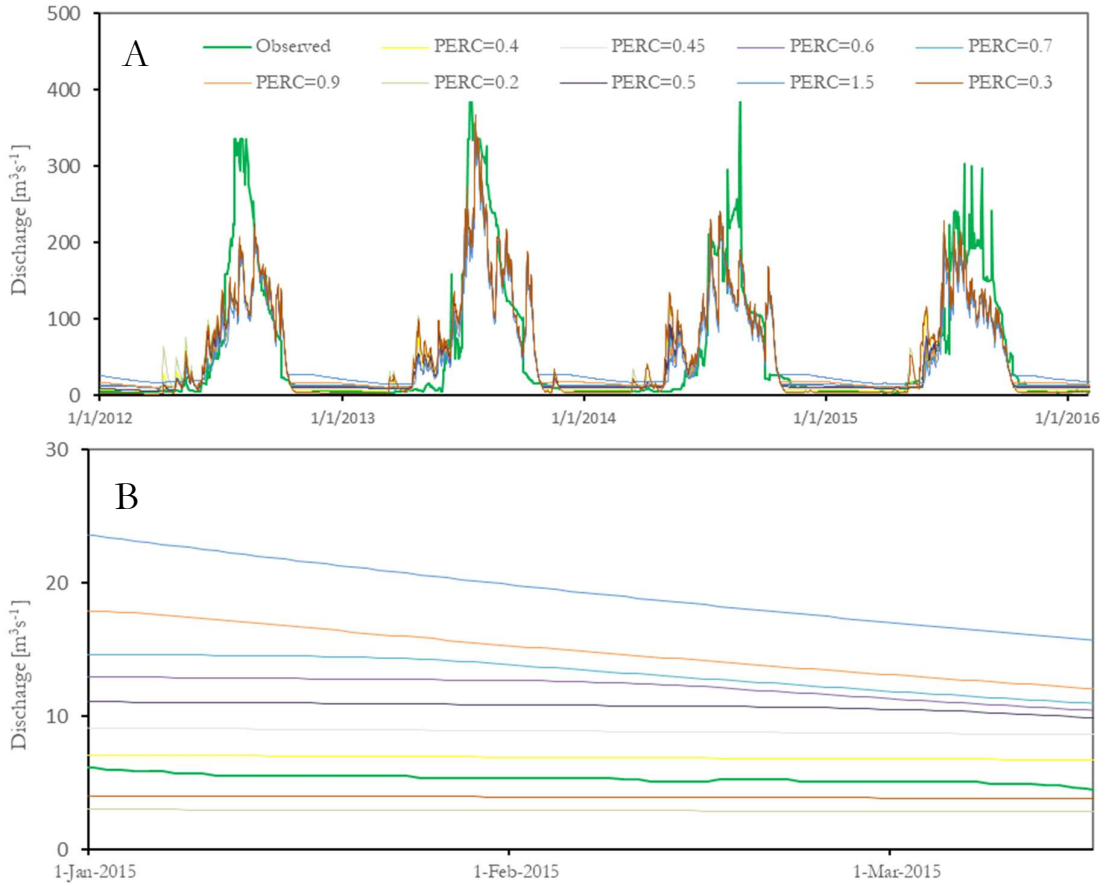


Figure 4-17 Effect of PERC parameter on simulated baseflow. A (2012-2016) and B (2015, Jan-Mar)

In the HBV Light version model, PERC value ranges from 0 to infinity (Seibert and Vis, 2012), while it was 0.1 to 2.5 in previous HBV version (Deckers et al., 2010). In this study area PERC,> than 1.5 does not show the hydrograph pattern of simulated streamflow.

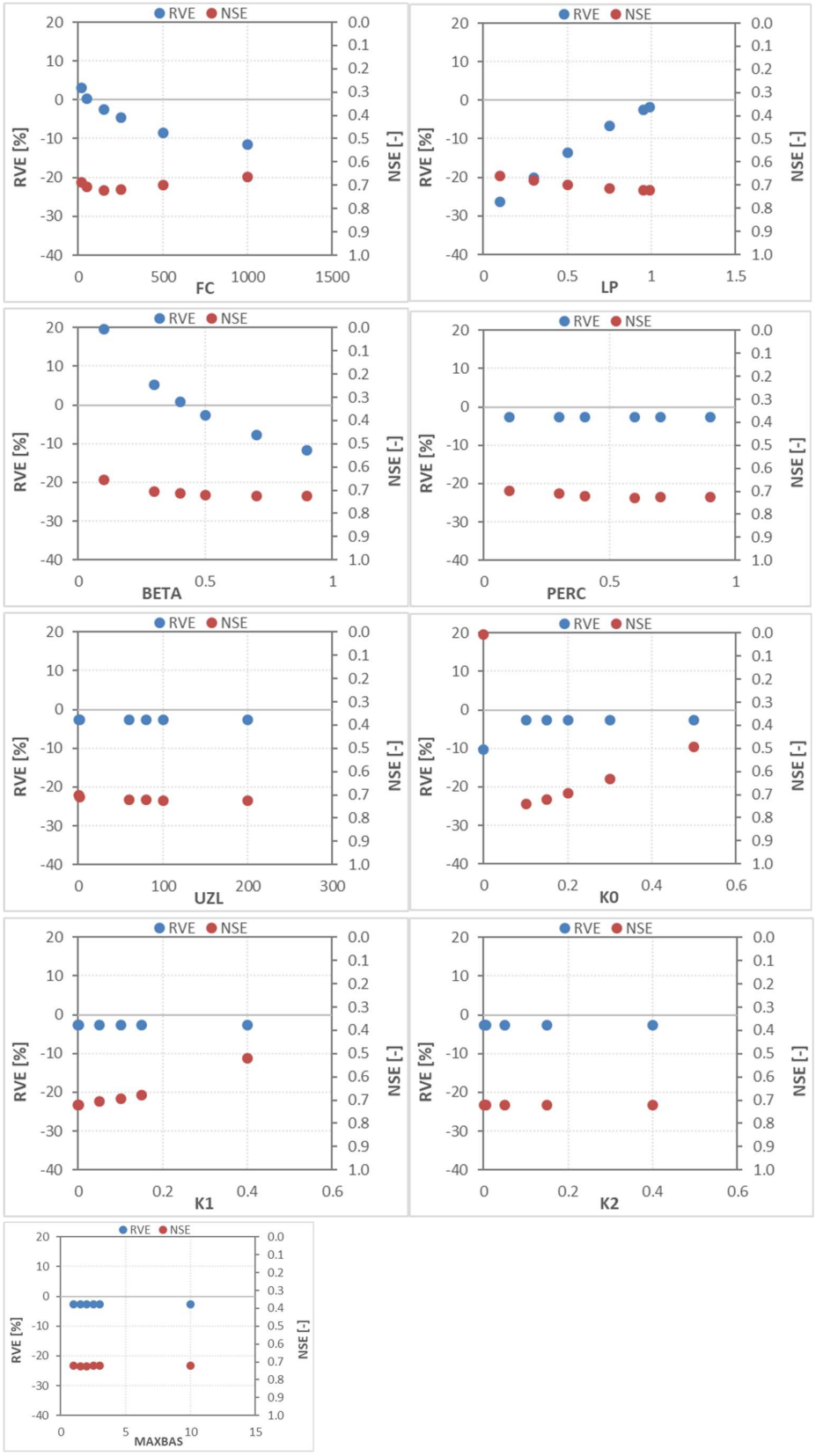


Figure 4-18 Sensitivity of parameters and their effect on model performance

UZL: Varying UZL between 5, 60 and 100 significantly affects the upper and lower groundwater box ( ). Uniform storage is seen for LGB for the dry and wet season at UZL = 60 and 100. However, there is an abrupt change in UGB at UZL =60 and 100. For instance, at UZL =60 and 100, the UGB matches for the beginning dry months (Jan, Feb and March of 2012 up to the beginning rainy month (June). Afterward, UGB increases in entire time series (2013-2016) for both dry and wet seasons at UZL =100.

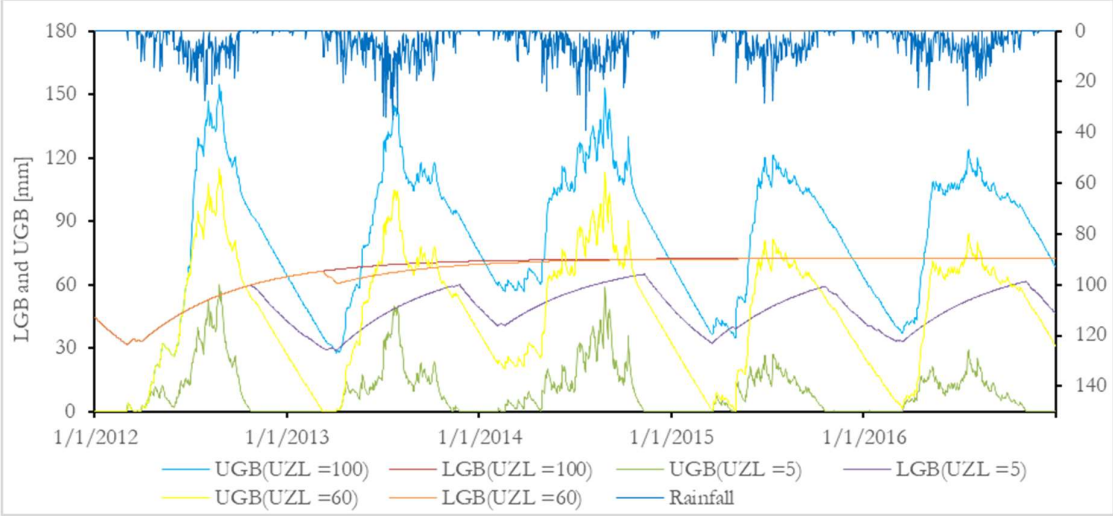


Figure 4-19 The effect of UZL parameter on upper and lower groundwater box

K0: Changing recession coefficient Ko between 0.0075 to 0.6 highly effects the overall volume and shape of the simulated hydrograph. For example, a smaller value of K0 = 0.0075 virtually does not show rising limbs, recession limbs, and peak flows while showing a straight line with RVE = -10.35 and NSE = 0.01. On the other hand, K0 value >0.15 effects peaks, rising and recession limbs. Ko has no visible effect on base flow and ETa. Besides underestimated volume at K0 =0.0075. K0 mainly affects the overall shape of the hydrograph with NSE ranges between 0.01 to 0.072**.** Most sensitive in peak flow.

K1: No effect on ETa and minor underestimation on overall volume. Varying recession coefficient K1 value within the range of 0.00015 to 0.15 affects NSE from 0.72 to 0.68 and RVE -2.6 to -2.7. K1 mainly affects the intermediate rising and recession limbs. Larger k1 increase LGB and highly decreases UGB, abrupted peak value and poor hydrograph shape.

K2: Virtually there is no visible change on baseflow, peak flow, rising and falling limb. Also, this is evidenced with no change in RVE and NSE value with changing K2 range between 0.0015 to 0.4. However, at a smaller value of K2 lower groundwater box increases and larger value LGB decreases.

MAXBAS: Length of triangular weighting function (MAXBAS) value changes between range 1 to 10 results without significant difference in NSE and RVE. At a smaller and larger value of MAXBAS constant value of NSE and RVE are observed in Figure 4-18. On the contrary, at larger value of MAXBAS virtually there is a shift in rising limbs, recession limbs, and peak flows.

4.8.2. Model calibration result

Figure 4-20 shows comparison of observed and simulated streamflow (2012-2016). The optimized parameter sets are shown in Table 4-4. As the peaks are too low also baseflows is not fitted. The mismatch of baseflow is checked for different values of PERC in sensitivity analysis section 4.8.1 (see Figure 4-17). The simulated streamflow shows consistent pattern with rainfall at the onset of each wet season (red dotted double arrow) showing reasonable model performance with NSE =0.72 and RVE = (-2.6%).

Compared to the observed streamflow (high peaks marked by red dotted boxes) the model continuously underestimates the peak flows for the entire time series exceptionally with a slight match in 2013. Likewise, the pattern of rising and recession limbs are not matched at the onset and end of the wet season in 2013 and 2014. This result is consistent with (Chaemiso et al., 2016) who forces the SWAT2005 model with eight-year (1988-1995) gauge-based data in the same river basin (whole Omo Gibe river basin see section 2.1) which results in too high base flow and too low peak flows. In the study of Chaemiso et al., (2016) the gap between (timing) simulated streamflow and observed discharge is too high with the model performance of NSE= 0.626. This study is consistent with Mohamed et al., (2010) who assess catchment behaviour with catchment modelling in Koga catchment (Ethiopia) with model performance of NSE =0.64. He found that the mismatch of low flow and high flow were affected by observed data quality.

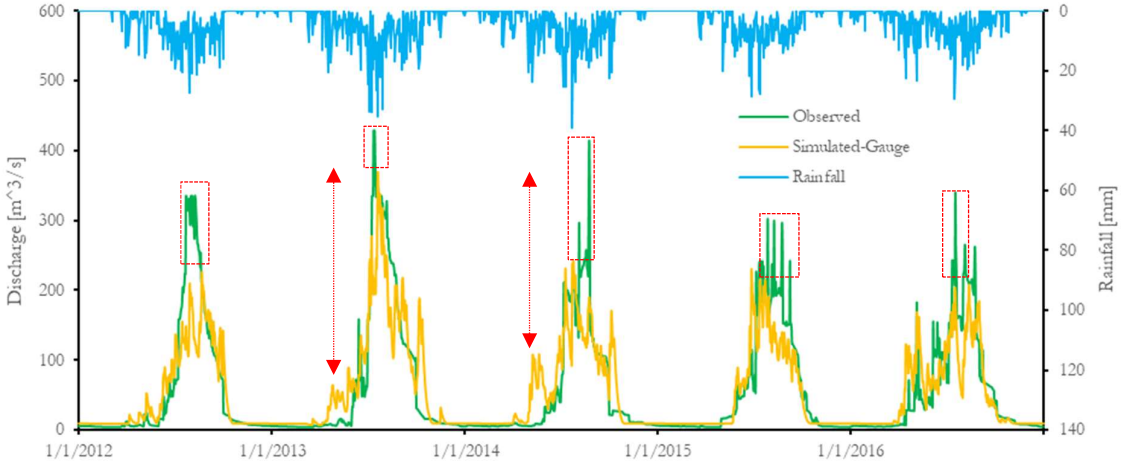


Figure 4-20 Calibration result for Wabe watershed (Jan.2012-Dec.2016)

Focusing on baseflow the calibration results for different catchment characteristics but, for same HBV model structure is compared with this study result. For instance, (Rientjes et al,2011) in their lake level simulation for nine gauged catchment used PERC = ranges between 1.02 to 2.32 (Ethiopia), (Deckers et al., 2010) used PERC=0.64 in their regionalization study for catchment variability (England), (Mohamed et al, 2010) used PERC =1.7 in their study of analysing catchment modelling (Ethiopia) and (Aguirre et al., 2013) used PERC=0.03 in their study on sensitivity and uncertainty of HBV model (Mississippi). In their study, they pointed out that PERC parameter mostly affects simulated baseflow while the model performance ranges between 0.43 to 0.86 with NSE. In this study, different values of PERC is tested and 0.4 is used as an optimized value see section 4.8.1 and Figure 4-17).

In practice, it is assumed that recorded values of observed discharge are accurate to compare the model simulation. However, as discussed in section 2.2.2 the shallow nature of the catchment and, the increase in water level during the wet season was noticed. Finally, and perhaps most importantly, the continuous mismatch in-terms of baseflow, peaks, rising limbs (onset of rainy season) could be due to, rating curve may not properly capture the stage-discharge relationship (which changes over multi annual cycle) see

section 3.3.1, the given network of rainfall stations may not capture spatial variability of rainfall (see Figure 2-5 and Table 2-1), spatial interpolation errors, model structure and considered landcover map effect.

Table 4-4 shows the optimized parameter values when the model is forced by gauge observation (condition1) and bias corrected SRE's (condition2). The three parameter values assigned for FC, LP and BETA represent parameter value at each of three vegetation zones. The first is for cereal crop second, Forest and the third is for Built up. For each independent forcing, the sensitive parameters are adjusted until baseflow, rising limb, recession limb, peak flow matches, and the difference between the simulated and observed streamflow is minimized. From the result in Table 4-4, it is understood that the optimized parameters are inevitably related to calibration forcing. The different sort of optimized parameters shows that calibrated parameters are data dependent. The comparable analysis was made by (Abebe et al., 2010b) who explains the change in parameter depend on model forcing data used and unsteadiness in dominant watershed process.

Table 4-4 Optimized model parameters of Wabe watershed. Condition 1 (calibration based on gauge rainfall) and condition 2 (calibration based on bias corrected SRE from CMORPH, ARC2, and CHRIPS 2012-2016.

				Calibration			
Parameter	Range		Default	Condition 1 (Gauge)	Condition 2 (Corrected-SRE)		
	Min	Max			CMORPH	ARC2	CHRIPS
FC [mm]	0	inf	200	(150,125,50)	(250,150,70)	(265, 185, 100)	(295,155,85)
LP [-]	0	1	1	(0.95,0.85,0.75)	(0.85,0.2,0.55)	(0.55,0.45,0.35)	(0.75,0.25,0.15)
BETA [-]	0	inf	1	(0.5,0.95,0.5)	(0.7,0.55,0.5)	(0.65,0.77,0.45)	(0.68,0.47,0.045)
PERC [mm Δt^{-1}]	0	inf	1	0.4	0.35	0.205	0.108
UZL [mm]	0	inf	20	60	56	77	69
K0 [1/Δt]	0	1	0.2	0.15	0.13557	0.15	0.125
K1 [1/Δt]	0	1	0.1	0.00075	0.0075	0.0085	0.00775
K2 [1/Δt]	0	1	0.05	0.0055	0.00055	0.0005	0.00577
MAXBAS [Δt]	1	100	1	1	1.11	1	1.2
Model performance							
Q_{bias}				0.974	1.017	0.997	0.994
NSE				0.72	0.63	0.66	0.65
RVE				-2.6	1.74	-0.3	-0.58
Validation							
Qbias	1.07			NSE	0.77	RVE	6.78

4.8.3. Validation result

Figure 4-21 shows validation hydrograph result (2009 and 2011). The analysis is based on the optimized parameter used to calibrate low flows and peak flow from gauge observation. The baseflow is well simulated by the model from the end of wet season onwards. High peaks are not well captured by the model. However, the model well simulated the patterns of rising limb high flows compared to calibration result (except in high peaks in rainy season). Baseflow is underestimated in the dry season of 2009 (January, June, and March) and becomes higher afterward (Sep 2009-Dec 2011). Perhaps, the underestimated baseflow could be due to infiltration of most rainfall from the forest and cereal crop (Enset) landcover types. And, the higher baseflows onwards could be the contribution from each land use type. Improved model performance is observed with NSE of 0.77 compared to 0.72 obtained during calibration. This is consistent with the study of (Mohammed et al, 2010) who validated HBV model with five-year time series(199-2000) in Koga (Ethiopia) and (Chaemiso et al., 2016) who validated SWAT2005 with four-year time series (1992-1994) in Omo Gibe (Ethiopia), and found out that increased model performance during validation. And also, (Rientjes et al., 2011) validated lake level simulation for six

catchments and obtained best NSE for four catchment compared to their calibration result (for example, NSE= 0.78/calibration and NSE=0.87/validation for Ribb catchment).

In this study, the reason for better model simulation performance during validation (NSE=0.77) compared to calibration (NSE=0.72), is likely to be better data quality of the validation period (less missed data). However, the RVE (6.78%) and Q_{bias} (1.07) deteriorates compared to calibration result (see Table 4-4). Although, as marked with red dotted box small peak flows in the onset of the rainy period is not simulated by the model. The qualitative and quantitative assessment reveals two principles of the model (Rientjes, 2015; Seibert and Vis, 2012) rubbish in much more rubbish out and good thinking and simple logic help. This is evidenced in this validation analysis result. Initially, the time period of 2009, 2010 and 2011 was used as validation time window and giving very poor hydrograph pattern of simulated streamflow as shown in

This is presumably due to the poor quality of observed discharge in 2010 also described in section 3.3.1. After removing 2010-year data the model performance reveals reasonable result with reference to observed discharge from 2009 and 2010.

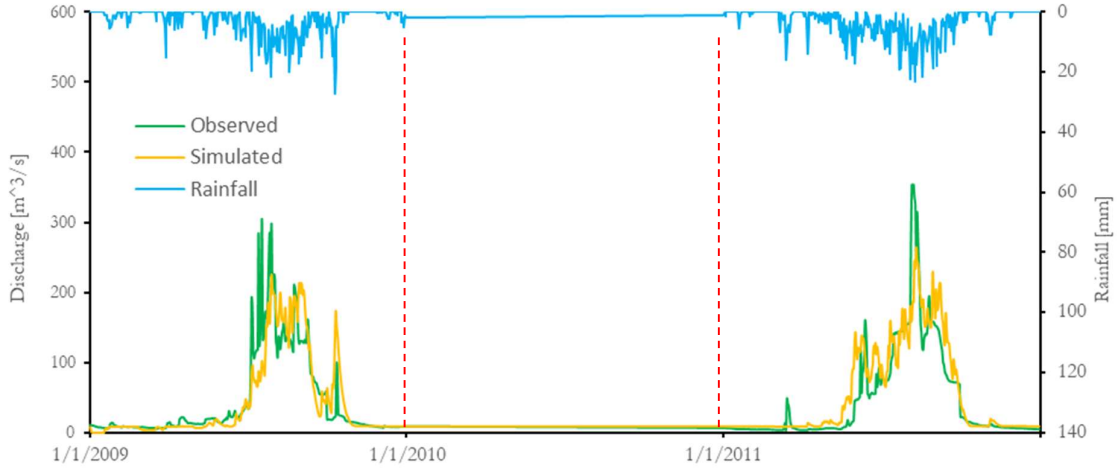


Figure 4-21 Validation result (2009 and 2011). Two vertical dotted lines indicate removed data (2010)

4.8.4. The effect of satellite potential evaporation on streamflow simulation

Figure 4-22 shows the comparison between observed, gauge simulated and simulated streamflow when in-situ potential evaporation is replaced by satellite PET. It is noticed that there are continuous high peaks in observed discharge during the wet season. Comparing the simulated streamflow from in-situ and satellite potential evaporation two differences were observed. The first is, increased peaks flows with satellite PET and second, delay in runoff transformation with in-situ PET. Perhaps, the highest peaks in streamflow during the wet season when using satellite PET could be, due to its underestimation of potential evaporation for the same season (see section 4.2). However, the slightly increased satellite PET retrieval for three years (2012,2013,2015) does not propagate in the streamflow result.

Furthermore, the actual evapotranspiration result from the model when forced with Insitu ETo and satellite PET has shown in Appendix 8A. Virtually ETa from satellite PET shows higher for entire 2016 up to dry season (June, July, and March in 2013) then it matches with Insitu (Aug 2013-Sep 2015) and afterward diminishes (0ct 2015-Dec 2016). Although, the seasonal pattern is seen from the same result. There is slightly increased bias in simulated streamflow when forced with satellite PET (Qbias=1.02)

compared with Hargreaves (Q_{bias} =0.974). Probably, the increased bias from satellite PET in consecutive time step could be, due to a smaller outflow term (actual evapotranspiration).

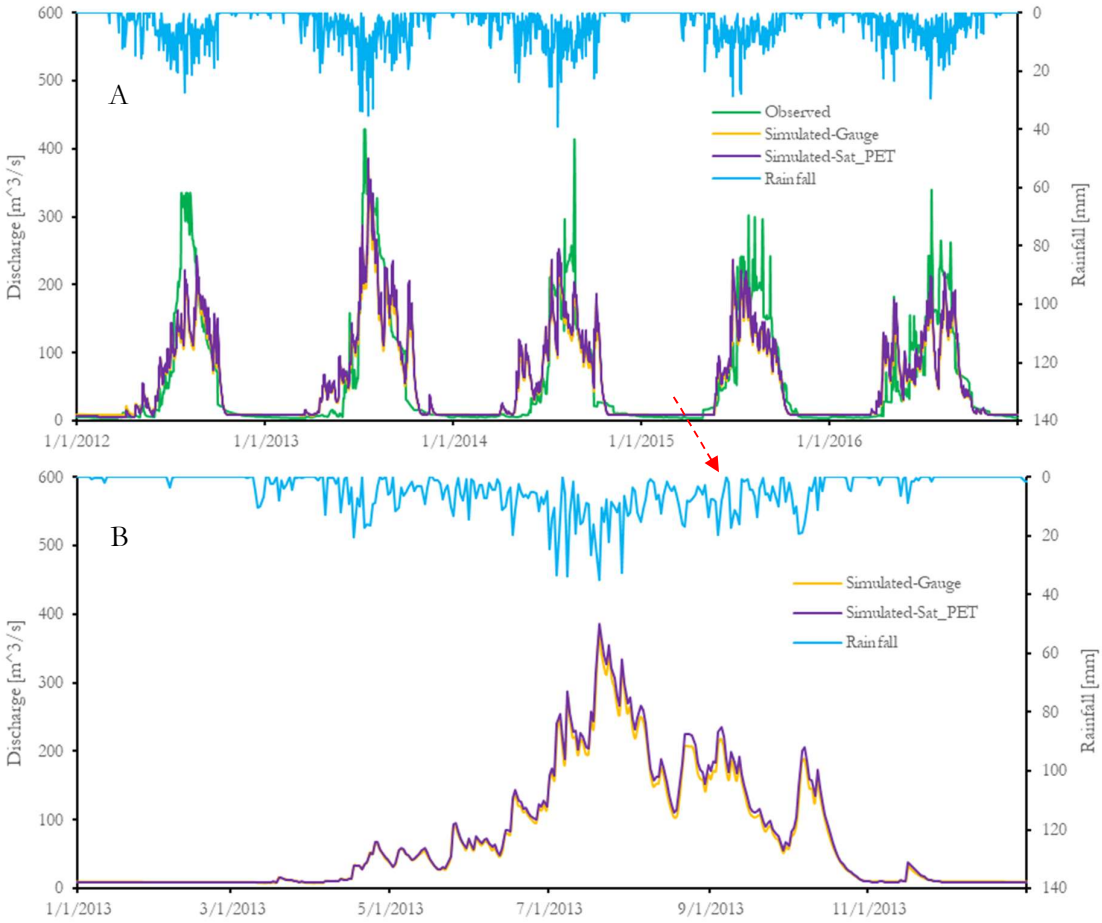


Figure 4-22 Comparison between observed, gauge simulated and satellite PET simulated hydrograph 2012-2016 (A) and zoomed to dry season of 2015 (B). The red dotted row shows zoomed data.

4.8.5. Comparing streamflow simulation with uncorrected and bias corrected SRE

Figure 4-23 A, B and C shows a comparison of simulated streamflow from uncorrected and bias corrected SRE's from CMORPH, ARC2 and CHRIPS respectively. Firstly, the result Figure 4-21 shows that all the uncorrected CMORPH, ARC2 and CHRIPS underestimates the simulated streamflow. There is a poor overall match in observed and simulated streamflow from uncorrected SRE. However, comparatively uncorrected CHRIPS shows less underestimation with RVE -3.8. The most underestimation is seen by ARC2 with RVE -25.2 followed by CMORPH with RVE -19.4. This indicates how the systematic error in satellite rainfall estimates propagates in runoff prediction. Similarly, poorly matched hydrograph characteristics are evidenced with NSE 0.63, 0.6 and 0.43 for CHRIPS, CMORPH, and ARC2 respectively as shown in Table 4-5. Uncorrected CMOPH (Figure 4-23 A) and ARC2 (Figure 4-23 B) fails to simulated peak flows in the entire time domain while uncorrected CHRIPS overestimates peak flows on onset of wet season in 2014 and 2016 (see Figure 4-23 C). In summary, in-terms of RVE, NSE, and Q_{bias} CHRIPS shows the highest performance followed by CMORPH and ARC2 (see Table 4-5).

Secondly, the effect of bias correction in CMORPH, ARC2 and CHRIPS streamflow simulation were evaluated. Particularly, substantial improvement in simulated hydrograph after bias correction was seen for

ARC2 and CMORPH those underestimated overall volumes of the hydrograph. This result is consistent with (Habib et al., 2014a) who apply TVSV bias correction for CMORPH and get an improved result (Gilgel Abay). Besides the improvement, after bias correction overestimation by three products were observed with RVE in Table 4-5. Unlike, in uncorrected case, the three SRE simulates high flows in the wet season and base flows in dry season reasonably well compared to observed as well as gauge simulated (see Appendix6
). Also, the rising and falling limbs are slightly matched than the uncorrected case. Perhaps, this could indicate the data dependency of calibration in modelling. This means the optimized set of parameters by gauge rainfall may not reproduce streamflow with different forcing term. For the corrected SRE products, in terms of accumulated error and fitting hydrograph shape in order CHRIPS, ARC2 and CMORPH perform well respectively.

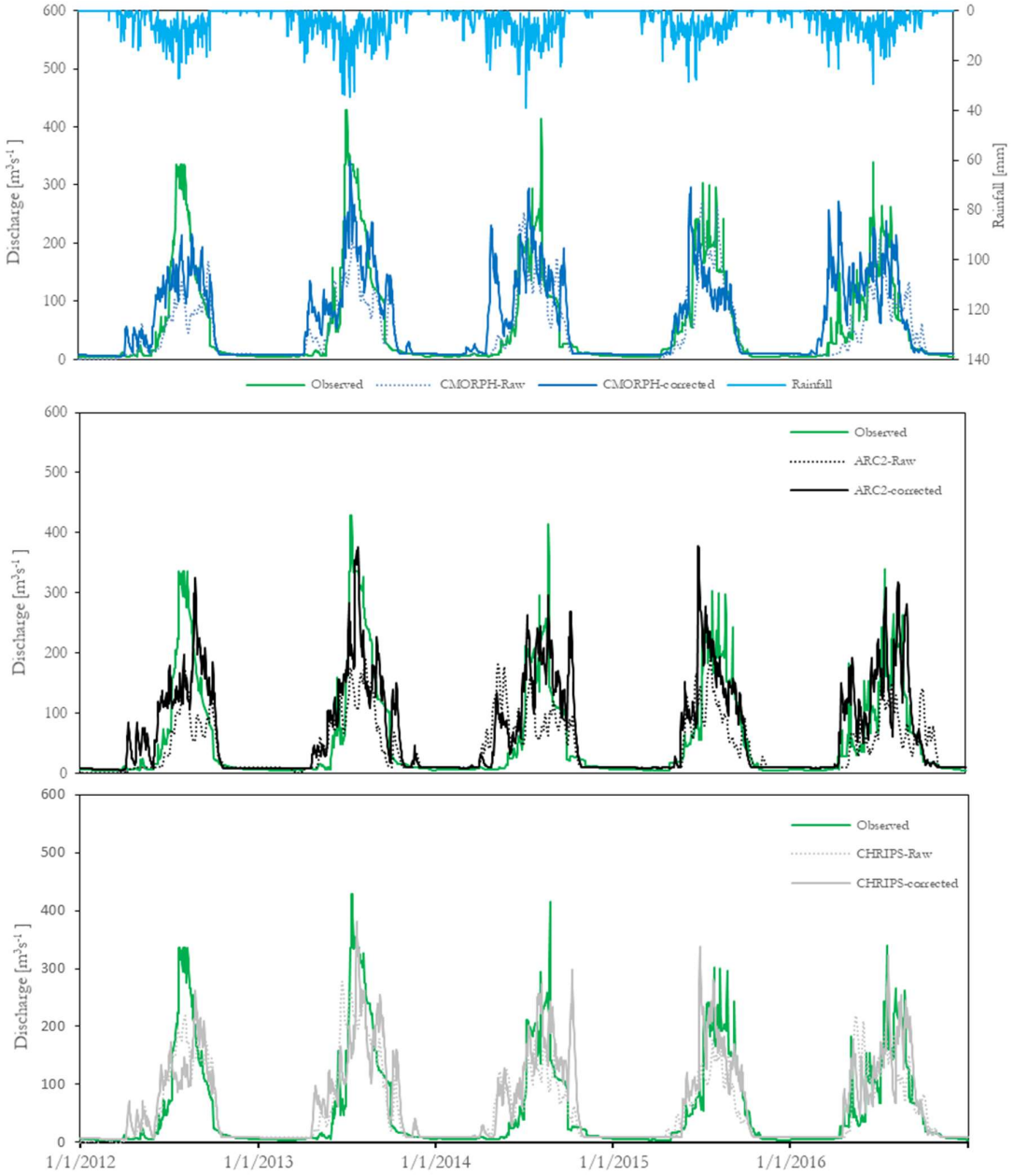


Figure 4-23 Comparison of simulated to observed streamflow by model forcing with uncorrected and bias corrected SRE's 2012-2016.

4.8.6. Calibration result for SDT scheme SRE (Condition 2)

Figure 4-24 shows stream flow comparison of observed with forcing the model by gauge and bias corrected SRE (2012-2016). Calibrating bias corrected SRE forcing results in less bias in simulating streamflow compared to gauge forcing. Calibrating the model with bias corrected satellite products reproduces the overall volume and fits most hydrograph characteristics compared with forcing the model with gauge observation (see Table 4-5). Also, the individual comparison with observed discharge and corrected SRE's products are depicted in Appendix7.

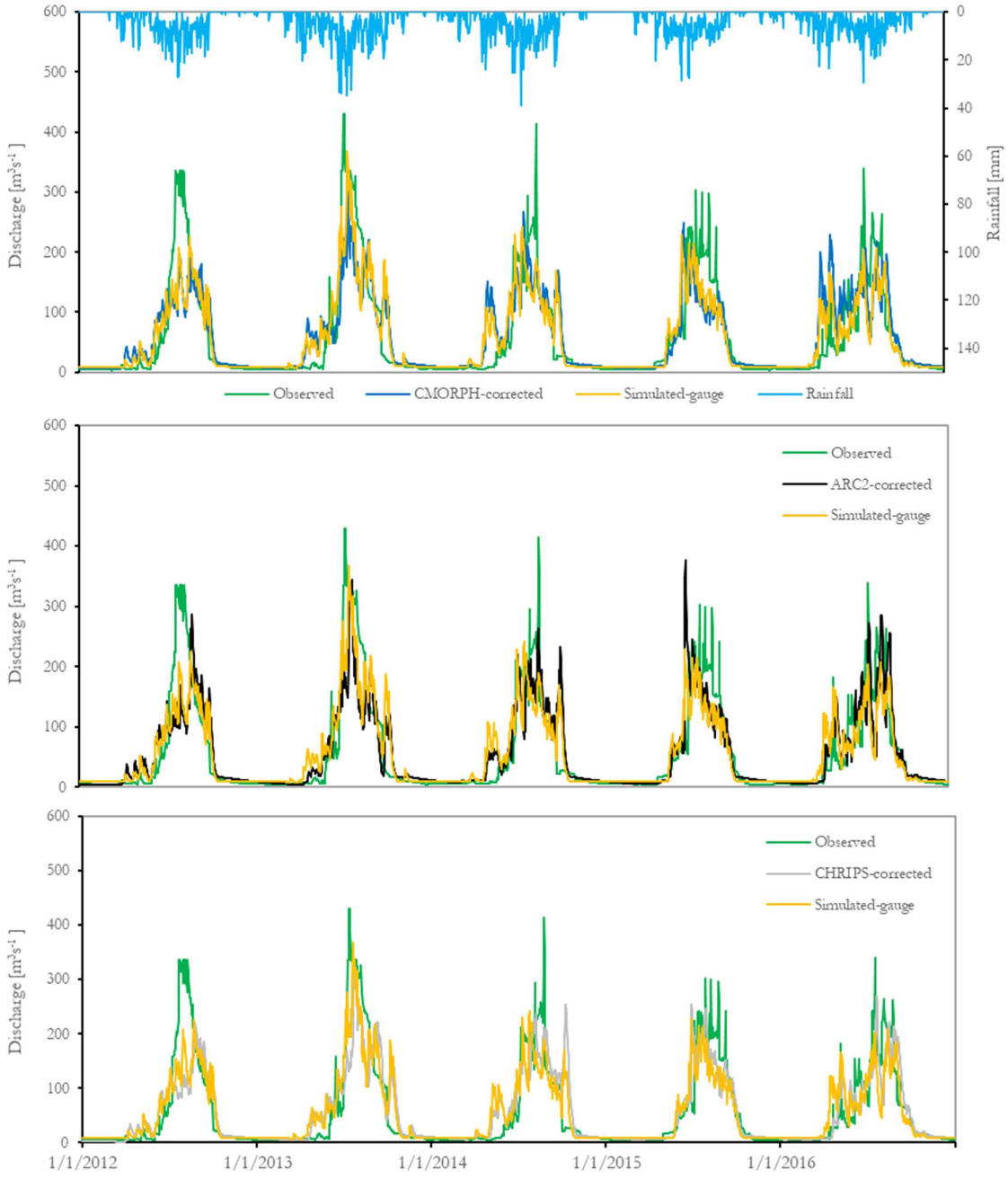


Figure 4-24 Comparison of observed and simulated streamflow by gauged rainfall and SDT bias corrected SRE

Compared to gauge-based simulation, the bias corrected ARC2, and CHRIPS shows reduced bias in simulating streamflow with Qbias=0.997 and 0.994 respectively, compared with gauge (Qb_{ias}=0.974). In contrast, fine tuning bias corrected CMORPH results in increased bias (Q_{bias} = 1.017) compared with gauge simulation but, reduced bias compared to uncorrected (Q_{bias}=0.81). However, the overall shape of simulated hydrograph is well captured with gauge based fine tuning (NSE=0.72) compared to bias corrected SRE products (NSE=0.63 for CMORPH, 0.66 for ARC2, 0.65 for CHRIPS) see Table 4-5. The simulated streamflow volume is well matched with the observed counterpart by fine-tuning the model with bias corrected ARC2 ($Qb_{ias=}$0.997 and RVE = -0.3) (see Table 4-4). This result conforms with (Habib et al, 2014) who forced the HBV model with time and space-variant bias corrected CMORPH forcing in Gilgel Abay basin (Ethiopia) results in reduced bias in simulated streamflow.

Comparing streamflow simulation using simulated streamflow from gauge forcing

Note that, the comparison is based on optimized parameters based on gauge observation. Appendix 7A and B shows a comparison of simulated streamflow based gauged rainfall, uncorrected, and SDT bias corrected CMORPH, ARC2, and CHRIPS (2012-2016). The comparison is an attempt to split up the impact of rainfall and rating curve (stage-discharge relationship error) from the uncertainty of model parameter on model streamflow simulation.

Comparison of accumulated rainfall and model simulated ETa (2012-2016)

Appendix 8A shows a comparison of streamflow (left side) and accumulated actual evapotranspiration (right side) from gauge forcing, a gauge with sat-PET, corrected and uncorrected SRE's products, and corrected and calibrated SRE's products. The accumulated streamflow from gauge forcing and gauge with satellite-PET matches with observed streamflow. Uncorrected CHRIPS matches with observed streamflow whereas CMORPH and ARC2 underestimate (Q1-left side). Bias corrected estimates of CMORPH, and ARC2 matches; however the introduction of error observed in corrected CHRIPS ((Q2-left side). Fine tuning the model with bias corrected CMORPH, ARC2 and CHRIPS reproduce the accumulated simulated streamflow correctly compare to observed discharge (Q3-left side).

4.8.7. Model water balance components and closure error analysis (2012-2016)

Table 4-5 showed a comparison of water balance components and closure error term from the gauge, a gauge with satellite-based PET-2Okm, uncorrected and corrected SRE's from CMORPH ARC2 and CHRIPS. The comparison is for the hydrological year. The hydrological year in the study area starts from the end of the dry season (1 June) from previous and ends the onset of the wet season (31 May) of the analyzed year. All values of water balance components are in $mmyr^{-1}$. Considering precipitation and ETa as the main model forcing the WB equation residual term is determined and WB closure error [%] is expressed as a percentage of precipitation. The mean annual water balance closure analysis result shows that, water is taken from the system over five years for the respective rainfall and potential evaporation forcing.

More water is taken from the system by satellite-based potential evaporation (-0.19) compared to Hargreaves ETo (-0.18). This is due to oversimulated streamflow compared to in-situ data. The lower closure error term from uncorrected satellite rainfall estimates CMORPH and ARC2 are due to underestimated rainfall which resulted in lower simulated streamflow for same in-situ based potential evaporation. Bias corrected satellite rainfall estimates show better closure term with gauge forcing for same ETo and optimized parameter set. In this way, CMORPH shows the perfect fit of closure term with in-situ. The result shows the reliability of the applied distribution transformation bias scheme.

On the other hand, recalibrating the model with bias corrected satellite rainfall estimates resulted in improved water balance closure compared to the optimized parameter set from the gauge. In this manner, ARC2 and CHRIPS show better RVE and Q_{bias} (see Table 4-5). This result is consistent with (Artan et al., 2007) who evaluate the adequacy of SRE in streamflow simulation (Sudan) and found that calibrating the model with SRE shows improved results.

Water Balance components	Calibrated model with in-situ measurement								Calibrated model with corrected SRE		
	In-situ		Un-corrected SRE			Corrected SRE			Corrected and calibrated		
	Hargreaves ETo	PET-20km	CMORPH	ARC2	CHRIPS	CMORPH	ARC2	CHRIPS	CMORPH	ARC2	CHRIPS
Simulated discharge [mm]	997	1052	795	750	968	1135	1182	1136	1012	982	990
Precipitation [mm]	1177	1177	1100	1021	1253	1317	1408	1416	1312	1397	1403
Actual ETa [mm]	386	343	377	368	422	423	417	424	523	572	542
WB Closure Error [mm]	-206	-218	-72	-96	-136	-241	-191	-145	-223	-156	-129
WB Closure Error [%]	-18	-19	-7	-9	-11	-18	-14	-10	-17	-11	-9
RVE [%]	-2.56	2.25	-19.4	-25.2	-3.8	13.2	19.5	14	1.74	-0.3	-0.58
NSE [-]	0.72	0.72	0.6	0.46	0.63	0.58	0.59	0.61	0.63	0.66	0.65
Qbias [-]	0.97	1.02	0.81	0.75	0.96	1.13	1.19	1.14	1.017	0.997	0.994

Table 4-5 Comparison of gauge, gauge with satellite-based PET-2Okm, uncorrected and corrected CMORPH ARC2 and CHRIPS water balance components and closure error analysis for Wabe watershed (2012-2016).

From the overall water balance analysis, it is concluded that satellite rainfall estimates are applicable for water resources management in ground measurement limed watershed like Wabe. This is true after appropriately correcting the systematic error from the SRE products.

5. CONCLUSION AND RECOMMENDATION

5.1. Conclusion

HBV Light version model was applied to simulate streamflow, quantify water balance components and closure error analysis with satisfactory result. In this study, three daily satellite rainfall estimates from CMORPH~8km, ARC2~10km, CHRIPS~5km was envaulted against daily rain gauge records in and around Wabe watershed over five-years, 01/01/2012-31/12/2016. Although, satellite-based potential evaporation estimates from FEWSNET~111km and PET~20km were evaluated against Hargreaves ETo.

The inter comparison result shows that three SRE's underestimate means and accumulated rainfall at the four-gauge location. The most underestimation was seen by ARC2. However, CHRIPS shows a slight overestimation of the maximum rainfall depth at four-gauge location (Agena, Fato, Wolkite, and Woliso). The coefficient of correlation of three uncorrected SRE's varies in the range between 0.2 to 0.4 for all gauge location. However, at higher elevation (Fato) ARC2 outperforms CMORPH and CHRIPS by CC (0.3), RMSE (7.4) and STDEV (5.2).

The study findings show that the detection capability of satellite rainfall estimates is seasonal and elevation dependent. Falsely detected rainfall depth in the wet season and missed bias in the dry season is the main source of error. The comparison result show in wet season CMORPH outperforms ARC2 and CHRIPS by detecting rainfall occurrence ranges between 60% at high elevated gauge location (Fato) to 80% at low elevated gauge location (Imdibir). The result shows that three SRE's fail to detect rainfall depth at highest elevation gauge location (Fato) compared to lowest elevated (Wolkite). This was shown by CSI (CMORPH=19% at Fato to 30% at Wolkite), ARC2 =17% at (Fato) to 20% at (Wolkite). The seasonal evaluation shows CMORPH, ARC2, and CHRIPS respectively well performs in detecting rainfall depth.

The analysis of the percentage of days belongs to six rain rate classes shows that the most predominant rain rate in the study area is light showers less than 1 mmd^{-1} which accounts for >than 65%. Although, the percentage of rain rates belong to 1-2.5, 2.5-5, 5-10, 10-20 and >25 are 6, 8, 9, 8 and 4 respectively. Intercomparison of satellite potential evaporation from PET-20km and FEWSNET shows that the former reasonably matches with selected in-situ Hargreaves ETo. The potential evaporation from PET-20km ranges between 1.5 to 5.9mmd^{-1} while FEWSNET ranges between 0.16 to 5.6 mmd^{-1} compared to Hargreaves which ranges between 3 to 5.9 mmd^{-1}.

Bias correction of SRE's was conducted using four bias correction methods TVSV, TVSF, TFSF, and DT. The systematic errors in SRE's are reduced after applying bias correction. However, there is an introduction of error depending on the selected indicator. For in instance spatio-temporal bias scheme (TVSV, TVSF, TFSF) increases the standard deviation compared to in-situ and uncorrected estimates. DT slightly overestimates accumulated rainfall compared to gauge station. This overestimation by DT bias scheme is due to increased rain rates of between 1 to 10mm rain rate class and decreased rain rates of <1mm compared to gauge reference. TVSF underestimates the accumulated rainfall at all gauge location while overestimating the maximum rain rate. The largest negative bias was shown by TVSF. TVSV shows the lowest daily bias followed by DT. The result shows DT has the most effective bias correction scheme with the highest correlation coefficient, lowest standard deviation, a coefficient of variation and root mean

square error. And hence, DT is selected for streamflow simulation. The result concludes the performance of bias correction methods depends on the quality of in-situ measurements (e.g., poor quality at Gubire station) and selected evaluation schemes.

The sensitivity analysis result showed that PERC, LP, BETA and Ko are the most sensitive parameters in Wabe watershed. The model reproduces streamflow with an acceptable range for calibration (NSE=0.72) and validation (NSE =0.77). The modelling result shows there is no perfect fit between the simulated and observed base flow and peak flows by the respective gauge and SRE's model forcing. The streamflow simulation from satellite-based potential evaporation (PET-2okm) results in high peaks, lower base flows, and high recession limbs compared with Hargreaves based ETo. The higher peak flow and lower base flow from satellite-based PET is presumably due to the products underestimation during wet season and overestimation in a dry season respectively. On the other hand, delay in simulated streamflow and larger base flow results from in-situ based ETo.

There is no perfect fit in water balance closure analysis (P-ET-Q≠0) instead, water is taken out of the system. Replacing in-situ potential evaporation with satellite-based PET resulted in more simulated streamflow and hence, closure error (-19%) compared with in-situ (-18%). This is probably due to less simulated ET from satellite-based PET. The study results shows recalibrating the model by bias corrected satellite rainfall forcing shows improved water balance closure error as low as -9% by CHRIPS followed by ARC2 (-11%) and CMORPH (17%).

5.2. Recommendation

The meteorological gauge location is limited and poorly distributed in the study area. For instance, there is no met-gauge observation in the north-eastern mountainous part of Wabe watershed. This gauge station could not adequately represent the spatial rainfall in the watershed and constrain to assess change of precipitation with elevation. An increase in meteorological gauge station is recommended over the high elevated mountainous region of Wabe watershed. Intercomparison of satellite-based rainfall estimates was done following point to pixel approach, we recommend further study on sub-basin scale.

In this study time limits to test different HBV Light model structure, it is suggested to test different model structures as it affects simulated streamflow. Although, it is suggested to replace the laborious trial and error parameter optimization with an automated procedure. Inconsistencies in observed streamflow time series affect the model simulated streamflow. This study suggests further analysis of the rating curve (stage-discharge relation). That was not evaluated in this study (it was not collected during field work due to data policy of MoWIE).

6. REFERENCES

Abebe, N. A., Ogden, F. L., & Pradhan, N. R. (2010a). Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: Implications for parameter estimation. *Journal of Hydrology*, *389*(3–4), 301–310. https://doi.org/10.1016/j.jhydrol.2010.06.007

Abebe, N. A., Ogden, F. L., & Pradhan, N. R. (2010b). Sensitivity and uncertainty analysis of the conceptual HBV rainfall-runoff model: Implications for parameter estimation. *Journal of Hydrology*, *389*(3–4), 301–310. https://doi.org/10.1016/j.jhydrol.2010.06.007

Aghakouchak, A., Mehran, A., Norouzi, H., & Behrangi, A. (2012). Systematic and random error components in satellite precipitation data sets. *Geophysical Research Letters*, *39*(9), 3–6. https://doi.org/10.1029/2012GL051592

Aguirre U., L. A., Miranda S., M. A., Urías L., M. A., Orona C., F., Almeyda L., I. H., Johansen N., R., & Tucuch C., M. (2013). Especies de trips (Thysanoptera) en mango, fluctuación y abundancia. *Revista Colombiana de Entomologia*, *39*(1), 9–12. https://doi.org/10.1016/j.jhydrol.2010.06.007

Allen, R. G., Luis, S. P., RAES, D., & Smith, M. (1998). FAO Irrigation and Drainage Paper No. 56. Crop Evapotranspiration (guidelines for computing crop water requirements). *Irrigation and Drainage*, *300*(56), 300. https://doi.org/10.1016/j.eja.2010.12.001

Allen, R. G., Pereira, L. S., Raes, D., Smith, M., & W, a B. (1998). Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. *Irrigation and Drainage*, 1–15. https://doi.org/10.1016/j.eja.2010.12.001

Alvarado-Barrientos, M. S., & Orozco-Medina, I. (2017). Comparison of satellite-derived potential evapotranspiration (MOD16A3) with in situ measurements from quintana roo, Mexico. *2016 IEEE 1er Congreso Nacional de Ciencias Geoespaciales: Sustainable Geospatial Technology at Service of Society, CNCG 2016 - Proceedings*, 13–17. https://doi.org/10.1109/CNCG.2016.7985078

Artan, G., Gadain, H., Smith, J. L., Asante, K., Bandaragoda, C. J., & Verdin, J. P. (2007). Adequacy of satellite derived rainfall data for stream flow modeling. *Natural Hazards*, *43*(2), 167–185. https://doi.org/10.1007/s11069-007-9121-6

Ashouri, H., Nguyen, P., Thorstensen, A., Hsu, K., Sorooshian, S., & Braithwaite, D. (2016). Assessing the Efficacy of High-Resolution Satellite-Based PERSIANN-CDR Precipitation Product in Simulating Streamflow. *Journal of Hydrometeorology*, *17*(7), 2061–2076. https://doi.org/10.1175/JHM-D-15-0192.1

Awulachew, S. B., Yilma, A. D., Loulseged, M., Loiskandl, W., Ayana, M., & Alamirew, T. (2007). *Water resources and irrigation development in Ethiopia* (No. 123). Colombo (Sri Lanka).

Bermúdez, A., Gómez, D., Muñiz, M. C., & Salgado, P. (2007). Transient numerical simulation of a thermoelectrical problem in cylindrical induction heating furnaces. *Advances in Computational Mathematics*, *26*(1–3), 39–62. https://doi.org/10.1007/s10444-005-7470-9

Bhatti, H. A., Rientjes, T., Haile, A. T., Habib, E., & Verhoef, W. (2016). Evaluation of bias correction method for satellite-based rainfall data. *Sensors (Switzerland)*, *16*(6), 1–16. https://doi.org/10.3390/s16060884

Chaemiso, S. E., Abebe, A., & Pingale, S. M. (2016). Assessment of the impact of climate change on surface hydrological processes using SWAT: a case study of Omo-Gibe river basin, Ethiopia. *Modeling Earth Systems and Environment*, *2*(4), 205. https://doi.org/10.1007/s40808-016-0257-9

Chen, F., Chen, H., & Yang, Y. (2015). Annual and seasonal changes in means and extreme events of precipitation and their connection to elevation over Yunnan Province, China. *Quaternary International*, *374*, 46–61. https://doi.org/10.1016/j.quaint.2015.02.016

Chow, V. T. R. M. adn L. W. M. (1988). *Applied_Hydrology_Chow_1988.pdf*. (B.J.Clark anad John Morris, Ed.). New York: McGraw Hill.

Congalton, R. G., & Green, K. (2009). Analysis of differences. In *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices* (pp. 121–129). Taylor & Francis Group.

D.Yates, K. S. (1994). *Potential Evapotranspiration Methods and their Impact on the Assessment of River Basin Runoff Under Climate Change* (A No. 2361). *International Institute for Applied Systems Analysis*. Laxenburg,Austria.

De Silva, R. P., Dayawansa, N. D. K., & Ratnasiri, M. D. (2007). A comparison of methods used in estimating missing rainfall data. *Journal of Agricultural Sciences*, *3*(2), 101. https://doi.org/10.4038/jas.v3i2.8107

Deckers, D. L. E. H., Booij, M. J., Rientjes, T. H. M., & Krol, M. S. (2010). Catchment Variability and Parameter Estimation in Multi-Objective Regionalisation of a Rainfall-Runoff Model. *Water Resources*

Management, *24*(14), 3961–3985. https://doi.org/10.1007/s11269-010-9642-8
Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., & Ropelewski, C. F. (2007). Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing*, *28*(7), 1503–1526. https://doi.org/10.1080/01431160600954688
Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., & Ceccato, P. (2018). Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quarterly Journal of the Royal Meteorological Society*, *144*(April), 292–312. https://doi.org/10.1002/qj.3244
Djaman, K., Balde, A. B., Sow, A., Muller, B., Irmak, S., Diaye, M. K. N., … Saito, K. (2015). Journal of Hydrology : Regional Studies Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley. *Journal of Hydrology: Regional Studies*, *3*, 139–159. https://doi.org/10.1016/j.ejrh.2015.02.002
Ebert, E. E. (2007). METHODS FOR VERIFYING SATELLITE PRECIPITATION ESTIMATES. In *Bureau of Meteorology* (pp. 345–356). Springer.
Eris, E., & Agiralioglu, N. (2012). Homogeneity and Trend Analysis of Hydrometeorological Data of the Eastern Black Sea Region , Turkey. *Journal of Water Resource and Protection*, *4*(February), 99–105. https://doi.org/10.4236/jwarp.2012.42012
Fang, G. H., Yang, J., Chen, Y. N., & Zammit, C. (2015). Comparing bias correction methods in downscaling meteorological variables for a hydrologic impact study in an arid area in China, 2547–2559. https://doi.org/10.5194/hess-19-2547-2015
FAO/UNESCO. (1974). Soil Map of the World. *Fao-Unesco*, *1*. https://doi.org/10.1038/1791168c0
Farr, T., Rosen, P., Caro, E., Crippen, R., Duren, R., Hensley, S., … Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, *45*(2005), 1–33. https://doi.org/10.1029/2005RG000183
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., … Verdin, P. (2014). A Quasi-Global Precipitation Time Series for Drought Monitoring. *U.S. Geological Survey Data Series*, *832*, 4. https://doi.org/http://dx.doi.org/110.3133/ds832
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., … Michaelsen, J. (2015). The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. *Scientific Data*, *2*, 1–21. https://doi.org/10.1038/sdata.2015.66
Gao, F., Feng, G., Ouyang, Y., Wang, H., Fisher, D., Adeli, A., & Jenkins, J. (2017). EVALUATION OF REFERENCE EVAPOTRANSPIRATION METHODS IN ARID , SEMIARID , *53*(4). https://doi.org/10.1111/1752-1688.12530
Gao, P., Li, P., Zhao, B., Xu, R., Zhao, G., Sun, W., & Mu, X. (2017). Use of double mass curves in hydrologic benefit evaluations. *Hydrological Processes*, *31*(26), 4639–4646. https://doi.org/10.1002/hyp.11377
George H. Hargreaves, & Zohrab A. Samani. (1985). Reference Crop Evapotranspiration from Temperature. *Applied Engineering in Agriculture*, *1*(2), 96–99. https://doi.org/10.13031/2013.26773
Germann, U., Galli, G., Boscacci, M., & Bolliger, M. (2006). Radar precipitation measurement in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, *132*(618 A), 1669–1692. https://doi.org/10.1256/qj.05.190
Gumindoga, W., Rientjes, T. H. M., Haile, A. T., Makurira, H., & Reggiani, P. (2016). Bias correction schemes for CMORPH satellite rainfall estimates in the Zambezi River Basin. *Hydrology and Earth System Sciences Discussions*, (February), 1–36. https://doi.org/10.5194/hess-2016-33
Habib, E., Haile, A. T., Sazib, N., Zhang, Y., & Rientjes, T. (2014a). Effect of bias correction of satellite-rainfall estimates on runoff simulations at the source of the Upper Blue Nile. *Remote Sensing*, *6*(7), 6688–6708. https://doi.org/10.3390/rs6076688
Habib, E., Haile, A. T., Sazib, N., Zhang, Y., & Rientjes, T. (2014b). Effect of Bias Correction of Satellite-Rainfall Estimates on Runoff Simulations at the Source of the Upper Blue Nile, 6688–6708. https://doi.org/10.3390/rs6076688
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
Haile, A. T., Habib, E., & Rientjes, T. (2013). Evaluation of the climate prediction center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile River. *Hydrological Processes*, *27*(12), 1829–1839. https://doi.org/10.1002/hyp.9330
Haile, A. T., Rientjes, T., Gieske, A., & Gebremichael, M. (2009). Rainfall variability over mountainous and adjacent lake areas: The case of Lake Tana basin at the source of the Blue Nile River. *Journal of Applied Meteorology and Climatology*, *48*(8), 1696–1717. https://doi.org/10.1175/2009JAMC2092.1

Hall, M., Baterson, L. and Jordan, C. (2004). ERDAS Imagine 8.5 and 8.6 training manual.
Hargreaves, G. H., Asce, F., & Allen, R. G. (2003). History and Evaluation of Hargreaves Evapotranspiration Equation, *129*(February), 53–63.
Hassan, A., & Jin, S. (2016). Water storage changes and balances in Africa observed by GRACE and hydrologic models. *Geodesy and Geodynamics*, *7*(1), 39–49. https://doi.org/10.1016/j.geog.2016.03.002
Huffman, G. J., Adler, R. F., Bolvin, D. T., Nelkin, E. J., & Code, N. G. (2009). The TRMM Multi-satellite Precipitation Analysis (TMPA), 1–23.
Islam, K., Jashimuddin, M., Nath, B., & Nath, T. K. (2018). Land use classification and change detection by using multi-temporal remotely sensed imagery: The case of Chunati wildlife sanctuary, Bangladesh. *Egyptian Journal of Remote Sensing and Space Science*, *21*(1), 37–47. https://doi.org/10.1016/j.ejrs.2016.12.005
Janssen, P. H. M., & Heuberger, P. S. C. (1995). Calibration of process-oriented models. *Ecological Modelling*, *83*(1–2), 55–66. https://doi.org/10.1016/0304-3800(95)00084-9
Jobard, I., Chopin, F., Berges, J. C., & Roca, R. (2011). An intercomparison of 10-day satellite precipitation products during West African monsoon. *International Journal of Remote Sensing*, *32*(9), 2353–2376. https://doi.org/10.1080/01431161003698286
Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, *5*(3), 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2
Kenward, T., Lettenmaier, D. P., Wood, E. F., & Fielding, E. (2000). Effects of digital elevation model accuracy on hydrologic prediction. *Remote Sensing of Environment*, *74*(March), 432–444. https://doi.org/10.1016/S0034-4257(00)00136-X
Khan, A., Richards, K. S., Parker, G. T., McRobie, A., & Mukhopadhyay, B. (2014). How large is the Upper Indus Basin? The pitfalls of auto-delineation using DEMs. *Journal of Hydrology*, *509*, 442–453. https://doi.org/10.1016/j.jhydrol.2013.11.028
Lidén, R., & Harlin, J. (2000). Analysis of conceptual rainfall-runoff modelling performance in different climates. *Journal of Hydrology*, *238*(3–4), 231–247. https://doi.org/10.1016/S0022-1694(00)00330-9
Likasa, H. G. (2013). Remote Sensing and Regionalization for Integrated Water Resources Modeling in Upper and Middle Awash River Basin , Ethiopia Supervisors : Remote Sensing and Regionalization for Integrated Water Resources Modeling in Upper and Middle Awash River Basin , Et.
Liu, W., Hong, Y., Khan, S., Huang, M., Grout, T., & Adhikari, P. (2011). Evaluation of Global Daily Reference ET Using Oklahoma's Environmental Monitoring Network-MESONET. *Water Resources Management*, *25*(6), 1601–1613. https://doi.org/10.1007/s11269-010-9763-0
LSA LISA Team. (2016). *Algorithm Theoretical Basis Document for Reference Evapotranspiration (DMETREF)* (Vol. 303).
Lyimo, P. B. (2015). Suitability of Satellite Rainfall for Hydrological Modelling in Kikuletwa Catchment, Tanzania.
Maathuis, B., & Mannaerts, C. (2013). In Situ and Online Data Toolbox, (September).
Madsen, H. (2000). Automatic calibration of a conceptual rainfall–runoff model using multiple objectives. *Journal of Hydrology*, *235*(3), 276–288. https://doi.org/10.1016/S0022-1694(00)00279-1
Mashingia, F., Mtalo, F., & Bruen, M. (2014). Validation of remotely sensed rainfall over major climatic regions in Northeast Tanzania. *Physics and Chemistry of the Earth*, *67–69*(February 2018), 55–63. https://doi.org/10.1016/j.pce.2013.09.013
Melesse, A. M., Abtew, W., & Setegn, S. G. (2013). Nile River Basin: Ecohydrological challenges, climate change and hydropolitics. *Nile River Basin: Ecohydrological Challenges, Climate Change and Hydropolitics*, (June 2018), 1–718. https://doi.org/10.1007/978-3-319-02720-3
Michael L, O. (1996). Multiple linear regression analysis using Microsoft Excel. Retrieved November 20, 2018, from http://chem-old.science.oregonstate.edu/courses/ch361-464/ch464/RegrssnFnl.pdf
Montero-Martínez, G., Zarraluqui-Such, V., & García-García, F. (2012). Evaluation of 2B31 TRMM-product rain estimates for single precipitation events over a region with complex topographic features. *Journal of Geophysical Research Atmospheres*, *117*(2). https://doi.org/10.1029/2011JD016280
Nash, E., & Sutcliffe, V. (1970). River flow forecasting through conceptual models Part I - A discussion of principles. *Journal of Hydrology*, *10*, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6
Nauditt, A., Birkel, C., Soulsby, C., & Ribbe, L. (2017). Conceptual modelling to assess the influence of hydro-climatic variability on runoff processes in data scarce semi-arid Andean catchments. *Hydrological Sciences Journal*, *62*(4), 515–532. https://doi.org/10.1080/02626667.2016.1240870

Nogueira, S. M. C., Moreira, M. A., & Volpato, M. M. L. (2018). Evaluating precipitation estimates from Eta, TRMM and CHRIPS data in the south-southeast region of Minas Gerais state-Brazil. *Remote Sensing*, *10*(2). https://doi.org/10.3390/rs10020313

Novella, N. S., & Thiaw, W. M. (2013). African rainfall climatology version 2 for famine early warning systems. *Journal of Applied Meteorology and Climatology*, *52*(3), 588–606. https://doi.org/10.1175/JAMC-D-11-0238.1

Oliveira, P. T. S., Nearing, M. A., Moran, M. S., Goodrich, D. C., Wendland, E., & Gupta, H. V. (2014). Water Resources Research, 7100–7114. https://doi.org/10.1002/2013WR015202.Received

Omondi, C. K. (2017). Assessment of bias corrected satellite rainfall products for streamflow simulation : A TOPMODEL application in the Kabompo River Basin. *Water Resources*, *MSc*, 88.

Pan, M., Li, H., & Wood, E. (2010). Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resources Research*, *46*(9), 1–10. https://doi.org/10.1029/2009WR008290

Radchenko, I., Breuer, L., Forkutsa, I., & Frede, H. G. (2014). Simulating water resource availability under data scarcity - A case study for the ferghana valley (Central Asia). *Water (Switzerland)*, *6*(11), 3270–3299. https://doi.org/10.3390/w6113270

Ramoelo, A., Majozi, N., Mathieu, R., Jovanovic, N., Nickless, A., & Dzikiti, S. (2014). Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa. *Remote Sensing*, *6*(8), 7406–7423. https://doi.org/10.3390/rs6087406

Rientjes, T. (2015). Hydrologic modelling for Intergated Water Resource. In *Lecture notes for Module 9-10 surface water stream* (pp. 1–269). Enschede.

Rientjes, T. H. M. (2016). *Hydrologic modelling for Integrated Water Resource Assessments*. *Lecture book for Modules 9-10 Surface water stream*. Enschede.

Rientjes, T. H. M., Haile, A. T., Kebede, E., Mannaerts, C. M. M., Habib, E., & Steenhuis, T. S. (2011). Changes in land cover, rainfall and stream flow in Upper Gilgel Abbay catchment, Blue Nile basin - Ethiopia. *Hydrology and Earth System Sciences*, *15*(6), 1979–1989. https://doi.org/10.5194/hess-15-1979-2011

Rientjes, T. H. M., Muthuwatta, L. P., Bos, M. G., Booij, M. J., & Bhatti, H. A. (2013). Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, *505*, 276–290. https://doi.org/10.1016/j.jhydrol.2013.10.006

Rientjes, T. H. M., Perera, B. U. J., Haile, A. T., Reggiani, P., & Muthuwatta, L. P. (2011a). Regionalisation for lake level simulation - The case of Lake Tana in the Upper Blue Nile, Ethiopia. *Hydrology and Earth System Sciences*, *15*(4), 1167–1183. https://doi.org/10.5194/hess-15-1167-2011

Rientjes, T. H. M., Perera, B. U. J., Haile, A. T., Reggiani, P., & Muthuwatta, L. P. (2011b). Regionalisation for lake level simulation – the case of Lake Tana in the Upper Blue Nile , Ethiopia, (2009), 1167–1183. https://doi.org/10.5194/hess-15-1167-2011

Romilly, T. G., & Gebremichael, M. (2011). Evaluation of satellite rainfall estimates over Ethiopian river basins. *Hydrology and Earth System Sciences*, *15*(5), 1505–1514. https://doi.org/10.5194/hess-15-1505-2011

Rwanga, S. S., & Ndambuki, J. M. (2017). Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, *08*(04), 611–622. https://doi.org/10.4236/ijg.2017.84033

Scheel, M L M; Rohrer, M; Huggel, Christian; Santos Villar, D; Silvestre, E; Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution, *15*, 2649–2663. https://doi.org/10.5194/hess-15-2649-2011

Searcy, J. K., & Hardison, C. H. (1960). Double-Mass Curves, Manual of Hydrology: part1. General Surface Water Techniques, U.S. Geological Survey, Water-Supply Paper 1541-B. *WaterSupply Paper 1541B*, 66. https://doi.org/http://udspace.udel.edu/handle/19716/1592

Seibert, J. (1997). Estimation of Parameter Uncertainty in the HBV Model. *Nordic Hydrology*, *28*(1982), 247–262. https://doi.org/10.2166/nh.1997.015

Seibert, & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrology and Earth System Sciences*, *16*(9), 3315–3325. https://doi.org/10.5194/hess-16-3315-2012

Sendama, M. I. (2015). Assessment of Meteorological Remote Sensing Products for Stream Flow Modelling Using HBV-light in Nyabarongo Basin , Rwanda. *ITC Thesis...*, 77 p.

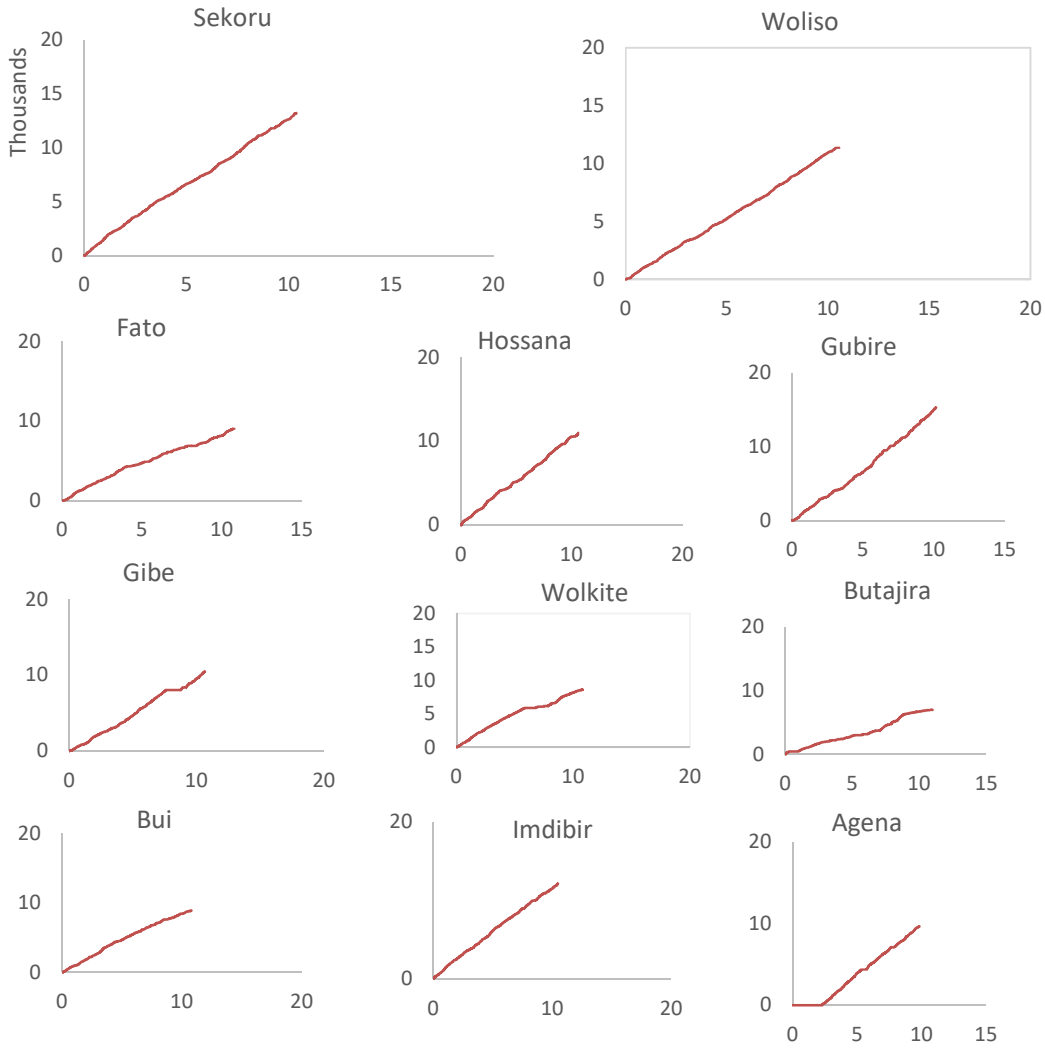
Steven, Mu, Q., Zhao, M., & Moreno, A. (2017). *NASA Earth Observing System MODIS Land Algorithm*.

Taylor, K. E. (2001). in a single diagram, *106*, 7183–7192.
Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., & Evans, J. P. (2015). How does bias correction of regional climate model precipitation affect modelled runoff ?, 711–728. https://doi.org/10.5194/hess-19-711-2015
Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., & De Roo, A. (2012). Validation of Satellite-Based Precipitation Products over Sparsely Gauged African River Basins. *Journal of Hydrometeorology*, *13*(6), 1760–1783. https://doi.org/10.1175/JHM-D-12-032.1
Thomas, J., Joseph, S., Thrivikramji, K. P., & Arunkumar, K. S. (2014). Sensitivity of digital elevation models: The scenario from two tropical mountain river basins of the Western Ghats, India. *Geoscience Frontiers*, *5*(6), 893–909. https://doi.org/10.1016/j.gsf.2013.12.008
Thyssen, C. (1911). DISTRICT No. 10, GR'EAT BASIN. *Monthly Weather Review*, (10), 1082–1084.
Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., … Zeng, J. (2009). Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research*, *114*(D24), D24101. https://doi.org/10.1029/2009JD011949
Uhlenbrook, S., Mohamed, Y., & Gragne, A. S. (2010). Analyzing catchment behavior through catchment modeling in the Gilgel Abay, Upper Blue Nile River Basin, Ethiopia. *Earth Syst. Sci*, *145194*(10), 2153–2165. https://doi.org/10.5194/hess-14-2153-2010
Uhlenbrook, S., Mohamed, Y., Gragne, A. S., & Medani, W. (2010). Analyzing catchment behavior through catchment modeling in the Gilgel Abay , Upper Blue Nile River Basin , Ethiopia, 2153–2165. https://doi.org/10.5194/hess-14-2153-2010
Vila, D. A., de Goncalves, L. G. G., Toll, D. L., & Rozante, J. R. (2009). Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. *Journal of Hydrometeorology*, *10*(2), 533–543. https://doi.org/10.1175/2008JHM1048.1
Wagner, S., Kunstmann, H., Bárdossy, A., Conrad, C., & Colditz, R. R. (2009). Water balance estimation of a poorly gauged catchment in West Africa using dynamically downscaled meteorological fields and remote sensing information. *Physics and Chemistry of the Earth*, *34*(4–5), 225–235. https://doi.org/10.1016/j.pce.2008.04.002
Ward, E., Buytaert, W., Peaver, L., & Wheater, H. (2011). Advances in Water Resources Evaluation of precipitation products over complex mountainous terrain : A water resources perspective. *Advances in Water Resources*, *34*(10), 1222–1231. https://doi.org/10.1016/j.advwatres.2011.05.007
Wilks, D. S. (2006). *Statistical Methods in the Atmospheric Sciences. 2nd ed. International Geophysics Series, Vol. 91*.
Xu, F., Jia, Y., Peng, H., Niu, C., & Liu, J. (2018). Temperature and precipitation trends and their links with elevation in the Hengduan Mountain region, China. *Climate Research*, *75*(2), 163–180. https://doi.org/10.3354/cr01516
Yang, Y., & Luo, Y. (2014). Evaluating the performance of remote sensing precipitation products CMORPH, PERSIANN, and TMPA, in the arid region of northwest China. *Theoretical and Applied Climatology*, *118*(3), 429–445. https://doi.org/10.1007/s00704-013-1072-0
Zhigang Sun, Qinxue Wang, Zhu Ouyang, Masataka Watanabe, 4Bunkei Matsushita1, & Fukushima1, T. (2007). Evaluation of MOD16 algorithm using MODIS and ground observational data in winter wheat field in North China Plain. *Okt 2005 Abrufbar Uber Httpwww Tldp OrgLDPabsabsguide Pdf Zugriff 1112 2005*, *2274*(November 2008), 2267–2274. https://doi.org/10.1002/hyp
Zotarelli, L., & Dukes, M. (2010). Step by step calculation of the Penman-Monteith Evapotranspiration (FAO-56 Method). *Institute of Food and …*, 1–10. https://doi.org/10.1016/j.scitotenv.2011.10.067

7. APPENDICES

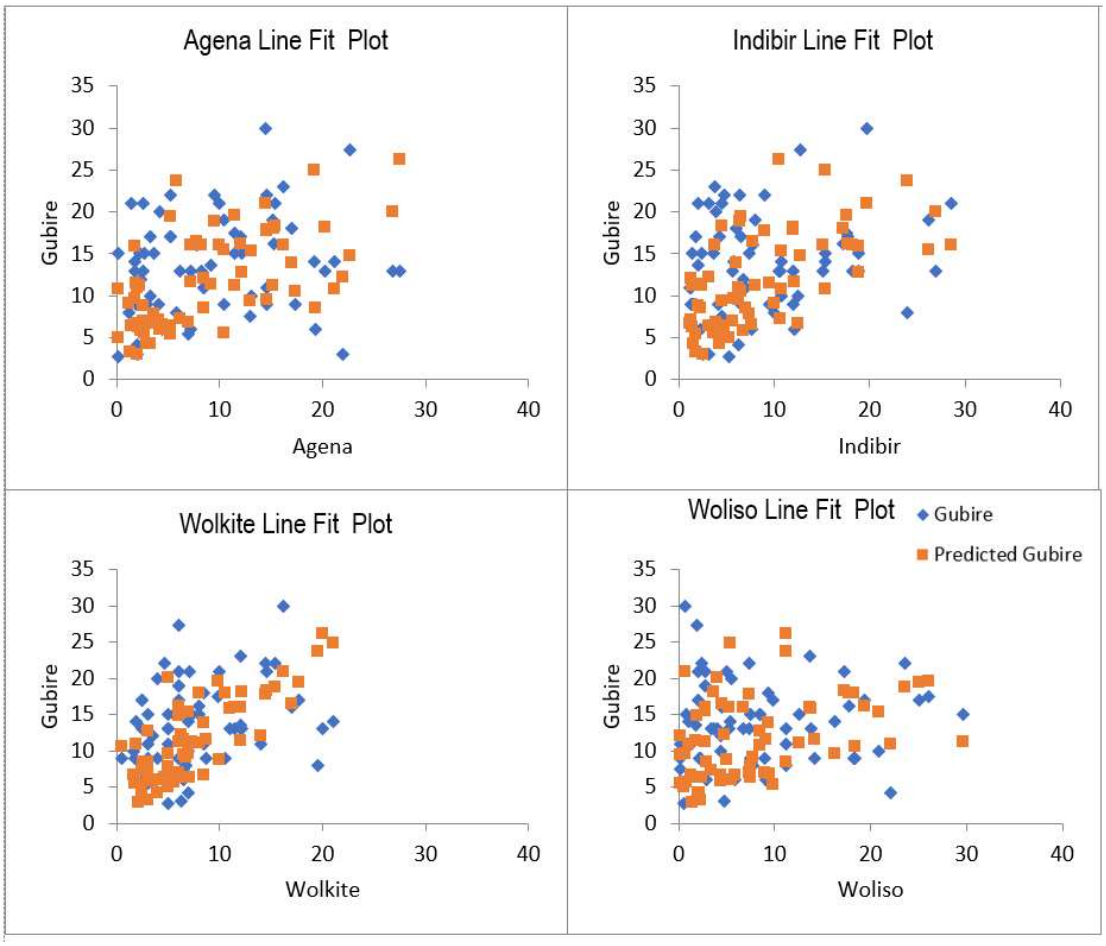
Appendix 1A

Double mass curve for collected precipitation data during fieldwork in daily bases in mm (2007-2016)



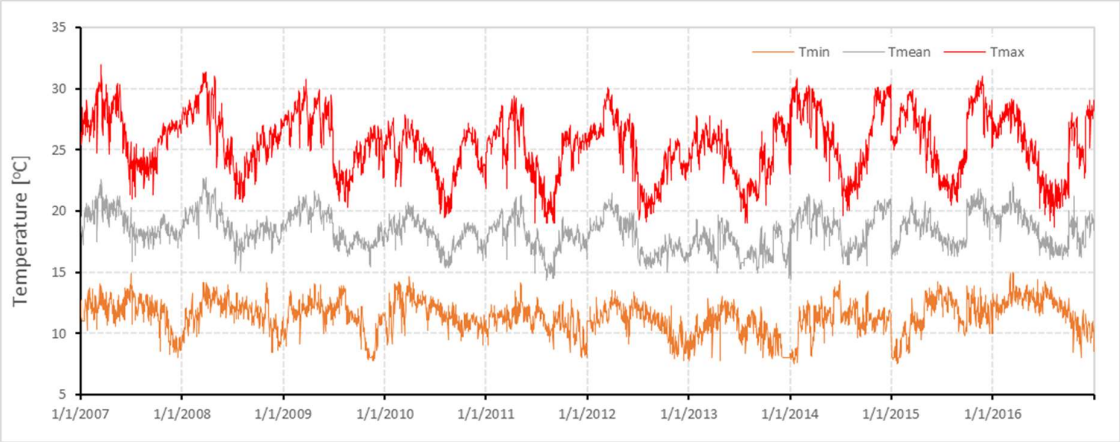
Appendix 1B

Correlation plot for fitted MLR model for Gubire station



Appendix 1C

Maximum, minimum and Average temperature from four stations (Woliso, Wolkite, Imdibir and Fato) used to determine Hargreaves potential evaporation in Wabe watershed



Appendix 2A

Statistical evaluations for uncorrected and bias corrected CMORPH, ARC2 and CHRIPS with reference to gauge for 6 stations for Wabe watershed 2012-2016

Station	Bias_Method	CMORPH								ARC2								CHRIPS							
		Sum	Mean	Max	stdDev	CC	CV	RMSE	(S/G)	Sum	Mean	Max	stdDev	CC	CV	RMSE	(S/G)	Sum	Mean	Max	stdDev	CC	CV	RMSE	(S/G)
Imdibir	R_SRE	5051	2.8	43.9	6.2	0.4	2.2	7.2	0.8	4494	2.5	48.6	6.1	0.3	2.5	7.6	0.7	5820	3.2	51.2	7.2	0.3	2.3	8.0	1.0
	Gauge	6083	3.3	58.4	6.8					6083	3.3	58.4	6.8					6083	3.3	58.4	6.8				
	BF_{TVSV}	6127	3.4	81.4	8.3	0.4	2.5	8.4	1.0	5985	3.3	86.4	9.0	0.3	2.8	9.5	1.0	6590	3.6	88.5	9.0	0.3	2.5	9.2	1.1
	BF_{TVSF}	6080	3.3	96.9	7.2	0.6	2.2	6.6	1.0	3795	2.1	58.3	5.4	0.5	2.6	6.1	0.6	4535	2.5	148.3	7.6	0.4	3.1	7.7	0.7
	BF_{TFSF}	6582	3.6	57.2	8.1	0.4	2.2	8.2	1.1	6338	3.5	68.5	8.5	0.3	2.5	9.1	1.0	6573	3.6	57.9	8.2	0.3	2.3	8.6	1.1
	SDT	6911	3.8	57.7	7.1	0.6	1.9	6.1	1.1	7220	4.0	67.5	8.0	0.6	2.0	6.9	1.2	7088	3.9	76.5	7.7	0.6	2.0	6.6	1.2
Agena	R_SRE	4731	2.6	47.5	6.0	0.4	2.3	7.2	0.7	4263	2.3	47.8	6.0	0.3	2.6	7.6	0.7	6031	3.3	71.8	8.1	0.3	2.5	8.8	0.9
	Gauge	6539	3.6	62.8	6.9					6539	3.6	62.8	6.9					6539	3.6	62.8	6.9				
	BF_{TVSV}	6636	3.6	97.4	9.4	0.4	2.6	9.3	1.0	6747	3.7	90.4	10.3	0.4	2.8	10.0	1.0	7112	3.9	82.5	10.3	0.3	2.6	10.4	1.1
	BF_{TVSF}	5953	3.3	66.8	7.4	0.6	2.3	6.6	0.9	3970	2.2	92.5	6.3	0.5	2.9	7.0	0.6	4229	2.3	73.7	6.8	0.4	3.0	7.5	0.6
	BF_{TFSF}	6165	3.4	61.8	7.8	0.4	2.3	8.1	0.9	6012	3.3	67.4	8.5	0.3	2.6	9.0	0.9	6811	3.7	81.1	9.2	0.3	2.5	9.5	1.0
	SDT	7276	4.0	65.4	7.1	0.6	1.8	6.0	1.1	7460	4.1	57.9	7.7	0.6	1.9	6.3	1.1	7520	4.1	66.0	7.7	0.6	1.9	7.0	1.2
Fato	R_SRE	4105	2.2	46.9	5.5	0.2	2.5	7.6	0.7	3713	2.0	39.2	5.2	0.3	2.6	7.4	0.6	5059	2.8	60.7	6.9	0.2	2.5	8.3	0.8
	Gauge	6020	3.3	48.5	6.6					6020	3.3	48.5	6.6					6020	3.3	48.5	6.6				
	BF_{TVSV}	5002	2.7	99.3	7.4	0.3	2.7	8.6	0.8	4525	2.5	89.4	7.0	0.3	2.8	8.2	0.8	5749	3.1	98.5	8.4	0.2	2.7	9.4	1.0
	BF_{TVSF}	4992	2.7	121.3	7.3	0.4	2.7	7.7	0.8	4774	2.6	81.3	7.8	0.3	3.0	8.4	0.8	4829	2.6	81.3	8.1	0.3	3.1	8.7	0.8
	BF_{TFSF}	5349	2.9	61.1	7.2	0.2	2.5	8.5	0.9	5236	2.9	55.2	7.4	0.3	2.6	8.6	0.9	5713	3.1	68.5	7.8	0.2	2.5	8.9	0.9
	SDT	7010	3.8	50.1	7.0	0.4	1.8	7.2	1.2	6683	3.7	52.4	7.2	0.5	2.0	7.0	1.1	6463	3.5	55.6	7.1	0.5	2.0	7.0	1.1
Gubire	R_SRE	5570	3.0	43.9	6.7	0.3	2.2	9.5	0.6	5259	2.9	46.9	6.8	0.3	2.3	10.1	0.6	5903	3.2	58.9	7.2	0.3	2.2	9.9	0.6
	Gauge	9468	5.2	67.0	9.0					9468	5.2	67.0	9.0					9468	5.2	67.0	9.0				
	BF_{TVSV}	6588	3.6	87.0	8.6	0.3	2.4	10.4	0.7	6568	3.6	93.5	9.6	0.2	2.7	11.5	0.7	7040	3.9	97.6	9.8	0.3	2.6	11.4	0.7
	BF_{TVSF}	6288	3.4	82.8	6.9	0.6	2.0	7.9	0.7	4375	2.4	70.8	6.0	0.4	2.5	8.8	0.5	4807	2.6	102.7	7.7	0.4	2.9	9.3	0.5
	BF_{TFSF}	7258	4.0	57.2	8.7	0.3	2.2	10.3	0.8	7416	4.1	66.1	9.5	0.3	2.3	11.4	0.8	6667	3.6	66.5	8.1	0.3	2.2	10.3	0.7
	SDT	10500	5.7	62.1	9.4	0.6	1.6	8.3	1.1	11217	6.1	79.1	10.4	0.5	1.7	9.5	1.2	10628	5.8	80.5	10.2	0.5	1.8	9.3	1.1
Woliso	R_SRE	4903	2.7	57.9	6.1	0.3	2.3	7.2	0.9	4749	2.6	47.7	6.1	0.4	2.3	6.8	0.8	5740	3.1	51.5	6.9	0.4	2.2	7.6	1.0
	Gauge	5719	3.1	60.9	6.6					5719	3.1	60.9	6.6					5719	3.1	60.9	6.6				
	BF_{TVSV}	5824	3.2	87.1	7.7	0.3	2.4	8.4	1.0	5754	3.1	104.2	8.3	0.4	2.6	8.4	1.0	6222	3.4	93.9	8.3	0.3	2.4	8.7	1.1
	BF_{TVSF}	6741	3.7	87.4	8.3	0.5	2.2	7.8	1.2	5145	2.8	68.3	7.5	0.5	2.7	7.2	0.9	5943	3.3	101.4	9.2	0.4	2.8	8.8	1.0
	BF_{TFSF}	6389	3.5	75.4	7.9	0.3	2.3	8.3	1.1	6698	3.7	67.3	8.5	0.4	2.3	8.3	1.2	6483	3.5	58.1	7.8	0.4	2.2	8.1	1.1
	SDT	6503	3.6	65.0	6.8	0.5	1.9	6.7	1.1	6741	3.7	83.5	7.3	0.6	2.0	6.4	1.2	6451	3.5	83.7	6.8	0.5	1.9	6.9	1.1
Wolkite	R_SRE	5532	3.0	48.8	6.6	0.3	2.2	7.3	1.1	5142	2.8	48.1	6.6	0.3	2.4	7.3	1.0	5958	3.3	59.3	7.3	0.3	2.2	7.7	1.2
	Gauge	5122	2.8	52.5	5.5					5122	2.8	52.5	5.5					5122	2.8	52.5	5.5				
	BF_{TVSV}	6427	3.5	101.8	8.6	0.3	2.4	8.8	1.3	6160	3.4	86.9	8.6	0.3	2.5	8.7	1.2	6432	3.5	94.2	8.7	0.3	2.5	8.8	1.3
	BF_{TVSF}	6590	3.6	68.2	7.3	0.4	2.0	6.9	1.3	4738	2.6	84.9	6.9	0.4	2.7	6.8	0.9	4849	2.7	100.6	7.6	0.4	2.9	7.4	0.9
	BF_{TFSF}	7208	3.9	63.5	8.6	0.3	2.2	8.9	1.4	7252	4.0	67.9	9.4	0.3	2.4	9.5	1.4	6729	3.7	66.9	8.2	0.3	2.2	8.5	1.3
	SDT	5968	3.3	61.2	5.8	0.5	1.8	5.4	1.2	6285	3.4	50.4	6.3	0.5	1.8	5.7	1.2	6312	3.5	69.8	6.5	0.5	1.9	5.8	1.2

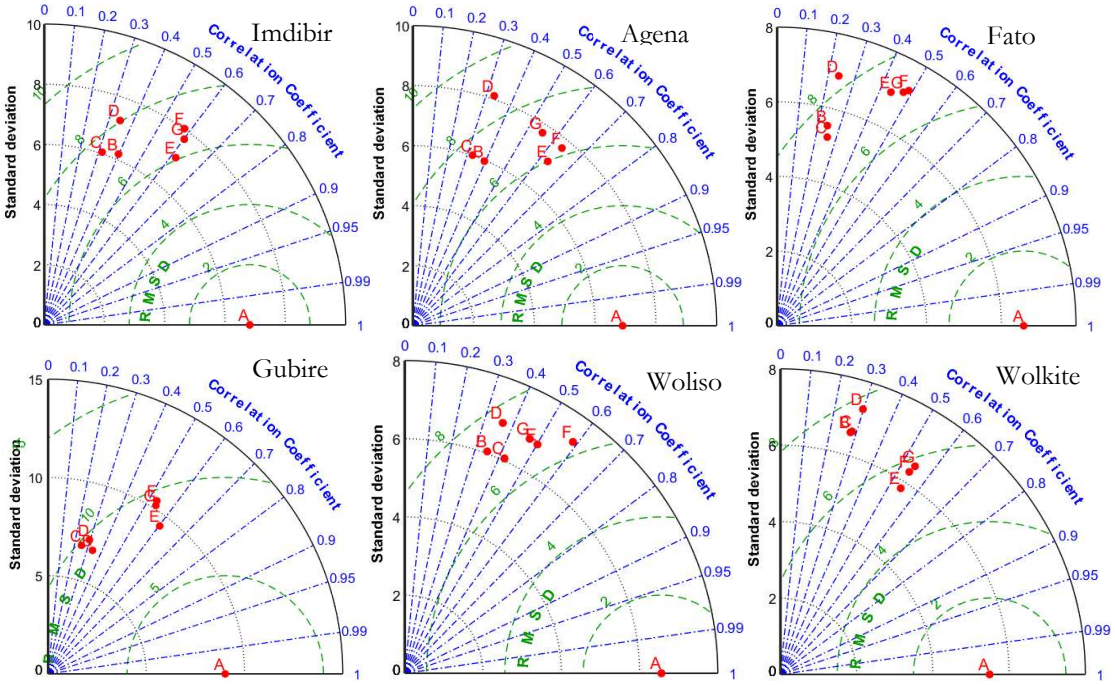
Appendix 2B

Accumulated difference in gauge, uncorrected and bias corrected satellite rainfall lumped, wet and dry season in Wabe watershed (2012-2016)

Appendix 2C

Taylor diagram showing statistical comparison of reference gauge against distribution transformation bias corrected satellite rainfall estimates from CMORPH, ARC2 and CHRIPS from (2012-2016)

The naming is A=gauge reference data for each station, B=CMOPRH, C=ARC2, and D=CHRIPS for uncorrected products whereas EGF stands for bias corrected CMORPH, ARC2, and CHRIPS respectively.



Appendix3A

HBV Light model governing equations

#	Equations	Definition
1	$P - ETa - Q = \frac{d}{dt}(SP + SM + UZS + LZS + Lake)$	General water balance equation
2	$R = IN\left(\frac{SM}{FC}\right)^{BETA}$	Recharge
3	$ETa = ET_p$ if SM $\geq$ (LP $*$ FC)	Actual evapotranspiration
4	$ETa = ETp\left(\frac{SM}{LP * FC}\right)\ if\ SM < (LP.FC)$	
5	$Cf = CFLUX * \frac{FC - SM}{FC}$	Capillary rise
6	$Q_q = K_1 * UZ^{(1+Alpha)}$	Quick flow
7	$Q_s = k2 * LZ$	Slow/base flow
8	$Q_{GW(t)} = K_2 SLZ + K_1 SUZ + K_0 \max(SUZ - UZL, 0)$	Computed runoff from response routine as a summation of three linear outflow equations

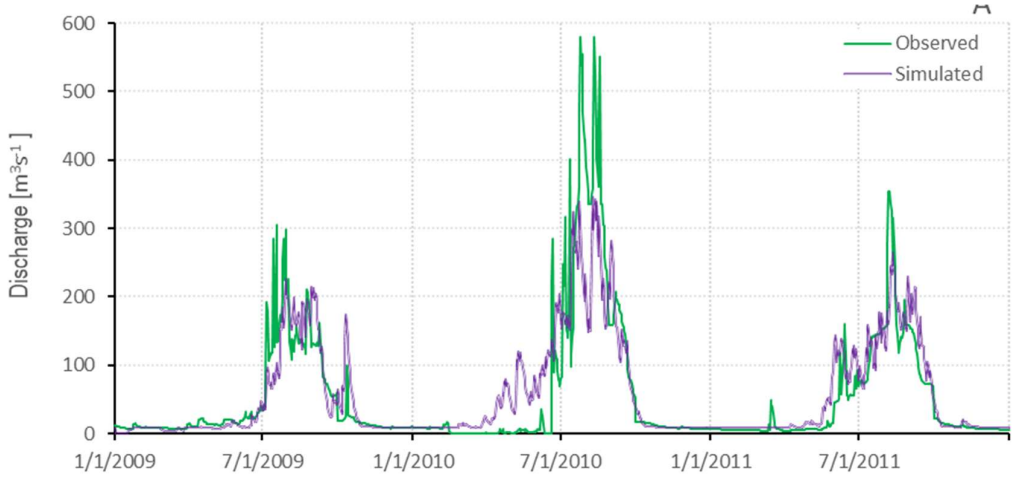
9 $$Q_{sim(t)} = \sum_{i=1}^{MAXBAS} c(i)\ Q_{(t)}(t-i+1)$$

Runoff from equation 8 is transformed by using triangular weighting function defined by parameter MAXBAS to give final simulated streamflow in [mm/day]

10 $$C(i) = \int_{i-1}^{i} \frac{2}{MAXBAS} - \left|u - \frac{MAXBAS}{2}\right| * \frac{4}{MAXBAS^2}\, du$$

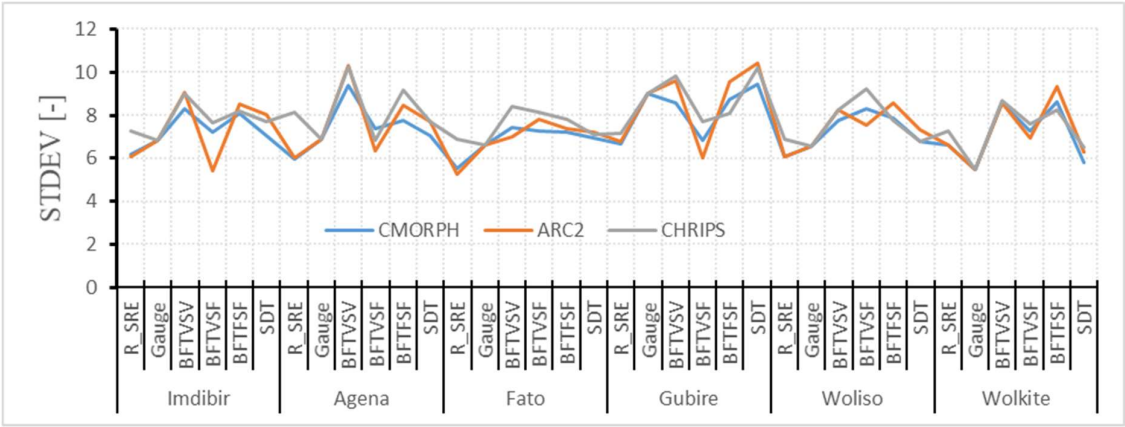
Appendix4

Validation result for Wabe watershed Jan.2009 to Dec.2011). This is to show overestimated observed streamflow in 2010.



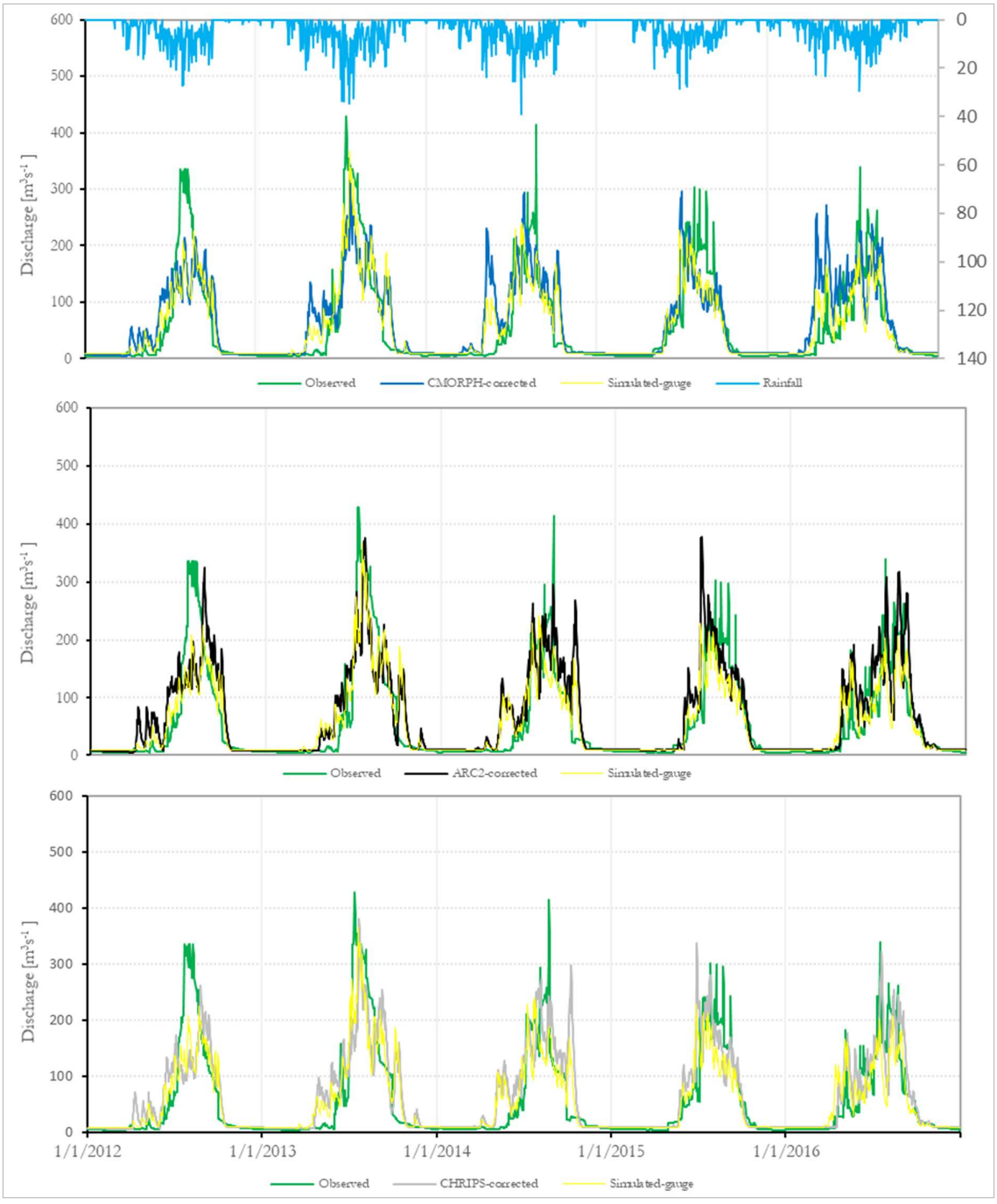
Appendix 5A

Standard deviation of gauge, uncorrected and SDT bias scheme corrected CMORPH, ARC2 and CHRIPS satellite rainfall estimates



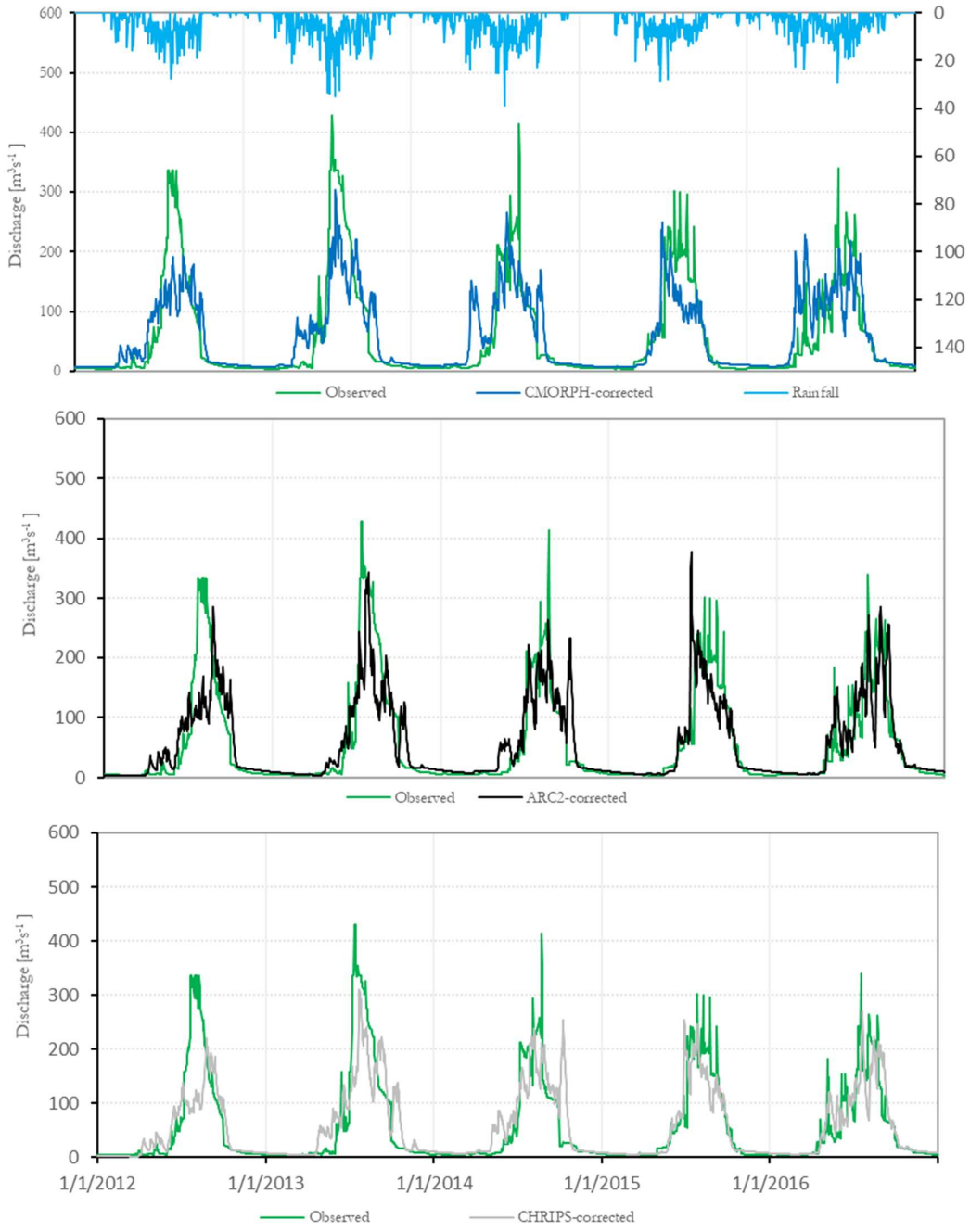
Appendix6

Comparing simulated streamflow result from gauge and SDT bias scheme CMORPH, ARC2 and CHRIPS against observed discharge in Wabe watershed 2012-2016



Appendix 7

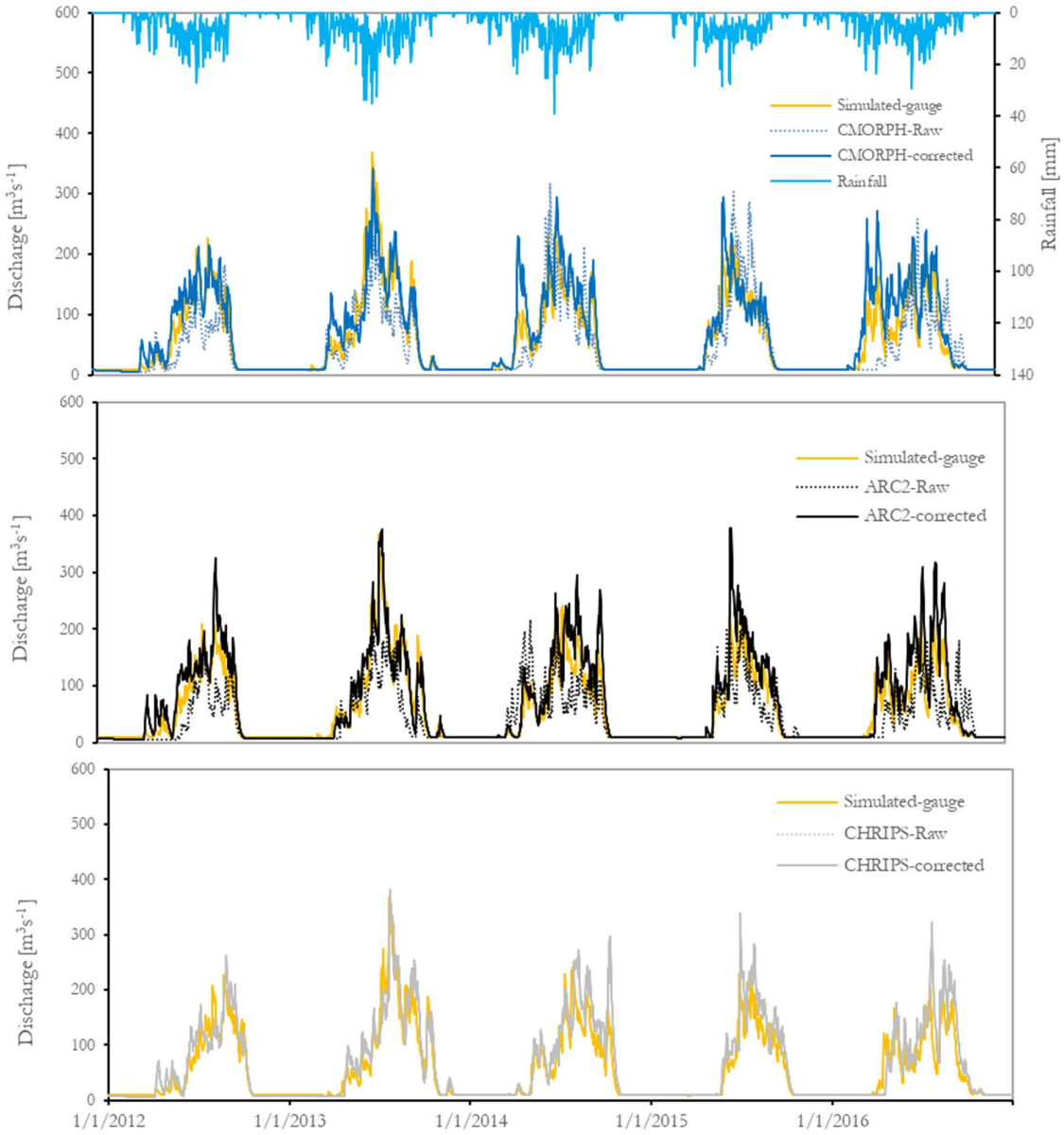
Comparing fine-tuned SDT bias scheme CMORPH, ARC2 and CHRIPS with observed discharge 2012-2016



Appendix7A

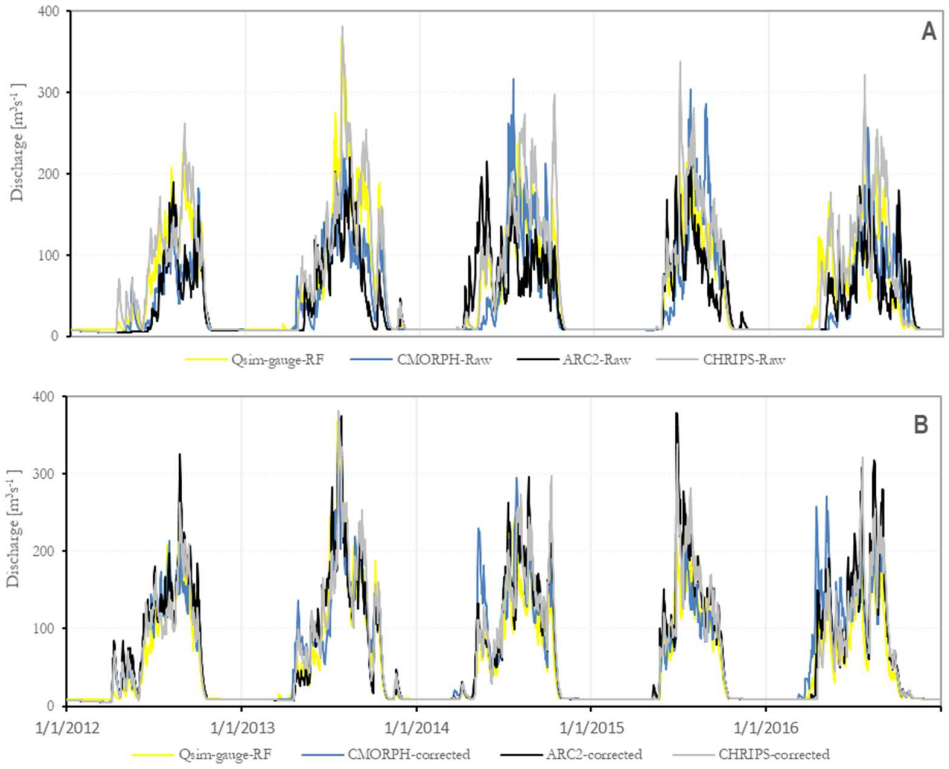
Comparison of simulated streamflow by gauged rainfall and, uncorrected and SDT bias corrected CMORPH, ARC2, and CHRIPS (2012-2016)

The motivation for this analysis is the reasonable virtually inspected pattern of simulated streamflow (from gauge rainfall) in-terms of base flows, recession limbs, peak flows as well as overall volume. The idea is first to replace observed streamflow by simulated streamflow based on gauged rainfall. Then, simulated streamflow based on satellite PET, uncorrected and corrected satellite rainfall, and from condition2 is compared against simulated streamflow based in-situ based gauged rainfall.



Appendix7B

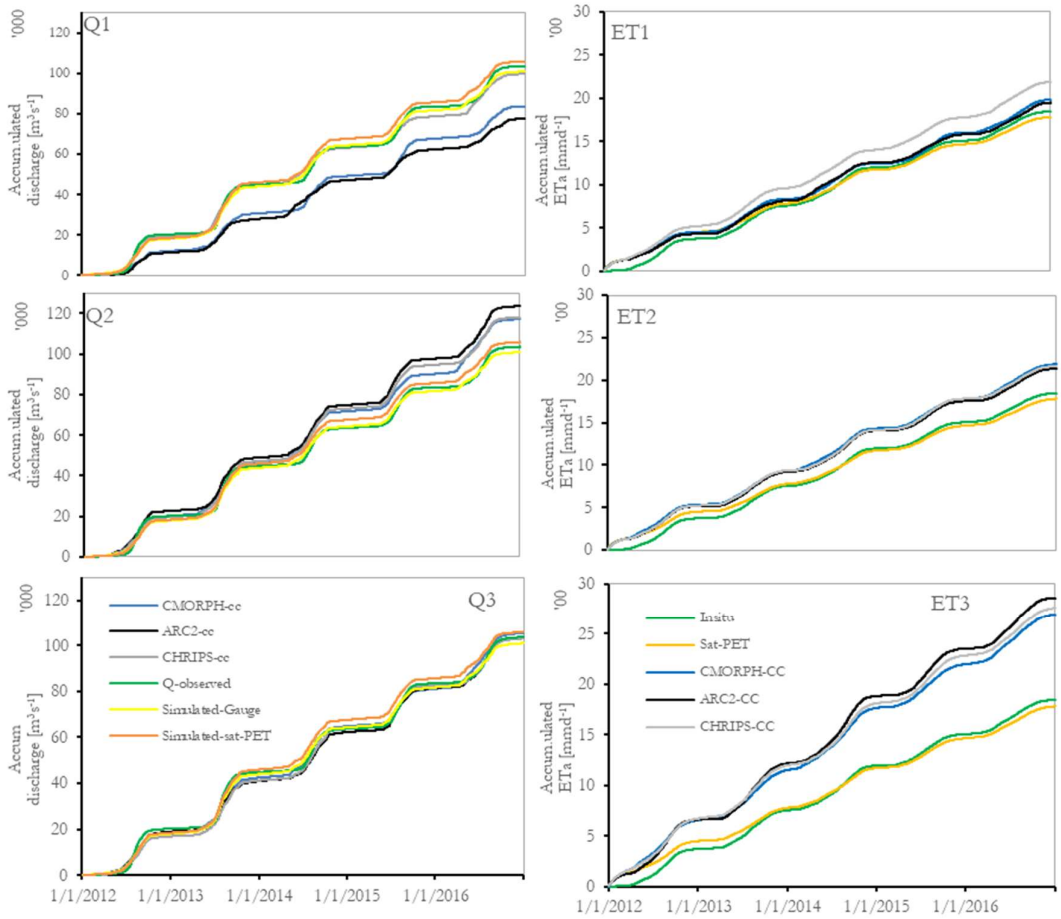
Comparing simulated streamflow result for uncorrected and distribution transformation bias scheme corrected CMORPH, ARC2 and CHRIPS against simulated discharge from gauge rainfall in Wabe watershed 2012-2016



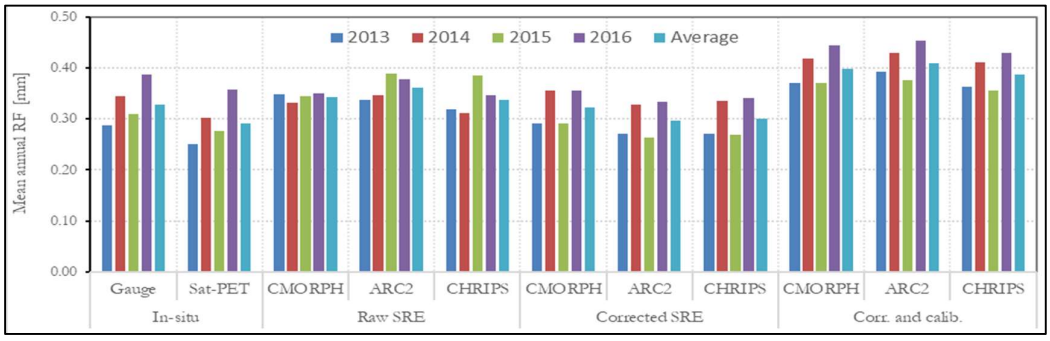
Appendix 8A

Comparison of accumulated model simulated streamflow and ETa in Wabe watershed 2012-2016.

Where: Q stands for streamflow as such Q1 refers to comparison with uncorrected SRE, Q2 refers to corrected SRE, Q3 refers to corrected and calibrated SRE and ET stands for model simulated actual evapotranspiration as such ET1 refers to comparison with uncorrected SRE, ET2 refers to corrected SRE, ET3 refers to corrected and calibrated SRE re CC stands for calibrated SRE after applying bias correction



A comparison of accumulated streamflow (left side) and accumulated actual evapotranspiration (right side) from gauge forcing, gauge with sat-PET, corrected and uncorrected SRE's products, and corrected and calibrated SRE's products. The accumulated streamflow from gauge forcing and gauge with satellite-PET matches with observed discharge. Uncorrected CHRIPS matches with observed discharge whereas CMORPH and ARC2 underestimate (Q1-left side). Bias corrected estimates of CMORPH and ARC2 matches however introduction of error observed in corrected CHRIPS (Q2 left side).
Surprisingly, fine tuning the model with bias corrected CMORPH, ARC2 and CHRIPS reproduce the accumulated simulated streamflow perfectly compare to observed discharge (Q3-left side).
In the right side shows the contribution of actual evapotranspiration in modelled streamflow.



Actual evapotranspiration coefficient (2012-2016) for the hydrological year

Appendix 9

Parameters sensitivity and their effect on simulated streamflow