NTNU

Norwegian University of
Science and Technology

Runoff Modelling For Bhutan Using Satellite Data

Sagar Ghimirey

Hydropower Development
Submission date: June 2016
Supervisor: Knut Alfredsen, IVM

Norwegian University of Science and Technology
Department of Hydraulic and Environmental Engineering

Sagar Ghimirey

Runoff Modelling for Bhutan Using Satellite Data

Master's Thesis: TVM4915 HYDROPOWER DEVELOPMENT Vår 2016

Trondheim, 10 June 2016

Supervisor: Professor Knut Alfredsen, IVM

Norwegian University of Science & Technology

Faculty of Engineering Science and Technology

Department of Hydraulic and Environmental Engineering

NTNU

Norwegian University of
Science and Technology

Forewords

The study titled "Runoff Modelling for Bhutan Using Satellite Data" is carried out in the partial fulfillment of the Programme "MSc in Hydropower Development" at the Department of Hydraulics and Environmental Engineering, Norwegian University of Science and Technology, Trondheim, Norway.

Sparse Hydro-meteorological network is a major challenge not only in planning water resource projects but also in many other applications where quality and representative information on precipitation and hydrology is an important input. Remote sensing and satellite technology has provided another platform to retrieve precipitation data in sufficient spatial and temporal resolution which is an important source especially in data scarce regions like Bhutan. Therefore, in this study analysis of one satellite precipitation product is carried out and evaluated for hydrological prediction for an ungauged catchment in Bhutan.

I hereby declare that the work presented in this report is my own and substantial outside input has been acknowledged.

Sagar Ghimirey 10th June 2016 Trondheim, Norway

Acknowledgement

I would like to express my deepest sense of gratitude to my supervisor, Professor Knut Alfredsen at the Department of Hydraulic and Environmental Engineering, NTNU who offered me an opportunity to work on the topic and provided me his continuous advice and encouragement throughout the course of this thesis. I thank him for the integrated guidance and leading me towards the successful completion of this thesis.

I would also like to express my sincere thanks to Yisak Sultan Abdella from Statkraft for introducing SHyFT which was the main basis for this thesis. His rigorous support and timely solutions to problems encountered with SHyFT was also a key to success. In addition, I also like to acknowledge the help and support of Kuganesan Sivasubramaniam regarding SHyFT.

I would also like to extend my sincere thanks to Professor Ånund Killingtveit, Professor in-charge at the Department of Hydraulic and Environmental Engineering, NTNU for his support and guidance during my study at NTNU.

My sincere thanks to Ms Dechen wangmo and Ms Wangmo of Deparement of Hydropower and Power Systems, Ministry of Economic Affairs, Thimphu, Bhutan for assisting me with the data and GIS.

I wish to thank Ababe Girmay Adera (Alumni) for introducing GIS and many other valuable inputs that he has provided during the initial phase of the Thesis. In addition, I would like to thank all my colleagues who in one or the other way have helped and supported me during my study and in the thesis.

Lastly, I would like to thank my beloved wife Chandrika Mongar, my daughter Anushka Ghimirey and my parents for their continued inspiration, and support during my pursuit of Master's program and bearing with me for having gone from their life for two years.

NTNU
Norwegian University of
Science and Technology

Faculty of Engineering
Science and Technology
Department of Hydraulic and
Environmental Engineering

M.Sc. THESIS IN

HYDROPOWER DEVELOPMENT

Candidate: Sagar Ghimirey

Title: Runoff Modelling for Bhutan Using Satellite Data.

1 BACKGROUND

Rainfall – runoff modelling can be a tool to develop runoff series from ungauged catchments using regional modelling approaches or parameter transfer methods. Access to high quality and representative precipitation data is crucial for the quality of rainfall-runoff modelling, and access to such data can be a problem for the applicability of the rainfall-runoff models. In the recent years a number of gridded precipitation products have been developed based on satellite and other remote sensing data. These have several promising features regarding areal coverage and access to precipitation in areas with few gauges, but the accuracy of the data and the applicability to modelling needs to be tested. The objective of this thesis is to evaluate satellite data for runoff modelling in Bhutan.

2 MAIN QUESTIONS FOR THE THESIS

1. Study literature on the application of satellite based precipitation, particularly with a focus on Bhutan or in other areas with similar terrain features. Evaluate sources for precipitation data and retrieve datasets relevant for the study site and time period available. The available data sources should be clearly described for future reference.
2. Assess existing precipitation data with a focus on data quality and the length of measurement period. For an evaluation of the satellite data, gauges with measurement periods that overlaps should be identified.

3. Comparison of ground measured precipitation and gridded data satellites at gauge locations. Statistical analysis and evaluation of data quality. Evaluate the use of the remote sensed data for model applications, and do bias correction of the data if necessary. The outcome of this process should be time series of gridded data ready for rainfall-runoff modelling.
4. Set-up a distributed hydrological model for one or more catchments in Bhutan and calibrate it using both gauges and satellite data. Evaluate the calibration and the goodness of the model using both sources of input. This should also include an evaluation of the differences between the two input sources.
5. Use the model to extract runoff series for ungauged locations. Evaluate the runoff generation in the catchments and the potential of using remote sensed data to model runoff in ungauged locations.

3 SUPERVISION, DATA AND INFORMATION INPUT

Professor Knut Alfredsen will supervise the thesis work and assist the candidate to make relevant information available.

Discussion with and input from colleagues and other research or engineering staff at NTNU, SINTEF, power companies or consultants are recommended. Significant inputs from others shall, however, be referenced in a convenient manner.

The research and engineering work carried out by the candidate in connection with this thesis shall remain within an educational context. The candidate and the supervisors are therefore free to introduce assumptions and limitations, which may be considered unrealistic or inappropriate in a contract research or a professional engineering context.

4 REPORT FORMAT AND REFERENCE STATEMENT

The thesis report shall be in the format A4. It shall be typed by a word processor and figures, tables, photos etc. shall be of good report quality. The report shall include a summary, a table of content, lists of figures and tables, a list of literature and other relevant references and a signed statement where the candidate states that the presented work is his own and that significant outside input is identified.

The report shall have a professional structure, assuming professional senior engineers (not in teaching or research) and decision makers as the main target group.

The thesis shall be submitted no later than 10th of June 2016.

Trondheim 14th of January 2016

_____

Knut Alfredsen

Professor

SUMMARY

Due to the limitations posed by the difficult terrain and lack of resources to install and manage the hydro-meteorological stations, Bhutan faces challenges related to sparse data network while planning water resource projects in ungauged catchments. Precipitation is an important source of input to predict runoff for an ungauged catchments using some model but, the distribution of the existing rainfall gauges in Bhutan is clustered at lower altitudes and are located in the valleys, which, neither represent well the high spatial variability of precipitation nor they provide information on the climatic conditions of the headwater regions. Therefore, as an alternative, an attempt is made in this study, to predict the precipitation induced runoff using satellite precipitation product for Dangchhu catchment in the Punatsangchhu basin.

TRMM 3B42 version 7 daily precipitation estimates with a spatial resolution of 0.25° x 0.25° was acquired for comparison and evaluation against the ground measurements. The daily CPOD_S of 0.79 on average showed that ability of TRMM to detect precipitation good but the R2 value was low and negative in most of the pixels with small trend of reduction with increasing elevation. On monthly and annual scales, better results were obtained with R2 and RR values ranging from 0.73 to 0.98 and 0.71 to 1 respectively. The monthly and annual data sets performed equally well with no trend of increasing or decreasing R2 with elevation. In all the time scales there was underestimation of precipitation by the TRMM mostly in the southern regions and overestimation in the central and northern regions. Overall, there was a general qualitative match of satellite data with the gauge in terms of timing but there were quantitative differences in all time scales and hence bias correction was necessary.

SHyFT was calibrated and validated for Kerabari catchment in the basin with three input cases; [1] gauge precipitation, [2] BCSE and [3] RSE. There was consistent underestimation of simulated runoff in most years with RSE thereby, indicating the need to adjust the biasness in the satellite estimates. The BCSE and gauge simulated hydrographs were able to match relatively well with the observed hydrographs at Kerabari. Average R2 for the three input cases were 0.78, 0.73, 0.65 respectively. The calibrated model with the input case 1 and 2 was used to simulate flow for and an ungauged catchment (Dangchhu) and for some gauged catchments in the upstream for the period 2003 to 2007. Simulated hydrograph with BCSE was estimating higher flows compared to gauge inputs at Dangchhu and it was otherwise when a new simulation was conducted at Wangdue flow gauging station which has more ground networks. R2 at Wangdue were 0.62 and 0.26 for case 1 and 2. It was learnt that the performance of simulation with input case 1 and 2 was increasing with the increase in catchment size and ground network. Since RSE is underestimating flows and the BCSE is purely dependent on quality and density of gauges, the product selected for study could not be used independently in small catchments therefore other satellite data with much finer resolution should be assessed.

Table of contents

List of Appendix

List of Figures

List of Tables

Abbreviations

APHRODITE:	Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation of Water Resources
BCSE:	Bias Corrected Satellite Estimates
CHIRP:	Climate Hazards Group InfraRed Precipitation
CMORPH:	Climate Prediction Center Morphing Technique
DEM:	Digital Elevation Model
GSMaP:	Global Satellite Mapping of Precipitation
PERSIANN:	Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks
RSE:	Raw Satellite Estimates (estimates without bias correction)
SHyFT:	Statkrafts Hydrological Forecasting Tool Box
TRMM:	Tropical Rainfall Measurement Mission

1 Introduction

1.1 Background

While Bhutan is honored by nature with gushing streams, deep narrow gorges and huge head accessible within little stretch of waterway for development of hydropower projects, it is expensive to construct and manage Hydro-meteorological stations in such a difficult topographical territory. Though huge efforts and investments are being made regularly to improve the data quality and density of ground hydro-meteorological stations, Bhutan still experiences challenges related to inadequate hydro-meteorological network especially at higher altitudes. Majority of the current hydrological stations are situated on the North-south streaming rivers and they give data recorded from 5 years to the length of 20 years. Projects are also being built on these rivers however, there are additionally numerous east-west streaming rivers which has enormous potential and yet needs good hydrological information for planning water resource projects.

In these circumstances, traditional methods such as regression analysis and transposition of historical data using a gauged catchment in the same region is a common practice to develop a runoff series for an ungauged catchment. Nonetheless, we have learnt in classes about the fact that with the increase in difference in area between the ungauged and gauged catchment, errors will normally have a tendency to be growing because of diverse spatial and temporal distribution of precipitation, diverse geography, vegetation spread, land use pattern, geology and so on. Another method would be to use the ground meteorological data with satisfactory spatial and temporal resolution and do rainfall runoff modelling. The parameters of hydrological models for catchments with few or no data on runoff can be evaluated utilizing regional data also. However, the performance of the model is purely dependent on the quality and density of rainfall gauges which is not always available, especially in developing countries like Bhutan. While planning important projects like hydropower in such an ungauged catchment, the first option should be to quickly think of installing a new gauging station. It will be worthwhile to spare some time and resources to carry out a short series measurement of the ungauged catchment for comparing with the long time series of the gauged catchment and assessing the outcome of the transposition or model studies.

Innovation in remote sensing and satellite technology has provided another platform of retrieving regional and global based precipitation data with high spatial and temporal resolution. This can be used effectively and economically to fill the gaps in the ground

instruments, predict precipitation induced runoff and in many other applications where in-situ data is sparse or not available.

Therefore, in this study, a focus is made to carry out a rainfall runoff modelling for Bhutan using satellite data and generate a runoff series for an ungauged catchment in the Punatsangchhu basin. The modelling will be performed using the Statkraft's Hydrological Forecasting Toolbox (SHyFT).

1.2 Objectives

The main objective of this study is predict runoff for an ungauged catchment in Bhutan using satellite data and the related sub-objectives are to:

1. Assess the sources of gridded satellite precipitation data and acquire datasets appropriate for the study area
2. Evaluate the quality of existing precipitation data from gauges, then compare the satellite data with the gauge data and do bias correction if required
3. Setup and calibrate the hydrological model using both input sources evaluate the results and
4. Access the potential of the model to generate runoff for an ungauged catchment in the data scares region of Bhutan using the satellite data.

1.3 Organization of report

The report contains eleven chapters which are organized as follows:

Chapter 1 gives a general background of the problem faced by Bhutan in relation to sparse network of hydro-meteorological stations with some highlights on the possible solutions and objectives of carrying out this study.

Chapter 2 includes findings from some of the studies carried out in Bhutan and in similar region based on which a suitable product is selected for this study. The chapter also evaluates the sources for satellite based precipitation products and gives some details on TRMM satellite data.

Chapter 3 describes the study area, its location, topography, climate, land cover, rivers and hydrological regime.

Chapter 4 assesses the required ground data with some focus on spatial coverage and data quality followed by some comparison and evaluation of satellite data against the gauge records through some statistical analysis and graphical methods.

Chapter 5 introduces SHyFT along with some description on how to setup and configure the model for the desired catchment and carry out rainfall runoff modelling.

Chapter 6 deals with the preparation of input data and their formats for SHyFT.

Chapter 7 deals with the calibration and validation of the model for Kerabari catchment in the Punatasngchhu basin

Chapter 8 discusses the results of calibration and validation, and then evaluates the calibration and goodness of the model using both satellite and gauge precipitation as input.

Chapter 9 deals with the uses of the calibrated model to generate runoff series for an ungauged catchment in the basin and gives some evaluations on the potential of using satellite data to model runoff in the data scarce regions of Bhutan.

Chapter 10 includes the conclusion on the study and gives some recommendations for further study.

2 Literature review

2.1 Background

Precipitation is an important element in the hydrological cycle and is very critical input for diverse areas such as water resource projects, weather forecasts, landslides mapping, flood forecasting and mapping, agriculture, industries, including researchers and decision makers. Therefore, an accurate measurement of precipitation on a range of space and time resolution is crucial. The most conventional technique of measuring precipitation is through using rain gauge networks and depending on the availability of resources weather radars are also quite common. However, it is very difficult to have a good distribution and density of these networks of rain gauge and radars due to the limitation provided by the earth's topography and resources. While some regions, especially in the developed countries have adequate ground network with good spatial and temporal coverage, most developing and under developed countries have very few or no ground networks (Kidd, 2009). Innovation in remote sensing and satellite technology has provided another platform of retrieving regional and global based precipitation data with high temporal resolution. This can be used effectively and economically to fill the gaps in the ground instruments, predict precipitation induced runoff and in many other applications where data is sparse (Bajracharya, 2014).

It is now possible to acquire the information about the earth's surface and atmosphere through Earth observation satellites. Satellite observation system have been used since 1960s by housing both active and passive sensors on the satellites. Passive remote sensing measures the natural energy or the radiation of the earth. Active remote sensing gathers data by actively sending out signals and interact with the target of interest. These sensors fall into two fundamental classes: visible (VIS)-infrared (IR) sensors accessible from geostationary (GEO) and Low-Earth circling (LEO) satellites and Microwave sensors, as of now just accessible from LEO satellites (Gruber, 2008). **Figure 2.1**below shows an observing system of meteorological satellites and **Table 1**shows key low Earth orbiting and geostationary satellites and sensors currently deployed by mainstream precipitation algorithms.

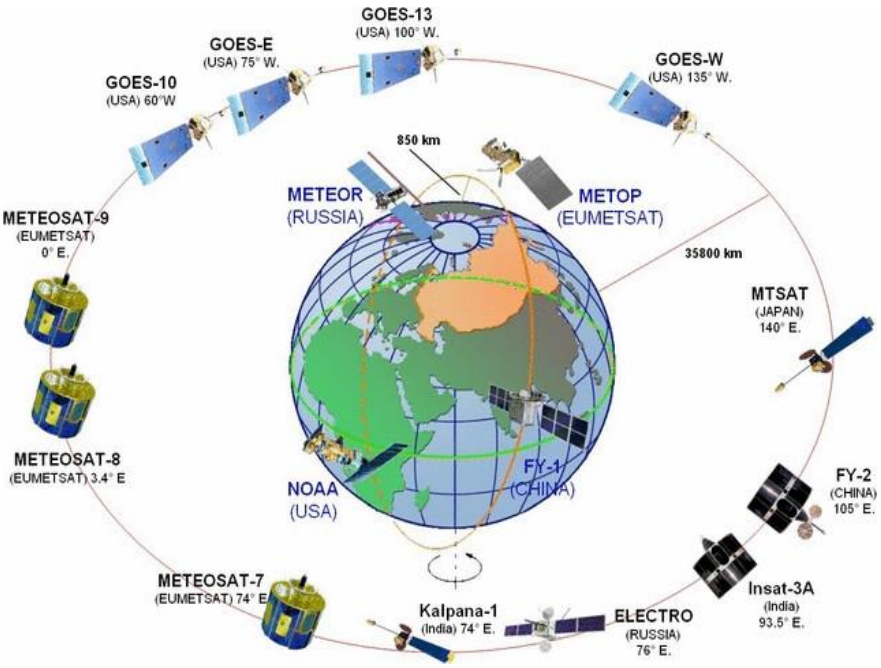


Figure 2.1: The observing system of Meteorological Satellites
The Geostationary satellites stationed about 35,800 km above the earth's surface while the Low Earth Orbiting satellites orbits the earth at about 850 km above the earth (Kidd, 2009).

Table 1: Key low Earth orbiting and geostationary satellites (Kidd, 2009)

Low Earth orbiting satellites				
Satellite	Sensor	Spectral range	channels	Resolution
NOAA 10/11/12	AVHRR	VIS & IR	5	1.1 km
	AMSU A and B	PMW	15/5	50 km (best)
	TOVS (HIRS/MSU/SSU)	Sounder		
DMSP F-13/14/15/16	SSMI/I & SSM/IS	PMW	7	
TRMM	TMI	PMW	9	5-50 km
	PR	Radar	1	4.3 km
Geostationary satellites				
GOES E/W	GOES I-M Imager	VIS & IR	5	1 & 4 km
Meteosat 5,7,8	MVIRI & SEVIRI	VIS & IR	3 & 12	1 & 4 km
MTSAT		VIS & IR	5	1 & 4 km

Various techniques and algorithms have been used to develop numerous satellite precipitation products as shown in **Table 2**. These products are prepared based on the data gathered from the sensors.

Table 2: Details of some satellite based data (Tamakar, 2011; Ghaju, 2010)

Product Name	Temporal resolution	Spatial resolution	Data extent	PMW data	IR data	Adjusted by gauge	Data format	One day definition (UTC)
CMORPH	3- hourly	0.25°	2003 - present	yes	yes	No	GIS	00:00-23:50
PERSIANN	6-hourly	0.25°	2001 - present	yes	yes	no	ASCII	00:00-00:00
TRMM 3B42	3-hourly	0.25°	1998 - present	yes	yes	yes	HDF	22:30-22:30
REF-2	Daily	0.1°	2001 - present	yes	yes	yes	GIS	06:00-06:00
GPCP-1DD	Daily	0.25°	1996 - present	Not directly	yes	Not directly	Binary	22:30-22:30
GSMap MVK +	1 -hourly	0.25°	2003-2006	Yes	yes	yes	Binary	00:00-00:00

Many have utilized the satellite precipitation products for a range of applications as these products have advantage against radar and gauged data because of global and spatial coverage. Despite that, the relationship between radiances and precipitation reaching the ground is hard to decide, it is critical to evaluate the extents of errors of the satellite estimates. Additionally, major algorithm inter-comparison projects have revealed that it is difficult to draw quantitative conclusions on the performance of the algorithms because the ground radar and gauges used to evaluate satellite estimates lacks the spatial and temporal coverage and the quality of ground data is also quite questionable (McCollum, 2002).

The following section will briefly highlight some of the findings and conclusions made by some of the researchers and scholars after evaluation of at least one or by inter-comparison of results from two to five products in relation to ground observation. Although availability of literatures on such studies in Bhutan was very limited, studies from Nepal or other similar Himalayan regions were of valuable source as they have similar climate and weather pattern.

2.2 Previous study in same and similar regions

Xue et al. (2013) did the evaluation of the TRMM Multi-satellite Precipitation Analysis product and also explored the improvements and error propagation of the 3B42V7 algorithm against the 3B42V6 using the Coupled Routing and Excess Storage (CREST) hydrologic model in the Wangchhu Basin of Bhutan. The comparison revealed that the 3B42V7 performed better against 3B42V6s underestimation for the whole basin and for grids of 0.25° x 0.25° with a modest enhancement of the correlation coefficients and also improved the occurrence

frequency across the rain intensity spectrum. A paper by Fakhurddin (2015) on development of a flood forecasting model for the same basin in Bhutan by integrating gauged data with satellite rainfall showed that the efficiency of the Weather Research Forecast (WRF) model was poor in the absence of upstream rain gauges. When the same model was run using the TRMM (3B42V6) precipitation data and calibrated parameters it showed that in case the observed data is not available, the TRMM rain data may be used to forecast floods.

Khandu et al. (2015) studied seasonal and interannual skills of the regional gauge-based APHRODITE and near-global satellite based products viz: TRMM, CMORPH and CHIRP over Bhutan against the gauged data for the period 1988-2012. The study showed that both gauge-based and satellite–based precipitation products were able to adequately yield the spatio-temporal patterns of rainfall variability and captured well the extreme precipitation and drought periods with correlations greater than 0.5. It was indicated that the APHRODITE and TRMM (3B42V7) performed relatively similar and better compared to other products.

S.Bajarcharya et al. (2010) conducted a validation of the CPC-RFE 2.0 satellite estimated rainfall in the summer monsoon dominated area of Hindu Kush Himalayan region which includes Bhutan, Afghanistan, Pakistan, Tibet, China, Nepal, Bangladesh and Myanmar. The study indicated that the satellite rainfall estimation overestimates the pre-monsoon rain and in the rain shadow area and also gave maximum negative bias and root mean square in the heavy rain days. It was also pointed out that satellite rainfall estimation and observations have vast difference. Shrestha et al. (2013) also compared the CPC-RFE2.0 satellite rainfall estimation over Nepal and mentioned that the events generally match qualitatively with tendency to substantial underestimation quantitatively. There were also some events of satellite estimates yielding higher than observed values mostly during cooler months and in case of moderate precipitation.

Krakauer et al. (2013) assessed five satellite precipitation products (APHRODITE, GSMaP, CMORPH, PERSIANN-CCS and TRMM) against the observed data over Nepal's mountainous region on a monthly basis. It was concluded that the TRMM product performed better and showed promising use in water resource applications while other satellite products performed substantially low in reproducing station precipitation.

Bajracharya et al. (2015) also evaluated five high- resolution satellite precipitation products (CPC RFE2.0, RFE2.0-Modified, CMORPH, GSMaP and TRMM 3B42) at different spatial and temporal resolutions with observed rainfall data over Brahmaputra Basin and concluded

that the RFE2.0-Modified performed best and TRMM 3B42 showed the second best performance. In the same study it was also demonstrated that there is a potential use of satellite precipitation in data scarce region.

In a study by Islam et al. (2010), comparison of rainfall calculated from the TRMM (3B42V6) with observed rainfall from 15 stations over Nepal was carried out on daily basis during 1998-2007. The study show that the rainfall estimated by TRMM follow a trend similar to that of historical observed data in monthly, seasonal and annual scales with underestimation on many days including overestimation on some days. The study also highlights that the TRMM satellite data can give a better spatial and temporal distribution of rainfall in mountainous region because it observes from the upper side and covers wide area.

Su et al. (2008) used TRMM (3B42V6) and showed that the model simulations using satellite-estimated precipitation for the La Plata basin in South America were able to reproduce the daily flooding events and also represented low flows but with some overestimation of peak flows. The study also says that there is a good agreement with the simulated flows in terms of reproduction of seasonal and interannual stream flow variability and has a potential for hydrologic forecasting in data sparse regions.

Duncan and Biggs (2012) assessed the TRMM (3B42V6) derived precipitation against the ground-based APHRODITE data seasonally from 2001 to 2007 and found that the satellite precipitation did not detect the rainy days, extreme events and the monsoon intensity of rainfall correctly. In the same research it is also mentioned that this satellite product has limited use in agriculture, water resource management and developing mitigation measures due to extreme events in Nepal.

2.3 Selection of Satellite Precipitation product

Findings from most researchers and scholars above reveals that the TRMM (3B42V6) gives a better spatial and temporal distribution of precipitation in the Himalayan region and it follows similar trend to that of historical data in monthly, seasonal and annual scales. Additionally, evaluation made by Xue et al., 2013 in the Wangchhu basin of Bhutan located adjacent to the current study area demonstrated that TRMM-3B42V7 performs better compared to the initial version 3B42V6. Therefore, TRMM-3B42V7 is chosen for further study. A brief over view of the TRMM Satellite rainfall estimation is presented in the following section.

2.4 TRMM-3B42 V7 Rainfall Estimate

Tropical Rainfall Measurement Mission (TRMM) is a combined work between National Aeronautics and Space Administration (NASA) of the United states and the Japnese Aerospace Exploration Agency (JAXA). The TRMM observatory, which housed the first-ever precipitation radar in space, was launched on 27th November 1997 from the Tanegshima Space Centre in Tanegashima, Japan (GES DISC, 2015). The main objective of the TRMM are to measure rainfall energy exchange of subtropical and tropical regions of the earth. The orbit of the TRMM satellite is nearly circular of approximately 350 km with a period of 92.5 minutes and is inclined at an angle of 35° (Kummerow, 1998).

The satellite is housed with multiple instruments such as Precipitation Radar (PR), TRMM Microwave Imager (TMI), and the Visible Infrared Sacnner (VIRS) throughout the TRMM satellite Constellation. The PR is intended to provide a three-dimensional maps of storm structure while the TMI measures the intensity of radiation at different frequencies and the VIRS senses the radiation in the visible infrared wavelengths (GES DISC, 2015). The details of instrument specifications are shown in the **Table 3** below.

Table 3: Specification of the three primary sensors

	PR	**TMI**	**VIRS**
Frequencies	13.8 GHz	10.7, 19.4, 21.3, 37, 85.5 GHz	0.63, 1.6, 10.8,12 µm of wavelength
Resolution	5 km horizontal, 250 m vertical	11 km x 8 km at 37 GHz	2.5 km
Scanning	Cross-track	conical	Cross-track
Swath width	250 km	880 km	830 km

In addition to above three primary instruments, the TRMM satellite also carries two related Earth Observing System (EOS) instruments in the Clouds and Earth's Radiant Energy System and the lightning Imaging System (LIS). **Figure 2.2** below shows a schematic view of the scan geometries of the three primary rainfall sensors.

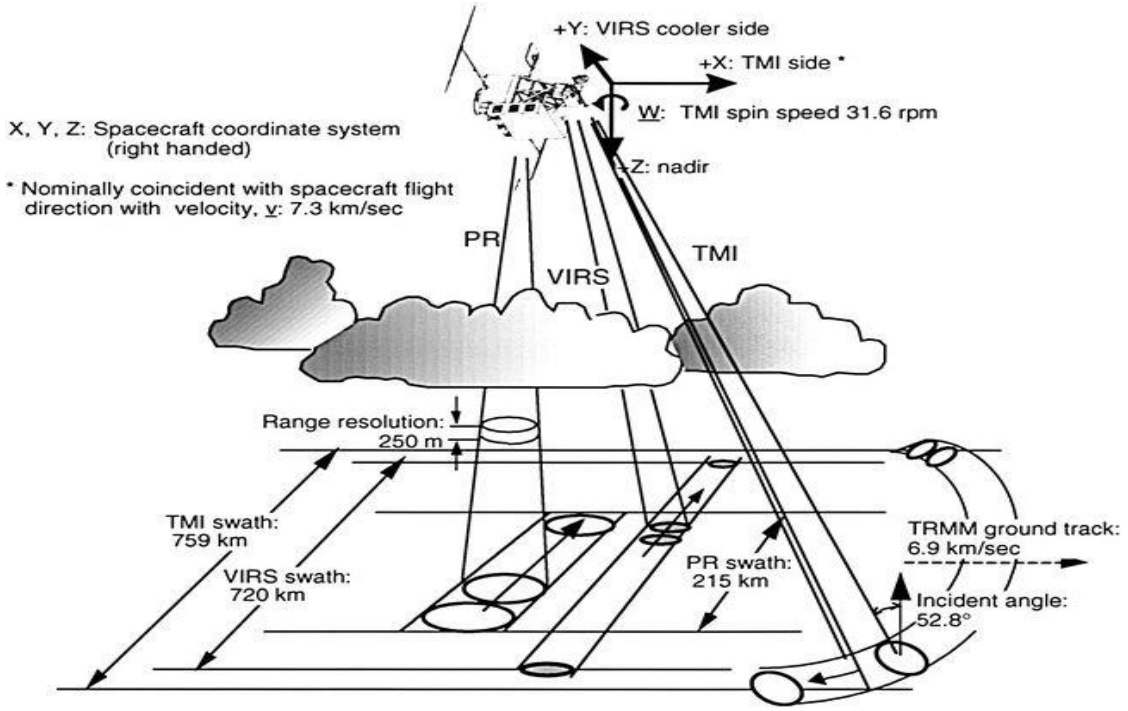


Figure 2.2: Schematic view of the scan geometries of the three TRMM primary rainfall sensors (Kummerow, 1998)

The data set 3B42 comprises of TRMM-adjusted merged infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. The calculation procedure comprises of two separate steps. The initial step utilizes the TRMM VIRS and TMI orbit data (TRMM products 1B01 and 2A12) and the monthly TMI/TRMM Combined Instrument (TCI) adjustment parameters (from TRMM product 3B31) to create monthly IR calibration parameters. The second step utilizes these determined monthly IR calibration parameters to modify the merged -IR precipitation data, which comprises of GMS, GOES-E, GOES-W, Meteosat-7, Meteosat-5, and NOAA-12 data. The last gridded, adjusted merged-IR precipitation and RMS precipitation -error estimates have daily temporal resolution and a 0.25° x 0.25° spatial resolution. Spatial scope reaches out from 50 degrees south to 50 degrees north (http://disc2.nascom.nasa.gov/s4pa/TRMM_L3/TRMM_3B42_daily/doc/TRMM_Readme_v3.pdf).

3 Study Area

3.1 Location and Topography

Kerabari is the last gauging point in the Punatsangchhu basin of Bhutan. The basin shares its northern boundary with Tibetan Autonomous Region of China and the southern boundary with the Indian state of Assam. The total area of the drainage basin is 9747.83 sq.km (all within Bhutan) of which 9626.6 sq.km (98.7%) constitutes the catchment at Kerabari gauging point. The catchment is located between latitude 26.75°N to 28.26°N and longitude 89.32°E to 90.40°E as shown in **Figure 3.1**below.

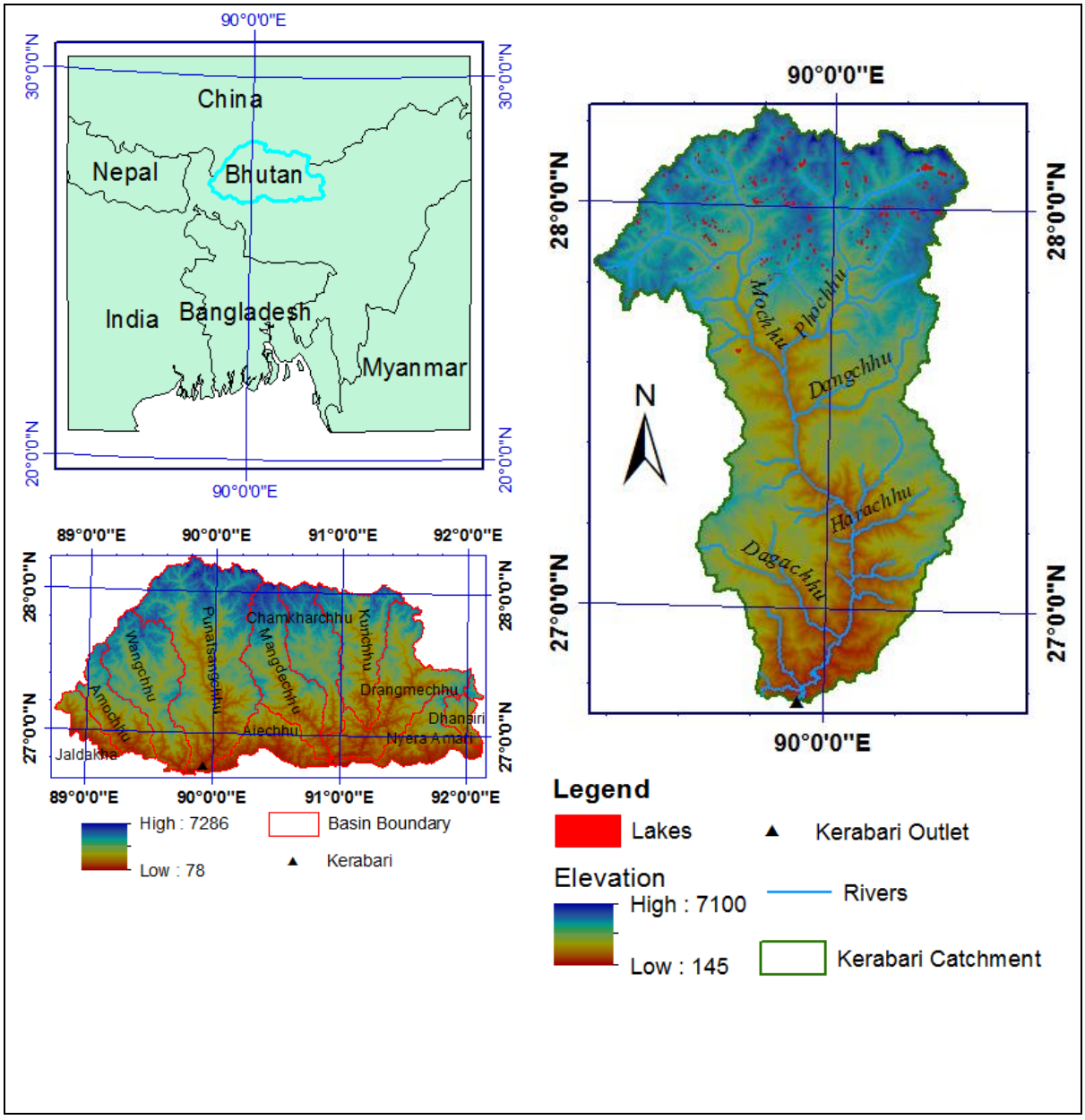


Figure 3.1: The Study Area

The area is predominantly mountainous with very high altitudes, irregularities and steep hills separated by narrow valley. The elevation ranges from 145 masl in the south to 7100 masl in the north and varies by several hundred meters within a short distance. The range of elevations for the entire country and the catchment is shown in the **Figure 3.1** above. In addition, the hypsographic curve in **Figure 3.2** shows the area elevation distribution of the catchment.

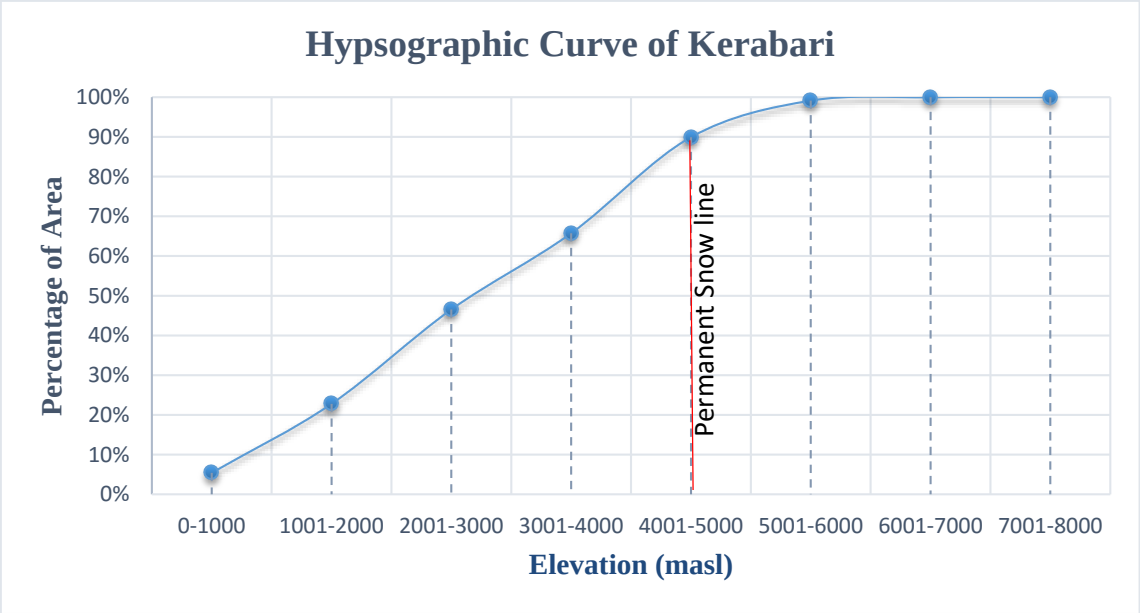


Figure 3.2: Hypsographic Curve of Kerabari

From the hypsographic curve, it can be seen that about 90 percent of the catchment area falls below the permanent snow line at 5000 m.a.s.l. The area increases uniformly by about 20 to 25 percent for every 1000 m increase in elevation until the permanent snow line. Beyond this line up to an elevation of 7100 m only about 10 percent of the catchment is under snow cover throughout the year.

3.2 Climate

Bhutan lies in the equatorial belt of high precipitation. In this belt warm easterly winds from both the hemisphere carrying enormous amount of moisture from tropical oceans converge. One of the most pronounced orographic precipitation formation effects globally occurs in conjunction with the monsoon over the Indian subcontinent (Dingman, 1994). In summer, warm moist southerly wind blows from Bay of Bengal over Bangladesh and India and approaches Bhutan thus producing heavy rains which dominates the climate of Bhutan.

Bhutan is divided into four climate zones: the sub-tropical zone, temperate zone, sub-alpine zone and the alpine zone due to the physical variation in the elevations and orientations of mountains. The subtropical southern foothills experiences mean annual rainfall ranging from 2500 mm to as high as 5500 mm. The mean monthly temperatures vary from 15°C to 30°C throughout the year and is hot and humid during summer and cool in winter. The temperate central valleys or inner hills above the sub-tropical zone are chilly during winter and warm during summer with as much as 1000 mm to 2500 mm of annual rainfall. The mean monthly temperature in this zone varies from 5 to 15 °C in winter and 15 to 30°C during summer. The sub-alpine zone experiences and annual mean temperature around 8° C with rainfall varying from 1000 to 1500 mm annually. The weather in this zone is marked by mist and fog, and cold winds during the summer with light showers and snow in the long winter. The alpine Greater Himalayas experiences year round snowfall and an annual rain of the range 500 mm to 1000 mm. The monsoon begins in June and continues until September with dry periods from November to March (Beldring, 2011). A summary of the climatic zones of Bhutan is given in **Table 4 Error! Reference source not found.**.

Table 4: Climate Zones of Bhutan

Climatic Zones	**Elevation** (masl)	**Mean temperature** winter/summer(°C)	**Annual rainfall** (mm**)**
Subtropical	< 2000	15/30 (monthly)	2500-5500
Temperate	2000 - 3000	5 to 15/ 15 to 30 (monthly)	1000 to 2500
Sub-alpine	3000 - 4000	8 (mean annual)	500 to 1000
Alpine	>4000		<500

Figure 3.3 below shows annual rain fall map of Bhutan and **Figure 3.4** shows the variation in annual precipitation with elevation from 11 selected stations in the catchment. As discussed above the rainfall in the study area is primarily influenced by the local geography, variation in the elevation and the southwest monsoon which blows from the Bay of Bengal. The large difference in elevation from south to north also contributes to the extreme regional variations in climate. The climate also varies between valleys and within valleys, depending on altitude and slope. Rainfall vary within short distances due to orographic lifting and rain shadow effects. Annual precipitation in the catchment varies from 400 to 700 mm in the upstream region of Gasa, 700 to 1500 mm for midstream region where Punakha and Wangdue Phodrang is located and exceptionally more than 2000 mm for steeply inclined topography in mid to downstream areas

of the basin. The climate in the downstream region near the border with India is also categorized to be subtropical with annual precipitation of 3000 to 5000 mm.

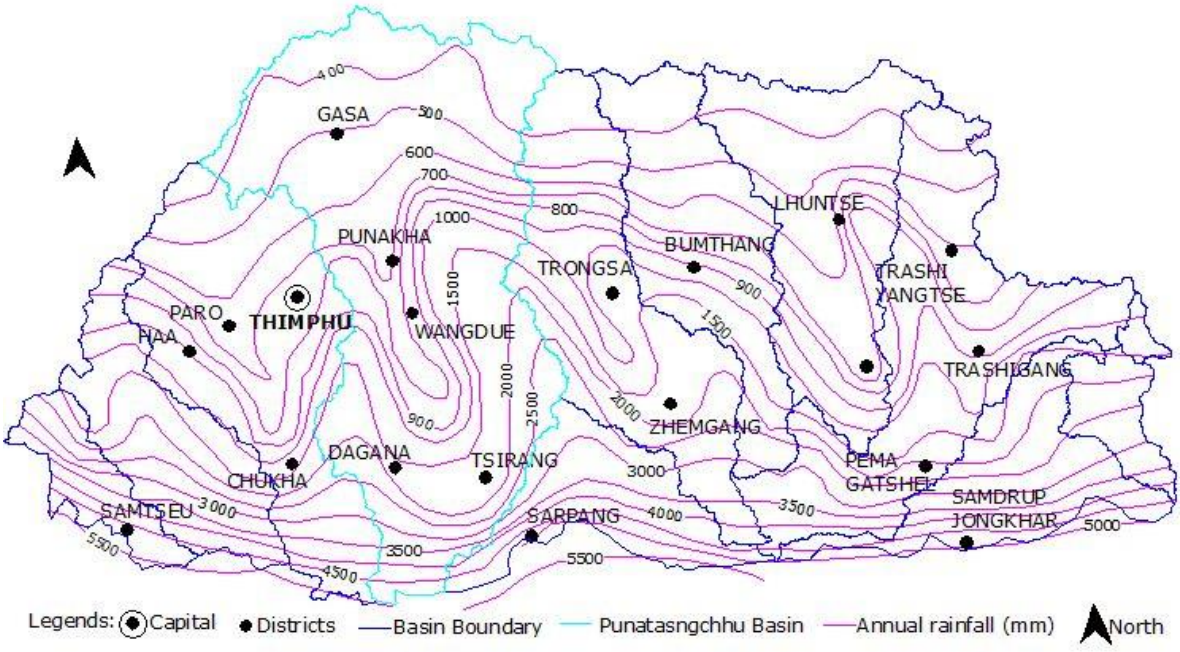


Figure 3.3: Annual rainfall map of Bhutan (CAPSD, 1994)

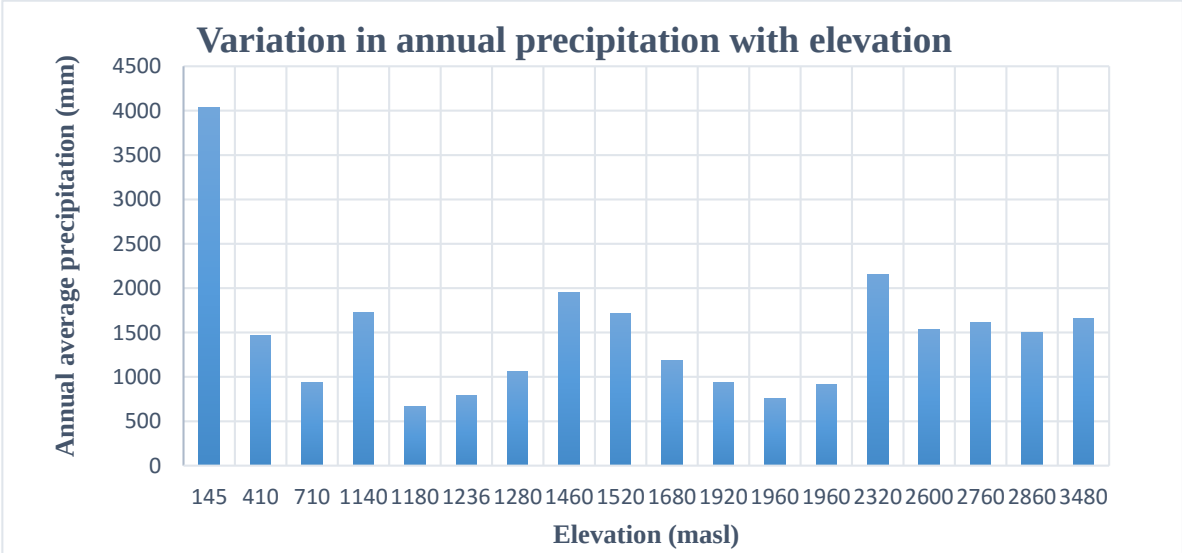


Figure 3.4: Variation in annual precipitation with elevation

The figure shows annual average precipitation from 11 stations in the catchment for the periods 2003-2012.

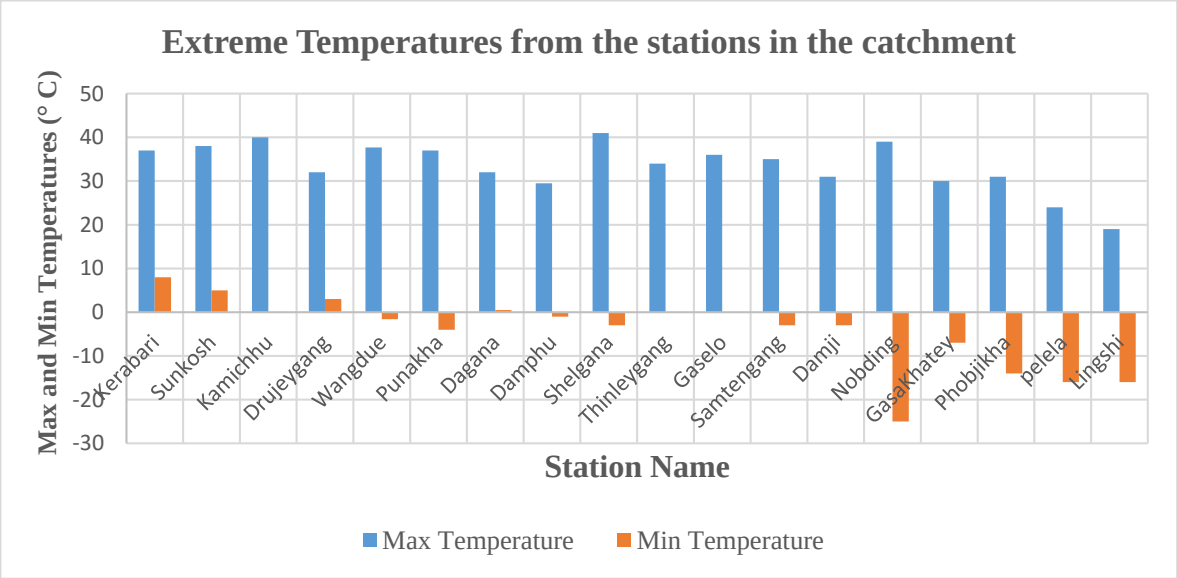


Figure 3.5: Extreme temperatures in the Catchment

The figure shows the extreme temperatures for the period 2003 to 2012 from 18 temperature stations in the catchment.

3.3 Land cover

Kerabari catchment consists of about 98.7% of the total area of the Punatsangchhu basin. As depicted by **Table 5** below there are no reservoirs in the catchment. Forest is the predominant land cover in the catchment with more than 63% of the total catchment area. Agricultural land constitutes 2.29 % and only 0.07 % of the built-up area. There are several hundreds of small and two large (potentially dangerous) lakes at high altitudes in the catchment. The total area covered by lakes is only 0.16% and glaciers and snow cover constitutes 13.76% of the total area. The remaining 20.71% of the area is made up of various types of land such as marshy, meadows, bare soil, landslide areas, screes etc.

Table 5: Details of Land Cover (NSSC, MoFA)

Details of Land Cover	**Percentage of Total Area**
Forest cover	63.01
Reservoirs	0.00
Lakes	0.16
Glaciers and snow cover	13.76
Agriculture	2.29
Built-up areas	0.07
Others	20.71

3.4 The rivers and hydrological regime

Figure 3.7 illustrates the rivers and the some flow gauging stations in the catchment. Mochhu and Pochhu are the two major contributing rivers that originates from the high altitude alpine region of the basin and confluences at an altitude of 1200 masl in Punakha District. From this point onwards the river is called as Punatsangchhu. Downstream of the confluence several small tributaries and rivers like Dangchhu, Harachhu, and Dagachhu also joins to the main river before it exits the Bhutan –India border. The main river Punatasngchhu is the longest river in the country with a length of about 250 km up to the Indian border in Assam. The river is called as Sunkosh in the Indian side of the border which finally drains to the River Brahmaputra in India. A tentative profile of the river along Pochhu and Mochhu is worked out and presented in **Figure 3.6** below. From the two figures it can be brought out that the physical features exist beyond 2000 meters above the river bed. The river along Mochhu and Pochhu has a gradient of 32 m and 28 m per each kilometer respectively.

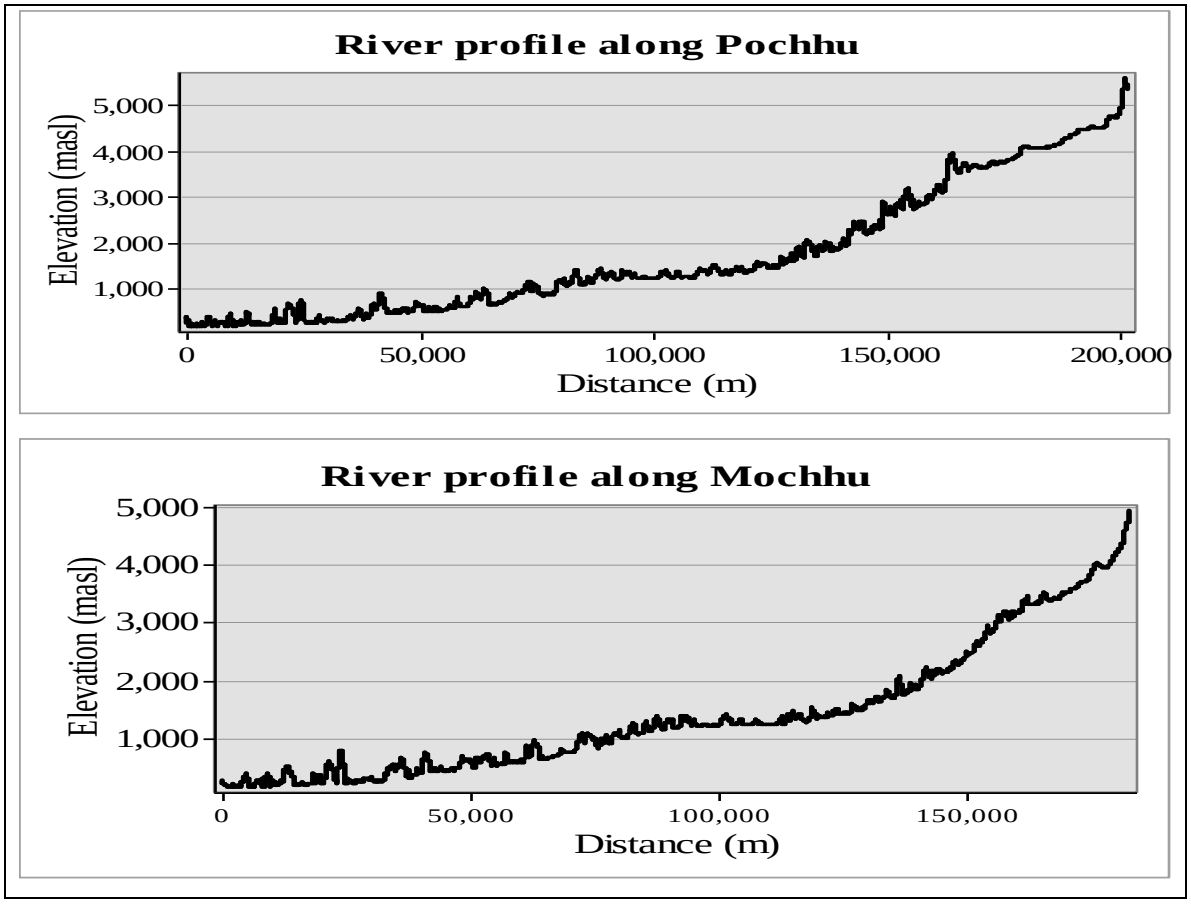


Figure 3.6: River profiles

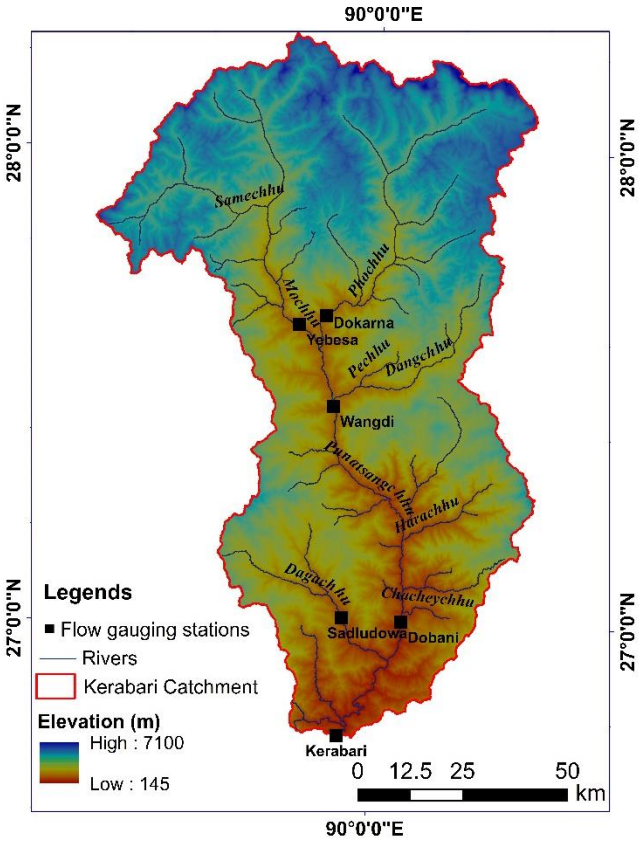


Figure 3.7: Rivers and Gauging stations

The yearly cycle of observed streamflow from three gauging stations in the catchment is shown in **Figure 3.8** below. The hydrological regime of the catchment is described by low stream flow in the winter within periods with little rainfall and low temperatures bringing about accumulation of snow at high altitudes, and high stream flow during summer created by monsoon precipitation and melting of snow (Beldring, 2011).

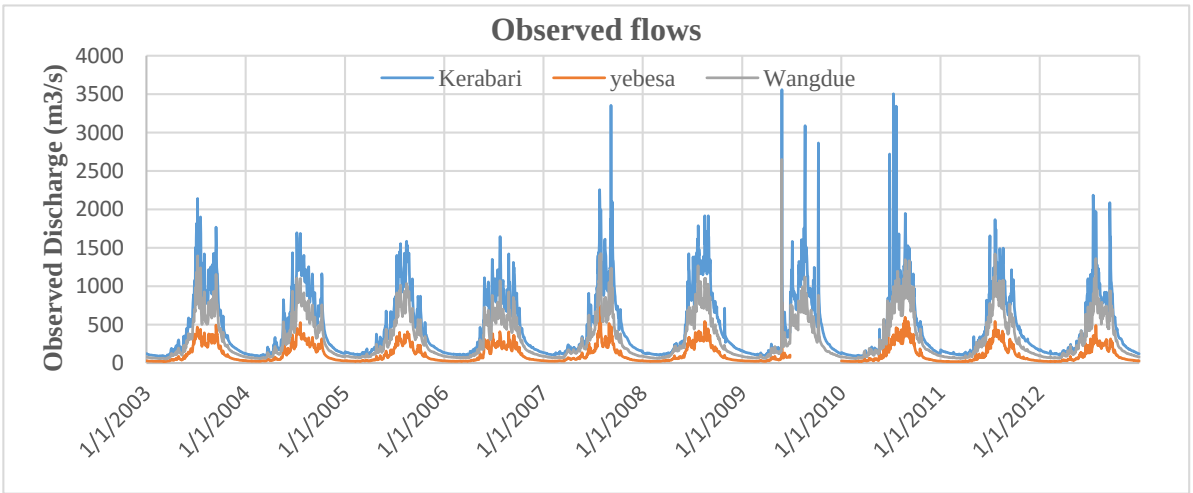


Figure 3.8: The yearly cycle of observed stream flows

4 Data and Methodology

4.1 Observed data and source

The observed hydro-meteorological data are collected from the Department of Hydromet Services (DHMS), Ministry of Economic Affairs of the Royal Government of Bhutan. All together there are 93 meteorological stations and 25 flow gauging stations installed by DHMS in the country. From a total of 93, 22 meteorological stations (**Table 6)** and 6 flow gauging stations are (**Table 8)** within the catchment in addition to 19 nearby meteorological stations (**Table 7**). The location of all these meteorological stations are shown in **Figure 4.1** and that of flow gauging stations are shown in **Figure 3.7.**

Table 6: Meteorological stations within the catchment

Sl.no	Station_Na	Dzongkhag	Lat (°)	Long(°)	Elevation (m)	Available since	Missing data between 2003-2012 (days)
G6	Daga Dzong	Dagana	27.07	89.87	1460	1996	90
G42	Damji	Gasa	27.81	89.73	2320	2005	1015
G30	Damphu	Tsirang	27.00	90.12	1520	1996	0
G7	Drujaygang	Dagana	26.97	90.04	1140	1990	67
G10	Gasakhatey	Gasa	27.90	89.72	2760	2003	175
G31	Gaselo	Wangdue	27.42	89.89	1960	1996	0
G43	Hetsothangka	Wangdue	27.45	89.90	1360	Data not available	
G32	Kamichhu	Wangdue	27.25	90.04	710	1996	245
G38	Lingshi	Thimphu	27.85	89.43	4080	2004	214
G46	Mendrelgang	Tsirang	26.96	90.12	890	Data not available	
G33	Nobding	Wangdue	27.55	90.15	2600	1996	882
G35	Phobjikha	Wangdue	27.47	90.18	2860	1996	92
G14	Punakha	Punakha	27.58	89.87	1236	1990	0
G41	Samdingkha	Punakha	27.63	89.88	1560	Data not available	
G36	Samtengang	Wangdue	27.52	90.00	1960	1996	0
G8	Sankosh	Dagana	27.02	90.07	410	1990	0
G15	Shelgana	Punakha	27.61	89.85	1680	1996	337
G39	Thinleygang	Thimphu	27.51	89.81	2265	2004	25
G9	Tsangkha	Dagana	27.03	90.05	1270	Data not available	
G45	Tsirabg Toe	Tsirang	27.06	90.10	1280	Data not available	
G37	WangdueRNRC	Wangdue	27.49	89.90	1180	1990	0
G40	Yebesa	Punakha	27.63	89.76	2040	2007	1462

Table 7: Nearby Meteorological stations

Sl.no	Station_Na	Dzongkhag	Lat (°)	Long(°)	Elevation (m)	Available since	Missing data between 2003-2012 (days)
G11	Betikha	Paro	27.25	89.42	2660	1996	733
G16	Bhur	Sarpang	26.90	90.43	375	1996	0
G24	Bjizam	Trongsa	27.52	90.46	1840	1995	0
G1	Chapcha	Chukha	27.20	89.55	2450	2000	92
G25	Chendebji	Trongsa	27.50	90.38	2660	1996	30
G2	Chukha	Chukha	27.07	89.57	1600	1990	275
G5	Darla	Chukha	26.88	89.57	1745	1996	601
G12	Drukgyel Dzong	Paro	27.50	89.33	2547	1996	0
G3	Gedu	Chukha	26.91	89.53	1980	2005	122
G21	Gidakom	Thimphu	27.38	89.57	2210	1996	0
G26	Kuengarabten	Trongsa	27.41	90.52	1780	1996	
G27	Langthel	Trongsa	27.37	90.57	1150	1996	31
G22	MoEA Complex(A)	Thimphu	27.47	89.64	2380	1996	0
G13	Paro (DSC)	Paro	27.38	89.42	2406	1996	0
G4	Phuntsholing	Chukha	26.86	89.37	220	1996	184
G17	Sarpang	Sarpang	26.86	90.27	330	1996	393
G23	Semtokha(A)	Thimphu	27.44	89.68	2310	1996	0
G18	Surey	Sarpang	27.02	90.54	1060	1996	0
G28	Trongsa	Trongsa	27.50	90.51	2120	1996	0

Table 8: Flow Gauging stations in the catchment

Sl.no	**Station**	**River/District**	**Elevation (masl)**	**Catchment area (sq.km)**	**Lat (°)**	**Long (°)**
1	Sadudowa	Dagachhu/Dagana	780	671	27.02	89.92
2	Sankosh	Sankosh/ Dagana	265	8593	27.01	90.07
3	Wangdue	Punatsangchhu/ Wangdue	1190	6271	27.46	89.90
4	Yebesa	Mochhu/ Punakha	1230	2320	27.63	89.82
5	Dokarna	Phochhu/ Punakha	1290	2295	27.65	89.88
6	Kerabari	Kerabari/ Dagana	145	9626.6	26.77	89.93

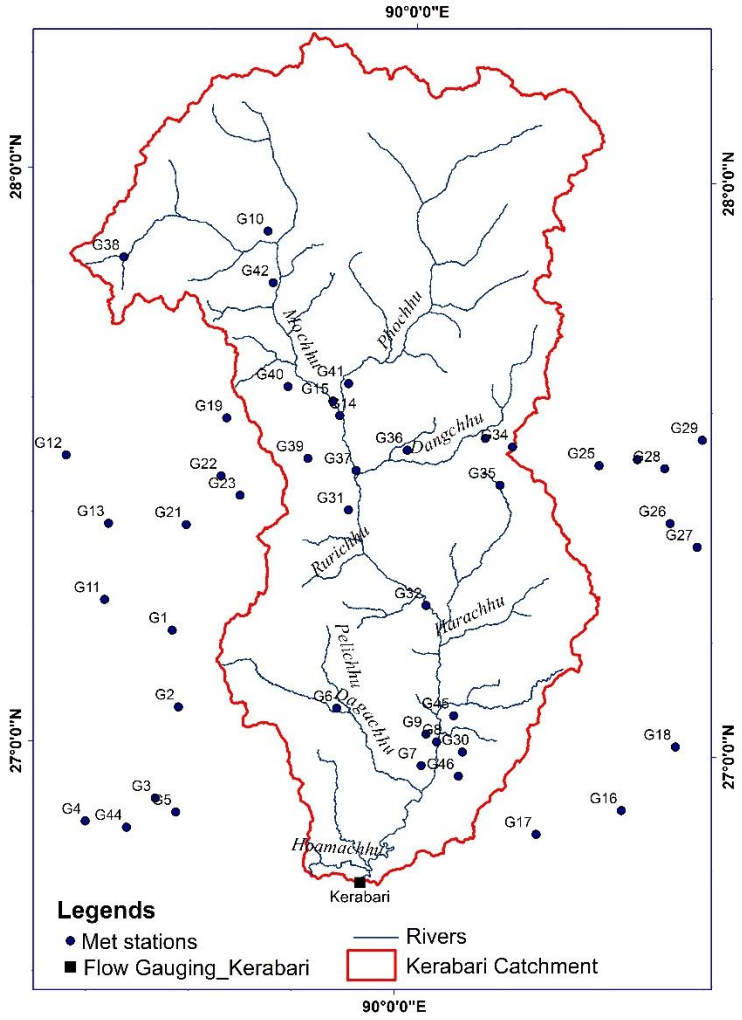


Figure 4.1: Meteorological stations in and around the catchment

4.2 Data coverage

As can be seen in **Figure 4.1** the spatial distribution of precipitation gauges are not uniform in the basin. Most of meteorological stations are clustered in lower altitudes in the valleys close to the populous areas and does not cover the headwater areas of the basin. Therefore, the available data does not provide representative information on the climatic conditions over the headwater regions of the basin. Majority of the stations in and around the basin have data since 1996 but only 6 stations within the catchment and 10 stations around the catchment have data without any gap. Rest of the stations have gaps ranging from 25 days to as high as 1462 days in the study period.

4.3 Selection of stations and study period

Out of 22 stations within the catchment it was not possible to get data from 5 stations. From the remaining, only 11(highlighted in **Table 6)** station were selected to be used in rainfall runoff modelling with SHyFT. The selection of station was purely based on the availability of data and minimum missing days with an objective of achieving relatively good spatial coverage. The period of study (2003 to 2012) was based on the availability of data during this period with an assumption that from a period of 10 years, first 5 years' data will be sufficient to calibrate the model and remaining 5 years will be used to validate the model as per general practice in modelling.

In order to compare gridded satellite data with ground measured precipitation and to evaluate the use of remote sensed data for model applications, 19 nearby meteorological stations (**Table 7**) were also identified in addition to the stations within the catchment. Only those stations were used when the ratio of the data available days between gauge and satellite was 0.8 and higher. In this regard, out of 36 gauges presented above, gauges G33, G40 and G42 were not used in the analysis.

4.4 Gauge data processing and quality control

4.4.1 Precipitation

Once the data is collected it is very important to inspect the quality. First of all, visual inspection was carried out to check the completeness of the time series and then all unphysical values like spikes, negatives etc.., were removed from the series. Wherever the gauge precipitation showed suspicious values in some of the days the corresponding value of flow records at flow gauges were checked vice versa to see if similar trend was followed. It was easy to see such problems in the data series by making a plot against time.

Some of the gaps in the time series were filled by calculating the correlation with nearby gauges and using the Station-Average Method when the annual precipitation value at each gauge differ by less than 10% or by using the Normal-Ratio Method when the annual precipitation at the gauges in the region differed by more than 10 %. As SHyFT model was used which will by itself do the interpolation of missing values using the Inverse Distance Weighing (IDW) Method internally in the model, manual filling of gaps was only considered if a stations miss records for a short period. After filling the gaps in the data, a trend analysis was conducted and accumulation plots were prepared to check if there is any inconsistency or inhomogeneity in data.

For example **Figure 4.2** shows a simple trend on mean monthly rainfall from some selected stations. As already discussed, the monsoon in the study area begins in June and continues until September with dry periods from November to March, similar trend is shown by the figure for all stations. In addition, **Table 9** show that percentage distribution of precipitation on seasonal basis is similar between the stations with as high as 59% during summer, 21% during autumn, 17% during spring and least of 3% during winter on average. Later a double mass curves (**Figure 4.3**) were also plotted to check for inhomogeneity and changes in the station location if any, but no change in the location of stations and inhomogeneity in data were found. Similar analysis was conducted on other precipitation stations which were used to validate the satellite data and no major problem were detected in the data sets.

Overall the quality of data can be described as sufficient quality and are reliable for further analysis except that, some stations have big gaps and the observed precipitation data set from Lingshi (G38) with annual records in the range of 5000mm to 12800mm is not at all realistic because the highest rainfall receiving stations in the country are Bhur and Sipsu (4500mm - 5500 mm) located in the southern foothills and faces direct monsoon without much barrier. This can also be cross checked from the annual rainfall map of the country presented in **Figure 3.3** which show that Lingshi is located in the annual rainfall range of 400 to 500 mm.

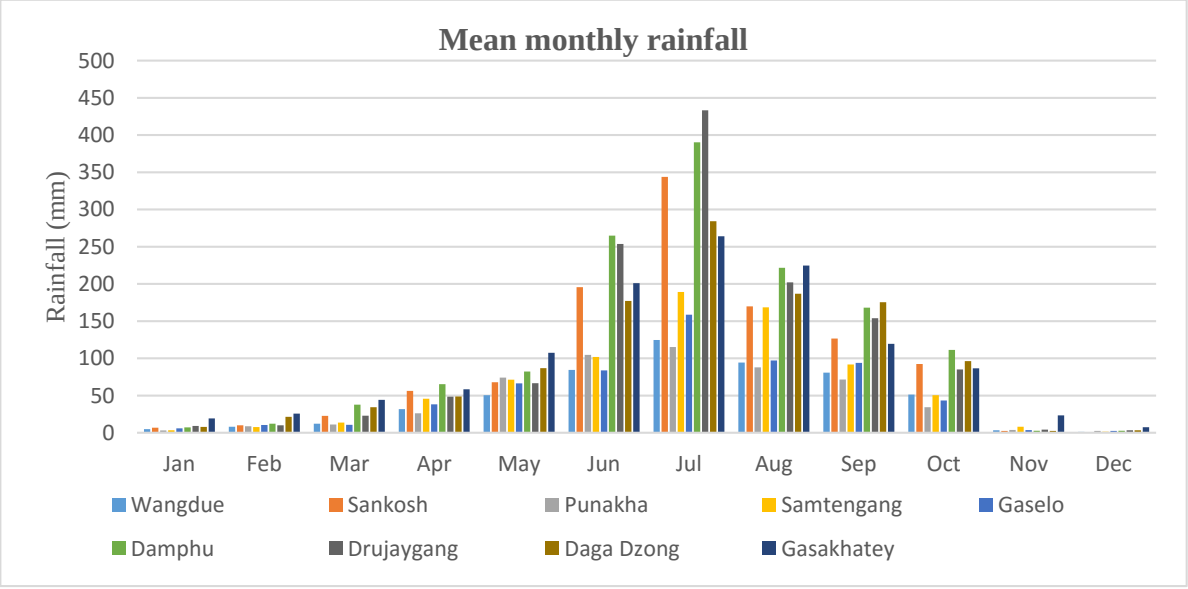


Figure 4.2: Mean monthly rainfall from 9 stations in the catchment

Table 9: Seasonal distribution of precipitation in the basin

Stations	Seasonal distribution (%)			
	Dec-Feb	Mar-May	Jun-Aug	Sept-Nov
Wangdue	3	17	55	25
Sankosh	2	13	65	20
Punakha	3	21	57	20
Samtengang	2	17	61	20
Gaselo	3	19	55	23
Damphu	2	14	64	21
Shelgana	2	18	61	19
Daga Dzong	3	15	58	24
Gasakhatey	4	18	58	19

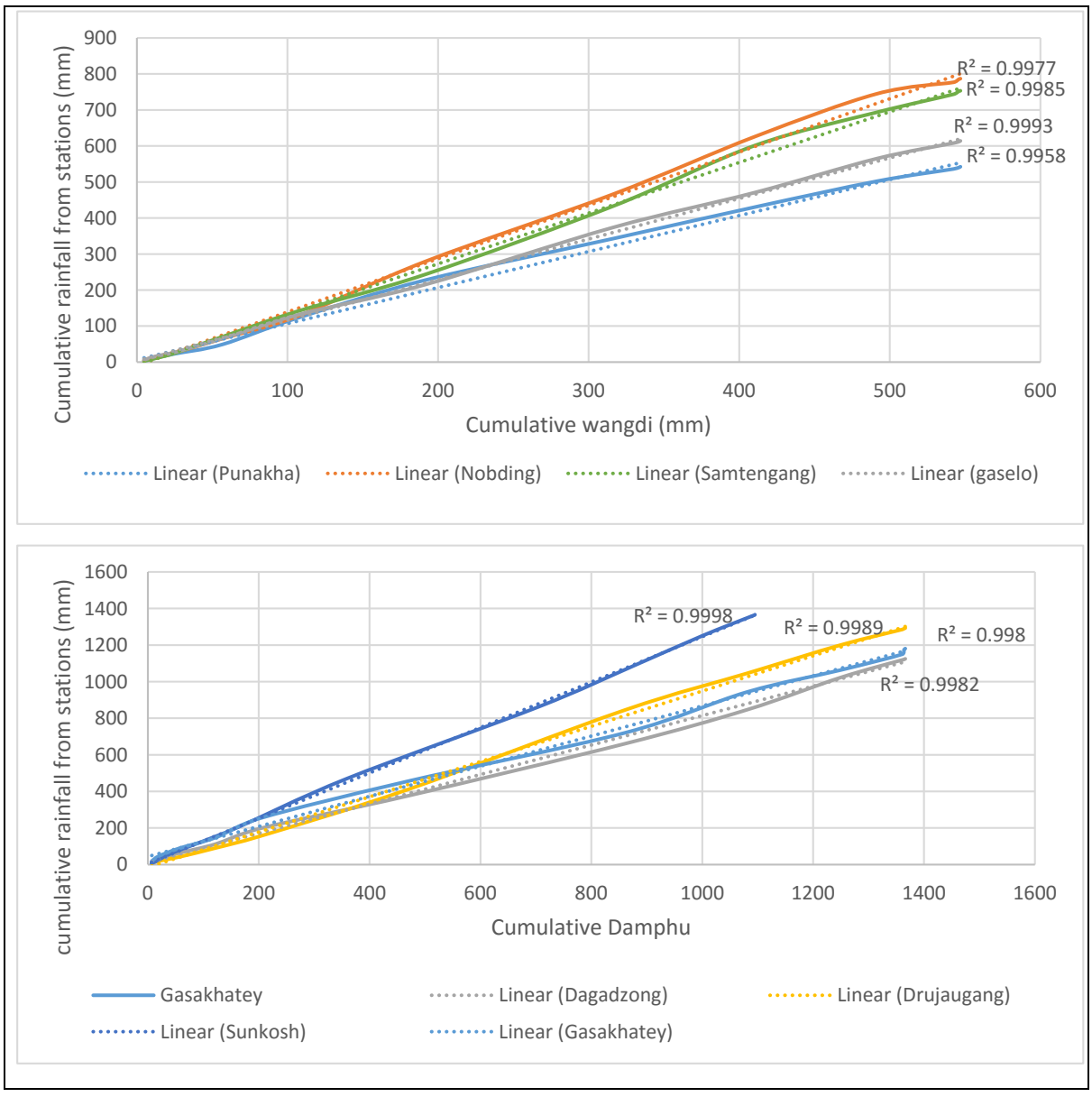


Figure 4.3: Double mass curves

4.4.2 Temperature

Analysis was conducted on the temperature data from 6 stations in the catchment and here also no major problems were detected besides some gaps in the data. Since the current setup of SHyFT does not accept missing values in temperature dataset, it was required to be filled manually using the equation shown below.

$$T_3 = T_2 + (T_{3avg} - T_{2avg}) \quad \text{(Rinde, 2015, lecture notes)} \qquad (1)$$

Where, T_3 is the calculated daily mean temperature for stations missing data, T_2 is the daily mean temperature from best correlated known gauging station, T_{3avg} is the average daily mean temperature from ungauged station and T_{2avg} is the average daily mean temperature from best correlated known gauging station.

Figure 4.4 shows the distribution of mean monthly temperatures from 6 stations located in the central and southern parts of the basin. It can be seen that the distribution is consistent with maximum temperature during summer and mimimum during winter. Small filling of gaps was required to be done for stations: Samtengang, Gaselo and Damphu. Temperature data from stations located in higher altitudes were not used because they had too much gaps.

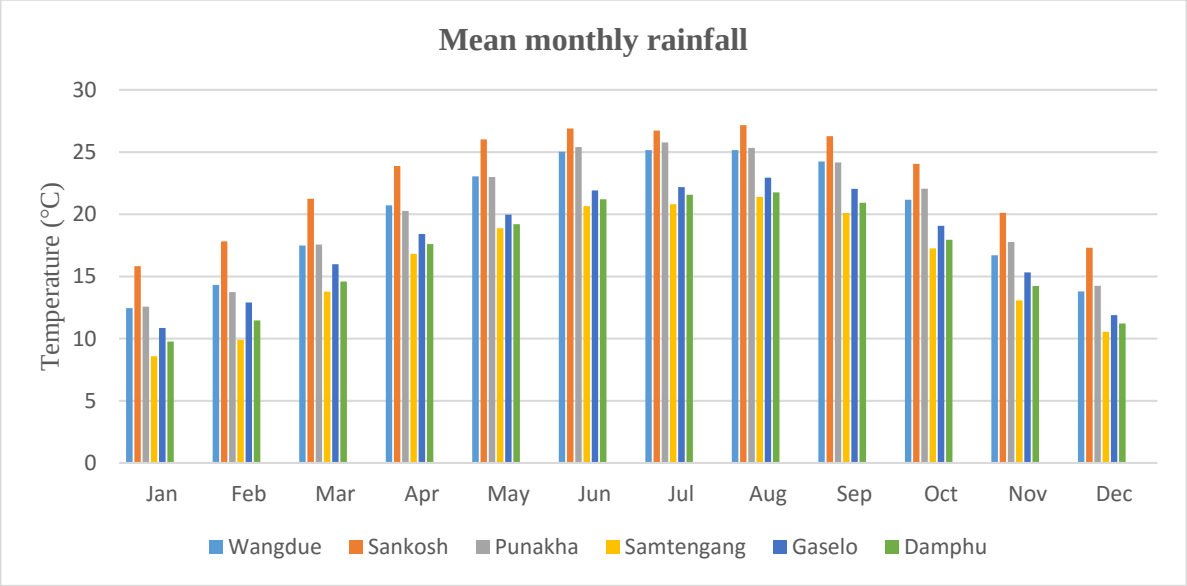


Figure 4.4: Mean monthly temperature from selected stations

4.4.3 Wind speed, relative humidity and radiation

The data on wind speed and relative humidity are taken from the meteorological stations at Wangdi, Punakha and Gaselo. The data on solar radiation was not available directly and therefore it was calculated manually for two stations (Damphu and Wangdue) using the data on latitude, time of the year and sunshine from the stations. Following relations from the snow melt model used for the course TVM4105 Hydrology (2014) was used for calculations.

$$Rs\left(\frac{w}{m^2}\right) = as + bs \times \frac{n}{N} \times Ra(\frac{w}{m^2}) \quad (2)$$

$$R_a\left(\frac{\frac{MJ}{m^2}}{day}\right) = \frac{24 * 60}{\pi} \times [Gsc \times dr(\omega_s \times Sin(\emptyset) \times Sin(\delta) + Cos(\emptyset) \times Cos(\delta) \times Sin(\omega_s)] \quad (3)$$

$$dr(radians) = 1 + 0.33Cos\left(\frac{2\pi J}{365}\right) \quad (4)$$

$$\omega_s(radians) = Acos[-\tan(\emptyset) \times \tan(\delta)] \quad (5)$$

$$\delta(radians) = 0.409Sin\left(\frac{2\pi J}{365} - 1.39\right) \quad (6)$$

$$N = \frac{24 \times \omega_s}{\pi} \quad (7)$$

Where:

Rs = Solar radiation

Ra = Extraterrestrial radiation

J = The time of the year

n = Actual duration of sunshine hours

dr = Inverse relative distance (earth-sun)/eccentricity

δ = Solar declination

ω_s = Sunset hour angle

N = Maximum possible sunshine hours or daylight hours

4.4.4 Hydrology

Before the available hydrological data could be used for calibrating the model for Kerabari catchment it was important to establish the consistencies of the observed flows. In this regard

the hydrographs from three catchments in the basin were compared and a double mass curve was plotted. The observed hydrographs and the double mass curve are shown in **Figure 4.5**. The comparison of hydrographs revealed that the observed flow at Kerabari and other gauging stations have similar variation pattern. The extreme events of 27th October 2009 was due to the cyclone in Indian ocean which hit all over Bhutan with heavy rainfall and flashfloods causing rivers to swell. This event was also recorded by upstream gauging satations. The event of 08/10/2009 was caused by heavy precipitation (6mm/hr) in the southern areas of Kerabari when precipitation was not recorded in the upstream areas. The double mass curve shown below infers that the observed data are consistent.

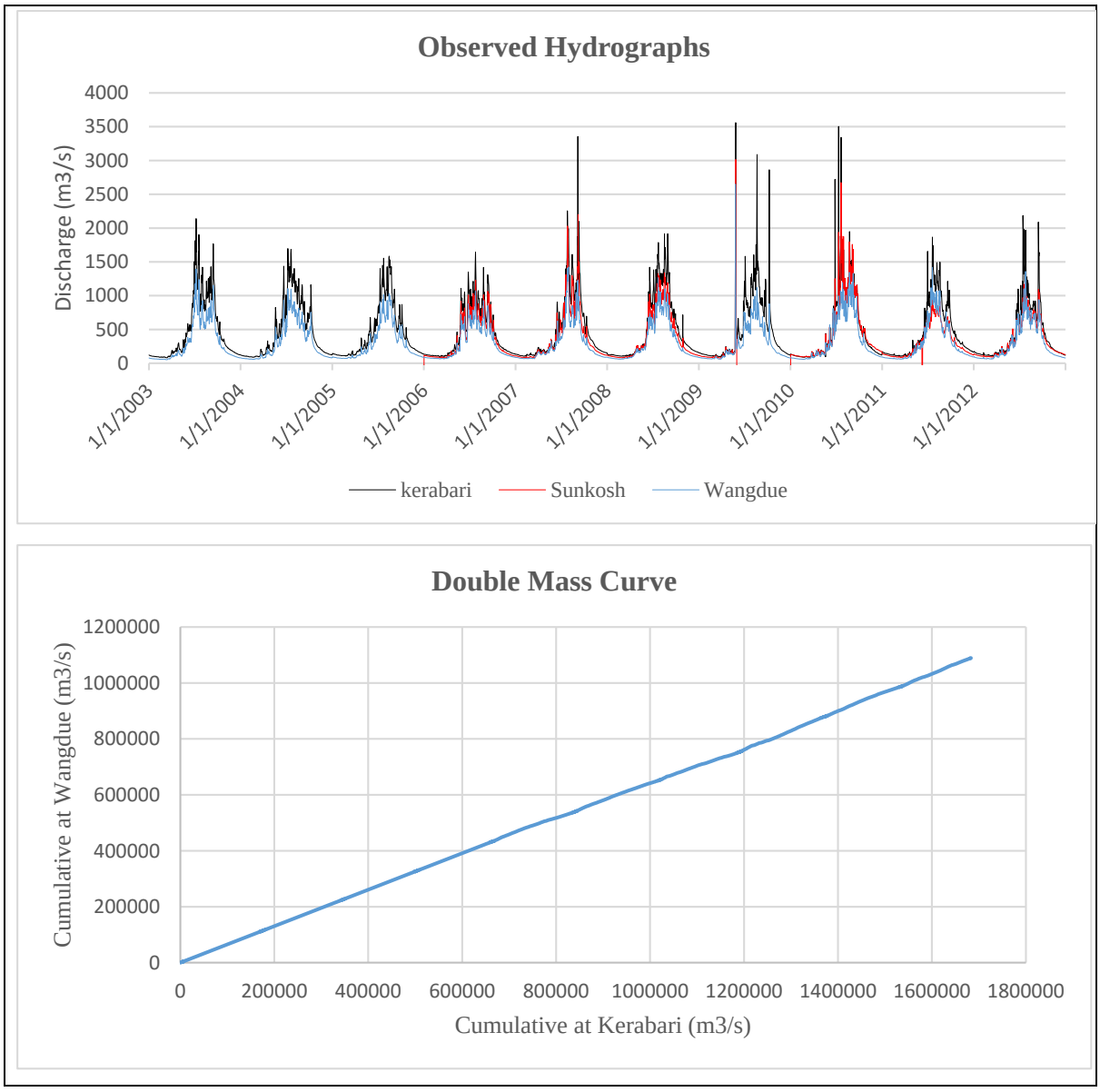


Figure 4.5: Observed hydrographs and Double mass curve

4.5 Satellite Precipitation Estimation product and source

As discussed in earlier chapters there are numerous precipitation products available freely over the internet. These products are managed and released by different organizations through their servers. One such product selected for use in this study is the TRMM 3B42 version 7 (TRMM-3B42V7) product released by Goddard Earth Sciences Data and Information Services Center, NASA.

The TRMM 3B42V7 provides precipitation estimates from 1998 until current date and in this study daily data from 2003 to 2012 is being used since adequate ground observation is available for this period. Mirador (http://mirador.gsfc.nasa.gov/) is used to locate and download all of the TRMM data for the period described above in NetCDF format. Originally the file comes in HDF (Hirerarchial Data Format) format with a temporal resolution of three hours. This means that there are 8 three hourly satellite observations daily and 2920 to 2928 observation yearly. To analyze 10 years data we will have nearly 30,000 files of 2.2MB each. For simplification in this study, daily data from 2003 to 2012 in NetCDF format will be used. Error! Reference source not found. **Error! Reference source not found.** exhibits one day precipitation estimate by TRMM satellite product in the Kerabari catchment area and **Table 10** shows some salient features of the TRMM 3B42V7 daily NetCDF data.

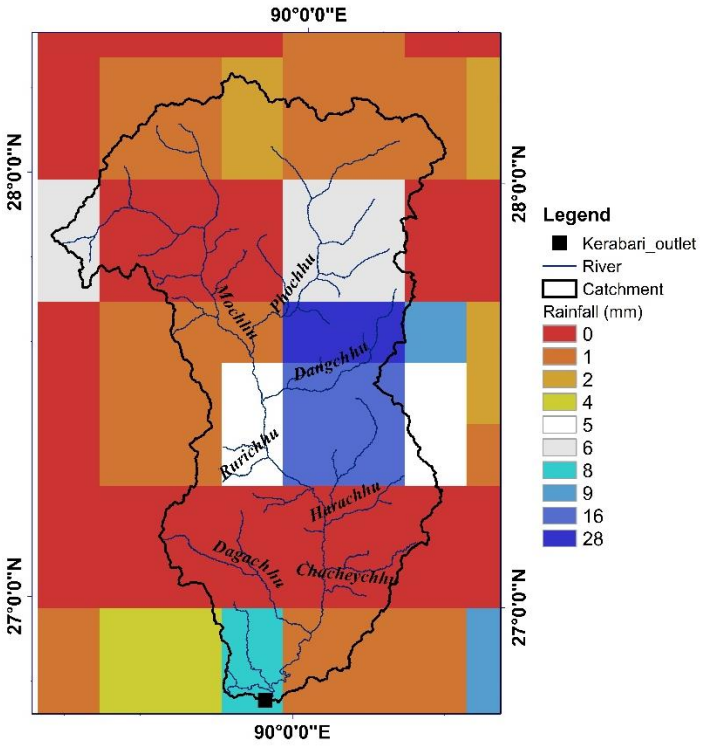


Figure 4.6: TRMM 3B42 daily estimate

Table 10: Salient features of TRMM-3B42V7

Period for which data is available	1998 to present
Geographic Coverage	Latitude: 50°S to 50°N, Longitude:180°W to 180°E
Temporal resolution	Daily
Spatial resolution	0.25° x 0.25° (25km x 25 km)
Grid size	400 x 1440 pixels
Average file size	2.2MB
Projection	Geographic WGS 1984
File type	NetCDF
Unit of measurement	mm
Fill value	-999.9f

4.6 Satellite Data Processing

Satellite data processing is the quite complex and time consuming process. As there are thousands of satellite files the data processing is not an easy job. Researchers in the past have written some scripts (e.g. Python scripts) which can make the processing job easy. The scripts can be used to clip the data for a specific area or catchment of interest, to perform statistical computations and to compare satellite precipitations with that of ground gauges and make corrections wherever necessary. However, to use these scripts one requires to have a sound knowledge of programming in such platforms. Optionally programs like Panoply of NASA, is another great tool if one has patience. Using this program, global time series of daily precipitation can be prepared and taken to Microsoft Excel to join and clip the time series of the area of interest as is done in this study. The extracted data are then processed and compared with ground observations before it could be used for runoff modelling.

Figure 4.7 Error! Reference source not found. is an illustration of how the satellite precipitations is extracted and used in this study. As shown in the figure a layer of satellite grid (27.5 x 27.5 km2) is placed over the catchment with one point in the center of each grid/pixel. The satellite precipitation at each point is computed as an average precipitation for that pixel

from daily estimates demonstrated in Error! Reference source not found. above. A total of 24 points (S1, S2…S24) are used to represent the catchments spatial variability in precipitation.

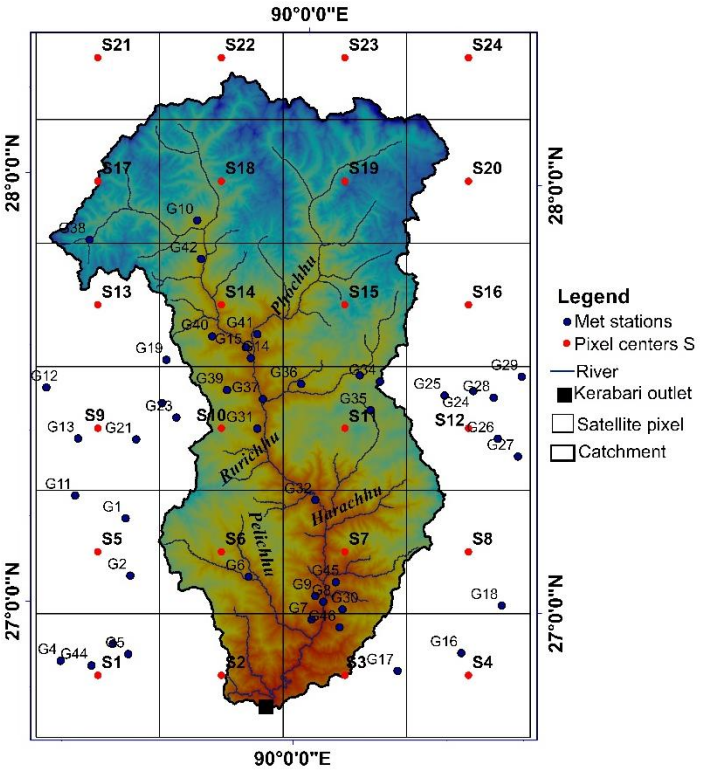


Figure 4.7: Satellite layer with an underlying layer of rain gauges

4.7 Verification of satellite precipitation estimates

Like any other observed data, it is vital to comprehend the precision and restrictions of satellite precipitation estimates. This is can be done by comparing the satellite estimates against the point observations from the rain gauges and radar estimates by visual evaluation of plots and using statistical methods. Standard statistical measures such as root mean square difference, bias and correlation are commonly used statistical methods for evaluating the errors in satellite precipitation estimates (Ebert, 2007). However, other methods are also employed in this study and are discussed in the following section.

4.8 Statistical and Graphical Methods

The comparison of the two data sets was done on daily, monthly and annual basis. While doing the statistical computations, all negative and missing values were removed from both the satellite and gauge data sets. In addition, only those gauging stations are considered for

comparison with the satellite estimates when the ratio of the gauge precipitation days and satellite precipitation days are more than or equal to 0.8. **Figure 4.7** above shows the locations of precipitation gauges in and around the catchment area. From these gauges an average observed precipitation is computed for each pixel center as a representative value for that pixel. In the absence of gauges in any pixel, the average value of precipitation is computed using the IDW from the nearby stations. However, it was not possible to do the validation of satellite estimates for the top six pixels (S15-S24) because there are no gauges in this area and interpolated values might not portray the actual scenario. However, trend in the variation pattern will be studied from the analysis of lower 15 pixels and correction of biases if required will be done accordingly for the remaining six pixels.

In addition to the three most widely used statistical validation methods stated above, other methods such as Scatter Plot, Nash-Sutcliffe Coefficient of Efficiency, Normalized Accumulated Difference, Mean Absolute Difference, Mean Relative Absolute Difference, Satellite Conditional Probability of Detection and Gauge Conditional Probability of Detection are also used for the comparison of the satellite estimates and gauge records. A brief description of such methods used in the study are given below and the related formulas are presented in **Table 11.**

1. **Scatter Plot**

Gauge precipitation and satellite precipitations can be plotted in x and y axis. If the plot is well concentrated around 45° line, then the two data sets are said have a strong correlation without any bias.

2. **Correlation coefficient (RR)**

The coefficient is the measure of linear correlation or dependence between two variables x and y. The coefficient may vary between -1 to +1. A correlation value of +1 means that the data points lie on a line for which increase in x results in increase in y. A correlation value of -1 means that the data points lie on the line and increase in x will result in decrease in y. If the correlation value is 0 then it means that there is no correlation between the two variables.

3. **Nash-Sutcliffe Coefficient of Efficiency (R2)**

The Nash-Sutcliffe Coefficient of Efficiency is a dimension less indicator which is used to assess the predictive power of hydrological models. The efficiency (R2) can vary from $-\infty$ to 1. An efficiency of 1 compares to a flawless match of satellite precipitation to the gage

information. An efficiency equal to 0 demonstrates that the satellite precipitations are as exact as the mean of the gauge data, though an effectiveness less than 0 happens when the gauge mean is a better indicator than the satellite determined precipitations. The coefficient is sensitive to extreme values and may yield imperfect results when the data set contains extreme events of precipitations.

4. **Normalized Accumulated Difference (NAD)**

NAD gives the percentile deviation of satellite precipitation with respect to gauge precipitation

5. **Root Mean Square Difference (RMSD)**

The root –mean-square difference is a normally used to measure the differences between the satellite estimates and observed rainfall. It represents the standard deviation of the differences between satellite estimates and point observed precipitations. The value of RMSD ranges from 0 to ∞. Lower value of RMSD implies that the satellite data is better correlated with the gauge data and vice versa. RMSD equal to 0 is a perfect match. The RMSD does not indicate if the data are over or under estimated

6. **Mean Absolute Difference (MAD)**

MAD is used to evaluate the average magnitude of error between the satellite and gauge precipitation. MAD is expressed in mm and it can range from 0 to ∞ without any direction of deviation. If MAD is 0, it means that there is a perfect match between the satellite and gauge data sets.

7. **Mean Relative Absolute Difference (MRAD)**

The average magnitude of error in satellite precipitation with respect to the gauge data can be evaluated using the MRAD. The value of MRAD ranges from 0 to ∞ and lower measure of MRAD implies that the satellite estimates are good.

8. **Estimation Bias (EB)**

The estimation bias is the normalized difference between the satellite and gauge precipitation data sets over a long period of time. The satellite estimate is said to be unbiased if EB is equal to 0. The Negative EB shows under estimation in the satellite data sets and positive EB implies over estimation in the satellite data.

9. **Satellite Conditional Probability of Detection (CPOD_S)**

CPOD_S is the measure of probability that precipitation recorded by a gauge is detected by the satellite. The value of CPOD_S ranges from less than 1 to 1. If this value is equal to 1 then it means that satellite data set has well detected the observed precipitation.

10. Gauge Conditional Probability of Detection (CPOD_G)

CPOD_G is the measure of probability that precipitation recorded by a satellite is recorded by the gauge. The value of CPOD_G ranges from less than 1 to 1. If this value is equal to 1 then it means that ground gauges have well recorded the precipitation detected by the satellite.

Table 11: Statistical formulas (Tamarkar, 2011; Ghaju,2010)

$RR = \frac{Covariance\ (SP, GP)}{\sigma_{GP} * \sigma_{SP}}$	(8)
$R2 = 1 - \frac{\sum_1^N (SP_i - GP_i)^2}{\sum_1^N (GP_i - Average\ GP)^2}$	(9)
$NAD = \frac{\sum_1^N SP_i - GP_i}{\sum_1^N GP_i} \times 100$	(10)
$RMSD = \sqrt{\frac{\sum_1^N (SP_i - GP_i)^2}{N}}$	(11)
$MAD = \sum_1^N \frac{\lvert SP_i - GP_i \rvert}{N}$	(12)
$MRAD = \frac{\sum_1^N \frac{\lvert SP_i - GP_i \rvert}{GP_i}}{N}$	(13)
$EB = \frac{\sum_1^N SP_i - \sum_1^N GP_i}{\sum_1^N GP_i} \times 100$	(14)
$CPOD\_S = \frac{No.of\ days\ when\ GP > 0.1\ and\ SP > 0}{No.of\ days\ when\ GP > 0.1\ and\ SP \geq 0}$	(15)
$CPOD\_G = \frac{No.of\ days\ when\ SP > 0.1\ and\ SP > 0}{No.of\ days\ when\ SP > 0.1\ and\ SP \geq 0}$	(16)
Where, GP = Gauge precipitation, SP = satellite precipitation, N = Number of data point	

4.8.1 Results and discussions

The statistical summary on comparison of satellite data with the ground gauges on daily, monthly and annual basis over 10 years (2003 to 2012) period are presented in the **Table 12** and the details are given in **Appendix-3.** The comparison is based on the average precipitation in a grid cell of size 0.25° x 0.25° from daily accumulated satellite estimates (TRMMV7) and gauge records.

Table 12: Summary of statistics for the catchment

Parameters	Verification basis					
	Daily		Monthly		Annually	
	Min	Max	Min	Max	Min	Max
NASH (R2)	-2.45	0.32	0.29	0.91	0.30	0.98
NAD(%)	-31.08	84.12	-31.54	84.12	-27.61	73.52
RMSD	5.65	28.28	35.79	280.05	165.58	1540.52
MAD	0.24	5.40	5.56	130.28	32.36	1368.47
MRAD	0.00	0.00	0.00	0.01	0.00	0.07
RR	0.34	0.58	0.87	0.93	0.36	1.00
EB(%)	-31.54	84.12	-31.54	84.12	-31.54	84.12
CPOD_S	0.60	0.89	0.98	1.00	1.00	1.00
CPOD_G	0.61	0.86	0.90	0.99	1.00	1.00

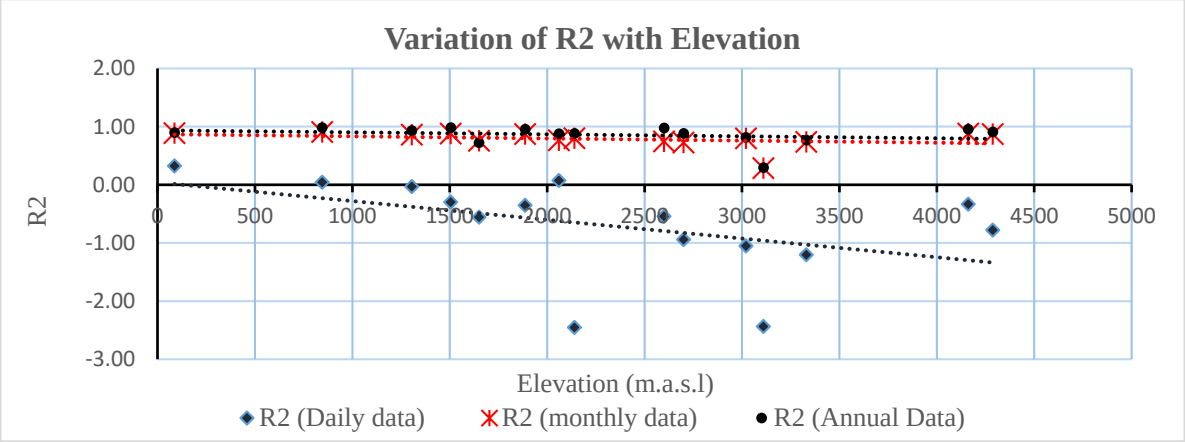


Figure 4.8: Variation of R2 with Elevation

The comparison of daily accumulated satellite data set showed various level of deviation from the gauges with the R2 value ranging from a minimum of -2.45 to a maximum of 0.32 and low correlation (RR) values in the range of 0.34 to 0.58. As can be seen in **Figure 4.8** above that from a total of 15 pixels compared, only three resulted in low positive R2 value. Low and negative values of R2 show that the satellite is not estimating the precipitation well. Furthermore, low positive value of RR indicates that there is a weak to fair correlation between the satellite and gauge precipitations. There is a small trend of lower R2 with increasing elevation in the daily data set. In contrast, when comparison was made with monthly and annual accumulated data, far better results were obtained with R2 and RR ranging from 0.73 to 0.98 and 0.71 to 1 respectively with exception of R2 value of 0.3 in one pixel (S9) and RR value of

0.51 and less in two pixels (S8, S9). Overall both monthly and annual data set performed equally well with no trend in R2 with elevation.

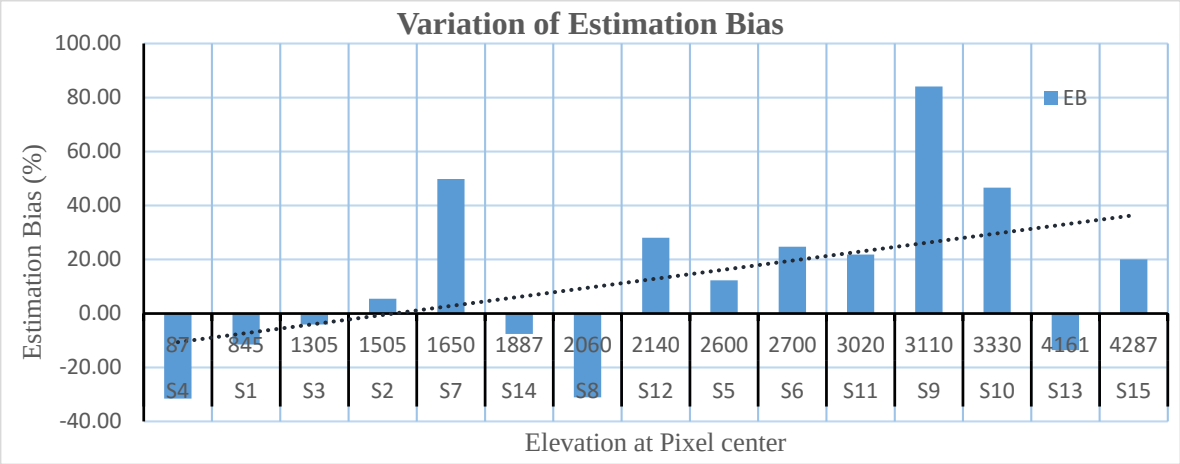


Figure 4.9: Variation in Estimation Bias in pixels

Figure 4.9 shows the variation in estimation bias in the satellite precipitation in each pixels along with corresponding elevation. The biases ranges between -31.54% to 84.12%. Amongst six pixels with negative bias five of them lie below an elevation of 2060 m.a.s.l with high mean absolute difference (MAD) and high root mean square difference (RMSD) which indicates underestimation of precipitation by satellite in the southern part of the basin which receives high precipitation. The result is similar to the findings by Khandu et al. (2015), Bajracharya et al. (2010) and Shrestha et al. (2008) that satellite based estimation underestimates heavy precipitation. However, an average satellite conditional probability of detection (CPOD_S) of 0.79 on the daily data sets, 1 in monthly and annual data set show that there is a good detection efficiency of precipitation by the satellite. In contrast, the case is opposite in the central and northern parts of the basin with positive biases, low MAD and low RMSD demonstrating over estimation.

An overall picture can be seen on the accumulation plots of both satellite and gauge precipitations on daily, monthly and annual basis as shown in **Figure 4.11**, **Figure 4.12**, and **Figure 4.13** below for the pixels S1, S12 and S4 along with scatter plots shown in **Figure 4.10**. The scatter plot in S1, S12 and S4 represent good estimation, underestimation and overestimation by satellite respectively. This can be noticed at a glance by looking at the concentration of plots on, below or above the 45° line. The scatter plots and accumulation plots for other pixels are given in **Appendix-1** and **Appendix-2** respectively. From these plots and statistics summarized in **Table 12**, it can be concluded that the data sets have some

correspondence because there is a general qualitative match between the satellite and gauge precipitations with quantitative differences. Therefore, the biases need to be adjusted before it could be used for further applications.

In case of TRMM3B42V7 daily estimates, the 3 hourly observations are made at Coordinated Universal Time (UTC) of 00:00, 03:00, 06:00, 09:00, 12:00, 15:00, 18:00, 21:00 and these observations are accumulated from 22:30:00 UTC to 22:30:59 UTC of next day. Whereas in case of Bhutan the gauge accumulation time is at 03:00 UTC (9Am to 9AM local time) which gives a delay in accumulation of satellite precipitation by 4.5 hours. This delay can be brought down to 1.5 hours and significant improvement in the R2 and RR can be achieved mainly in comparison of daily data sets if 3 hourly estimates are used and accumulated at 3:00 UTC. This is not an issue when analysis is made with monthly and annual statistics that's why their performance is better.

The distribution of rain gauges in the basin have followed the pattern of human settlement which is clearly depicted by **Figure 4.7**. There are no gauges or the gauges are sparsely distributed in low populous northern portions of the basin. Whereas, in the high populous central and southern foothills the gauges are clustered in a small area. The comparison did not yield good results because the distribution of gauges was inadequate to cover the spatial variability of rainfall in the northern areas. Additionally, since the rain gauges in mountainous regions tend to be in the valleys, the network of ground gauges have the tendency to underestimate the orographically enhanced precipitation that occurs in the mountainous topography, Ebert et al. (2007). The case is well portrayed by **Figure 4.9** which shows that the estimation biases are positive and high indicating that the satellite overestimates the precipitation in the northern regions.

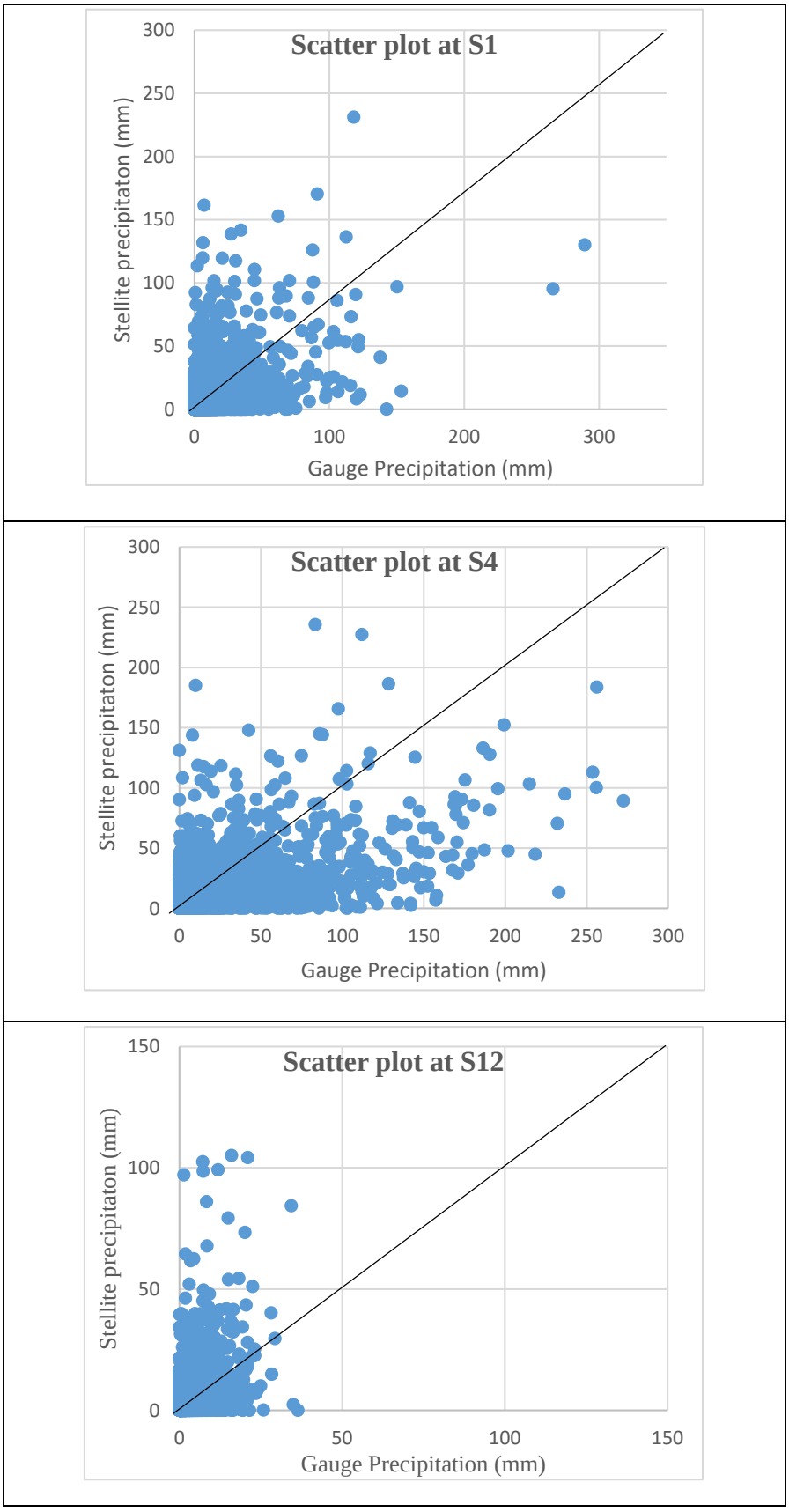


Figure 4.10: Scatter plots for pixels S1, S4 and S12

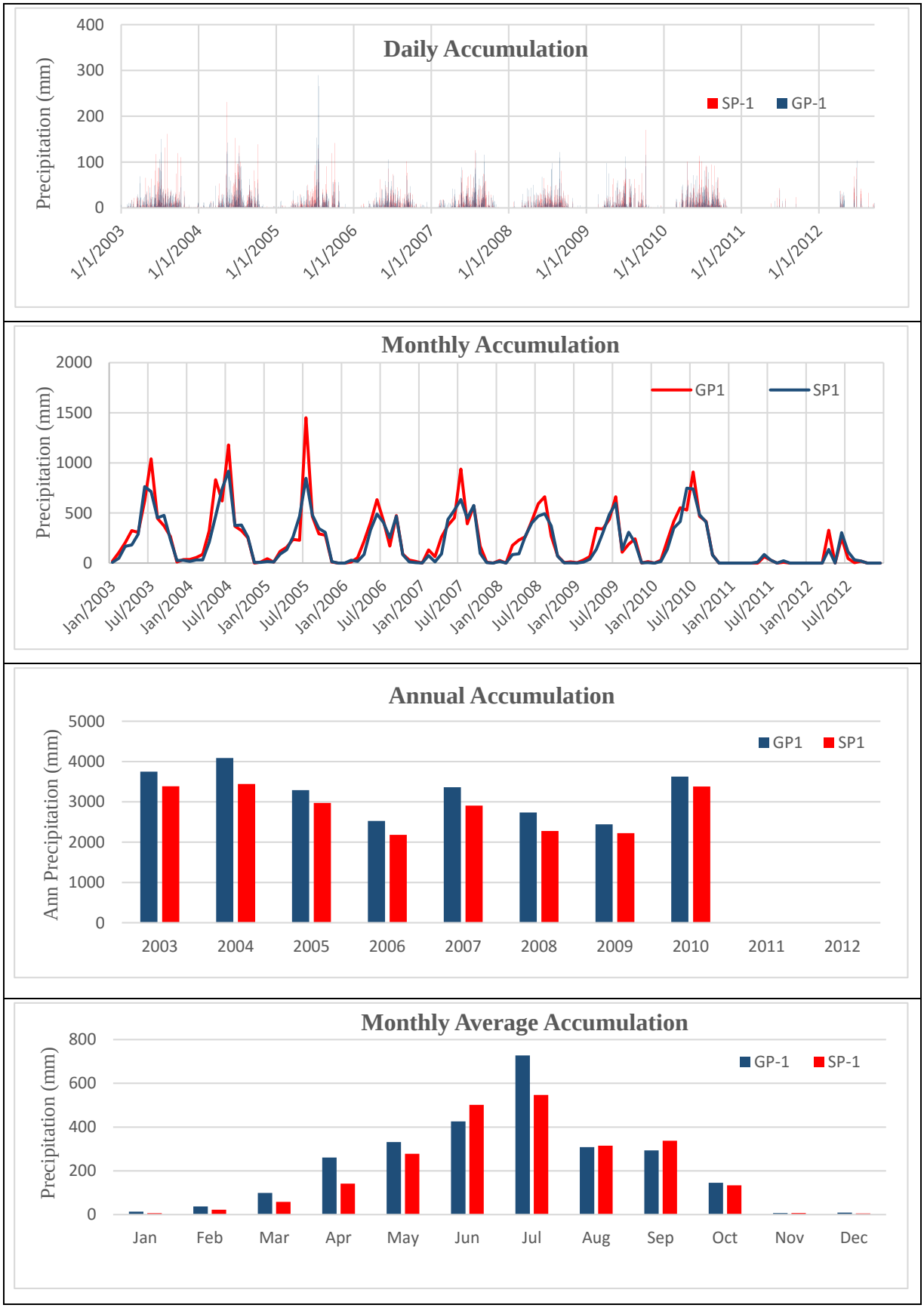


Figure 4.11: Accumulation Plots at S1

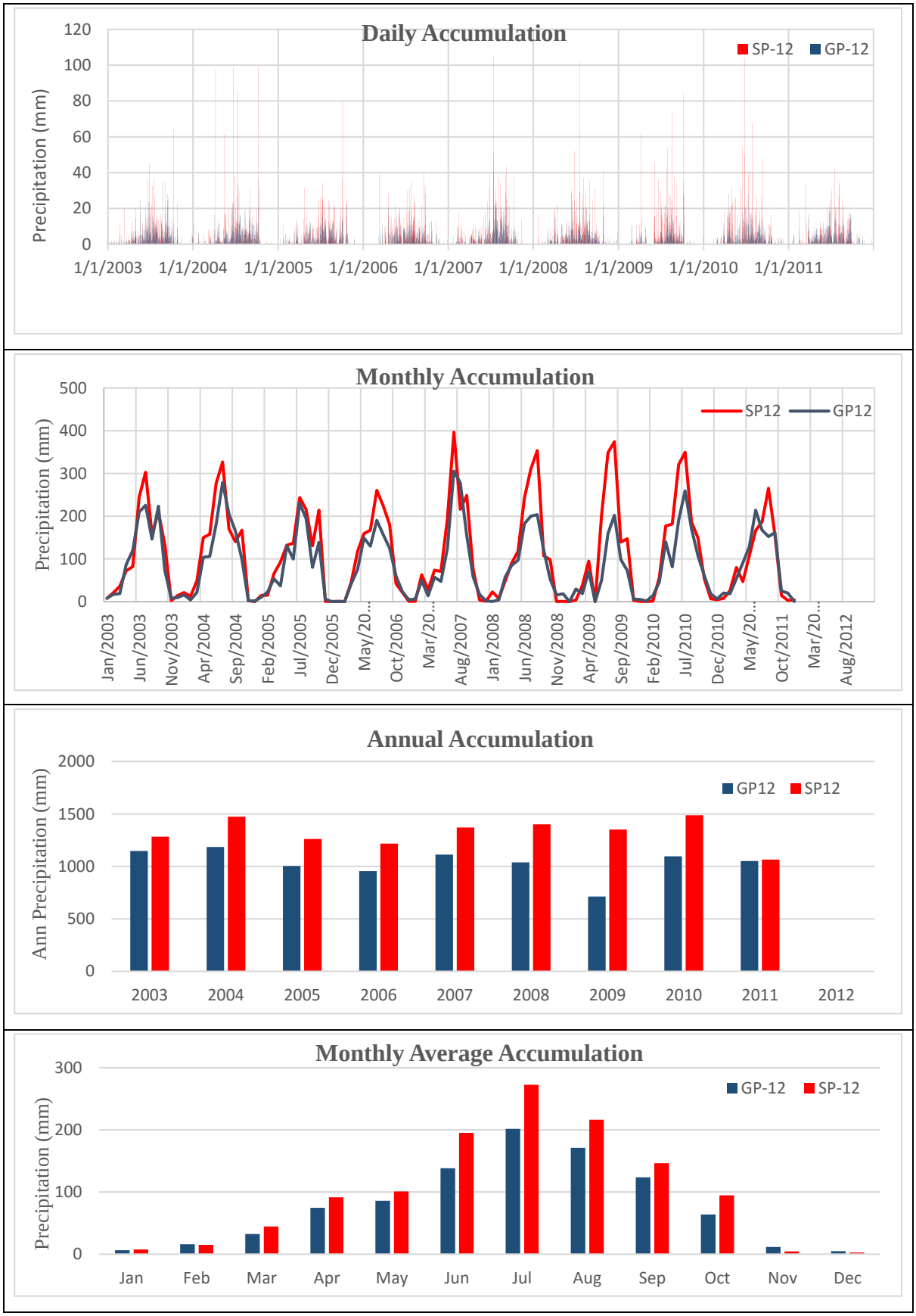


Figure 4.12: Accumulation Plots at S12

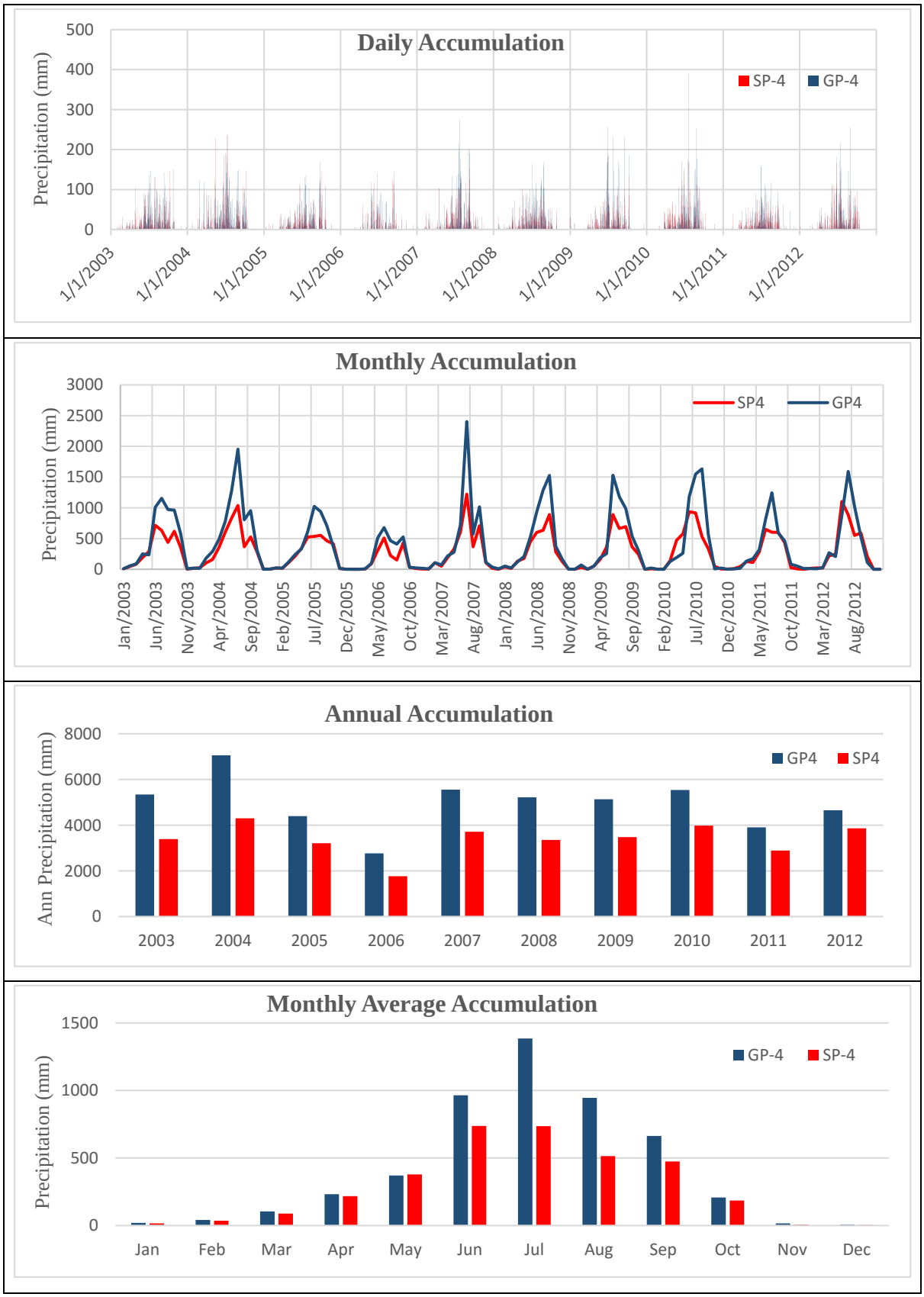


Figure 4.13: Accumulation plots at S4

4.9 Spatial variability within same pixel

When satellite grid is placed over an underlying layer of rain gauges **(Figure 4.7)** in the catchment, it can be seen that there are pixels that has 0 to 6 gauges. In some of the pixel theses gauges have a good spread over the pixel and in some they are clustered in an area smaller within the pixel. To understand the spatial variability within a pixel, one pixel S10 **(Figure 4.14)** with 5 gauges is studied. The gauges are within a radius of 9 to 16 km from the pixel center.

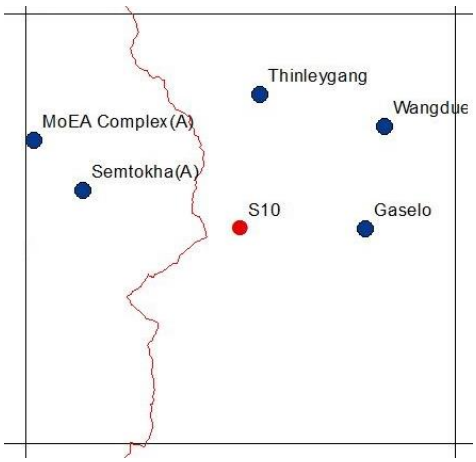


Figure 4.14: Spatial variability within same pixel (S10)

The variation in daily, mean monthly and annual precipitation in a small pixel of area of 27.5 x 27.5 km2 is revealed in **Figure 4.15** which shows daily, mean monthly and annual variation. It can be seen that from November to March all five stations experiences similar precipitation amounts in the range of 0mm to a maximum of about 15mm a month with December being the driest month. The variation in precipitation amounts opens up after march with Simtokha experiencing minimum of 23 mm and Gaselo experiencing maximum 47 mm in April. This gap slowly widens up as monsoon approaches with peak in July where Simtokha still receives the least (131 mm) and Thinleygang experiencing the highest (225 mm) precipitation amounts. After July the gap in the accumulation amounts vary almost consistently until September with least at Simtokha (64 mm) and highest at Thinleygang (154 mm) and suddenly closes between October and November. The precipitation pattern at MoEA and Wangdue vary quite similarly throughout the year. The annual average, maximum and minimum precipitation recorded by the gauges in this pixel is shown in **Table 13.**

Table 13: Annual spatial variability in precipitation

Station	Annual precipitation (mm)		
	Max	**Min**	**Average**
MoEA	704	482	626
Thinleygang	1231	621	928
Wangdi	898	467	643
Simtokha	593	412	533
Gaselo	953	559	726
SP10 (satellite)	1138	823	922

SP10 in mean monthly and annual plots below represents the average satellite precipitation for the pixel S10. It can be seen that due to inadequate gauge especially in the lower parts of the pixel the gauge average is far below the satellite average in addition to the spatial variability within a small pixel of size 27.5 x 27.5 km when comparison is made with 5gauges. Similarly, Tamrakar & Alfredsen (2012) compared 11 gauging stations within a single pixel in Nepal and showed that even in a small pixel area of 27.5 x 27.5 km there is wide variation in precipitation. In the same study it was also demonstrated that the pixel resolution also rules the competency and precision of estimation by satellite. This is mainly because, the pixel area for which the satellite precipitation is averaged will get reduced to an area with more uniformly distributed gauges which could represent well the spatial variability and average precipitation for that area.

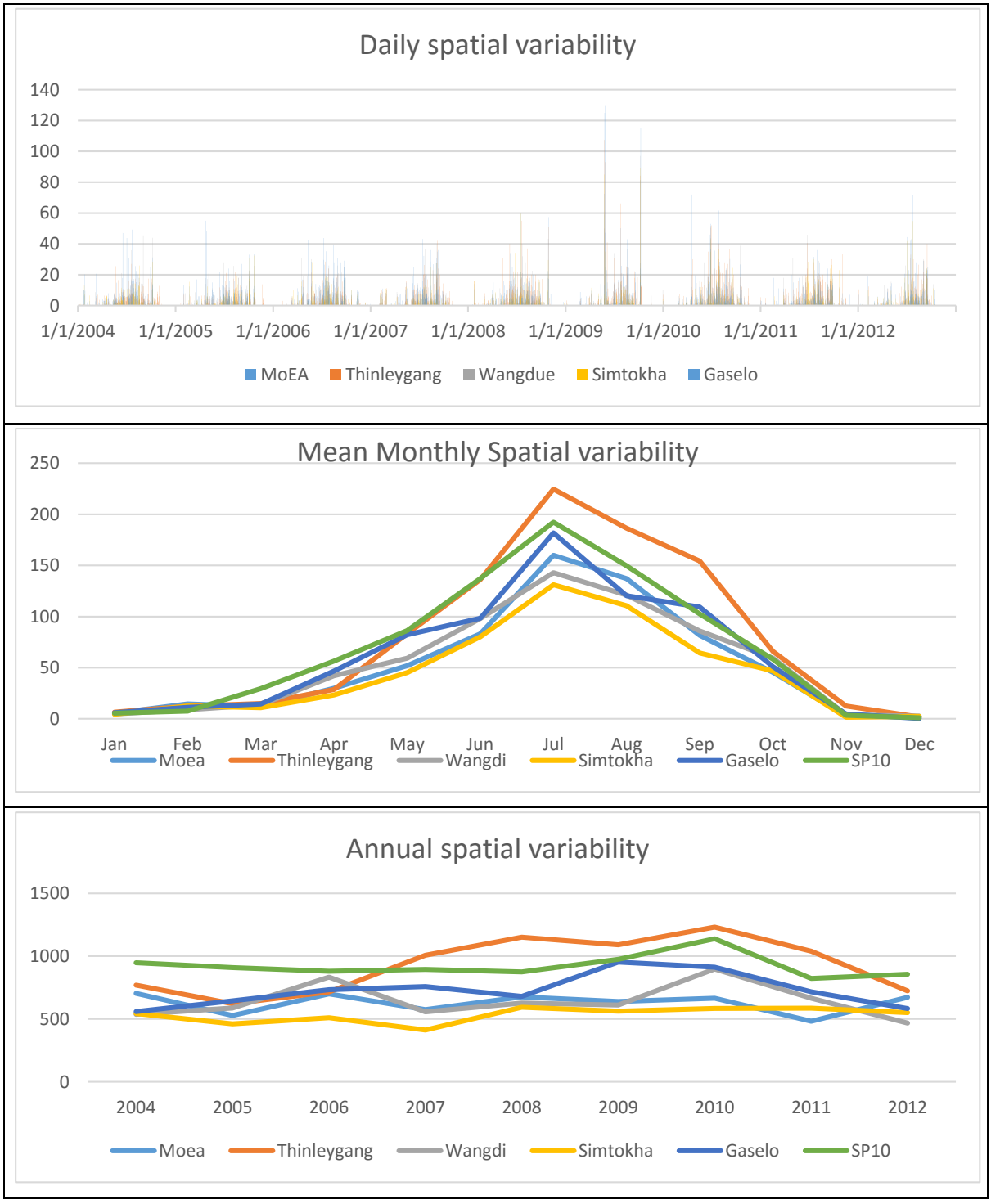


Figure 4.15: Daily, Mean Monthly and Annual Spatial Variability within a pixel

4.10 Areal Precipitation

The areal precipitation for the Kerabari catchment has been generated by SHyFT. The accumulation plots are shown in **Figure 4.16** for both gauge recorded precipitation and satellite estimated precipitation. Statistical analysis was performed for comparison and the results are summarized in **Table 14.**

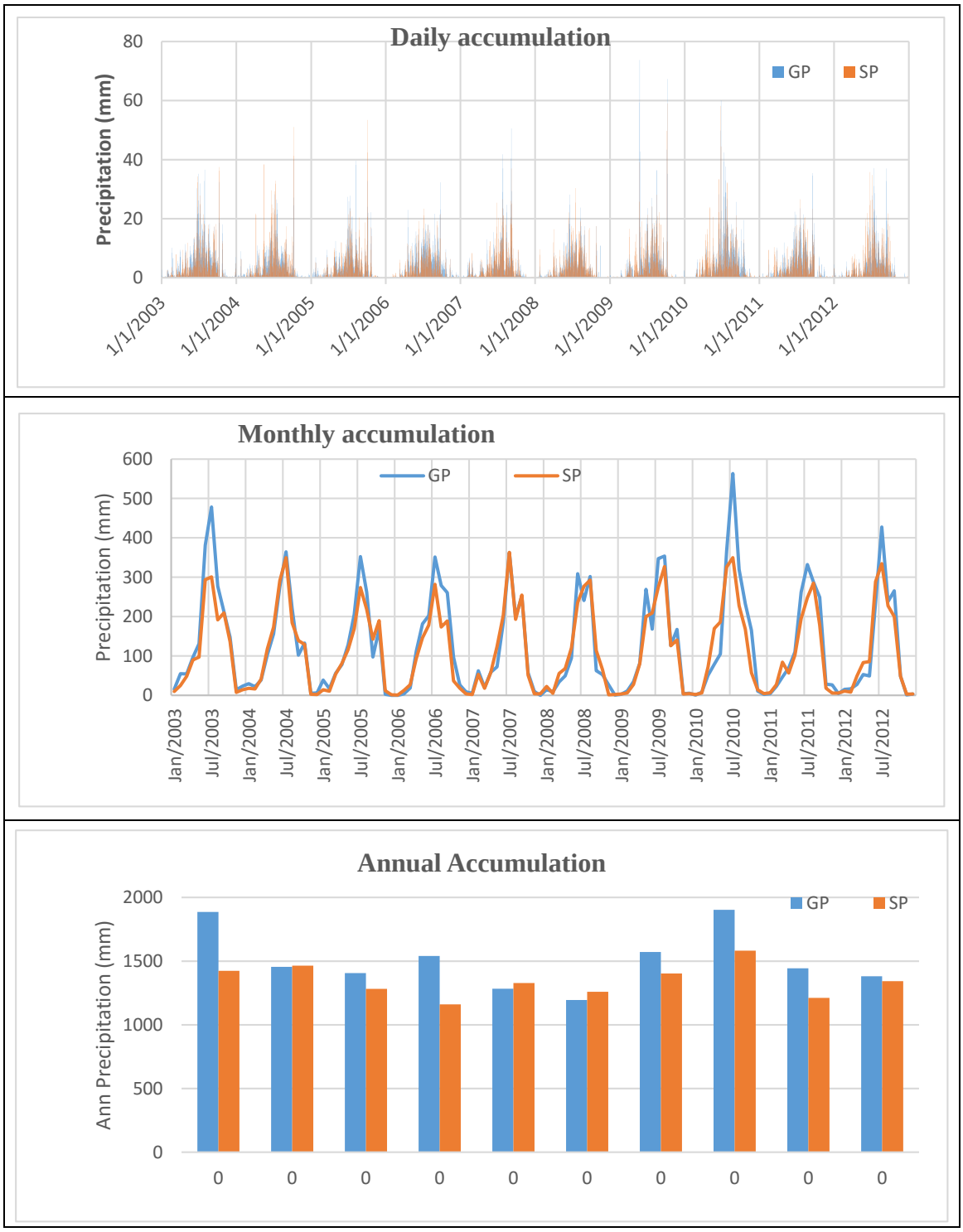


Figure 4.16: Areal precipitation (accumulation plots)

Table 14: Statistics on Areal Precipitation

Statistical parameters	Kerabari		
	Daily	**Monthly**	**Annual**
NASH (R2)	0.49	0.93	0.98
NAD(%)	-10.66	-10.66	-7.60
RMSD	5.12	45.20	187.68
MAD	0.56	13.39	114.46
MRAD	3.7E-05	8.9E-04	7.6E-03
RR	0.73	0.95	0.61
EB(%)	-10.66	-10.66	-10.66
CPOD_S	0.99	1.00	1.00
CPOD_G	0.86	0.98	1.00

The efficiency of detection in terms of NASH-R2 is at an average on daily and better on monthly and annual statistics. The CPOD_S is 0.99 and CPOD_G is 0.86 on daily data which show that the satellite is able to detect rainfall well for most days and it is also detecting precipitation for many days even when the gauges did not record any precipitation event. However, peak events during monsoon in some of the years have been underestimated by the satellite but there are also some years in which there was overestimation of peaks. An overall estimation bias of -10.66% indicates underestimation of areal precipitation by satellite. The biasness mainly occurred during the monsoon months when the satellite was not able to detect peak events.

It can also be seen from the accumulation plots that the areal precipitation in the Kerabari catchment is well represented by the satellite estimates in terms of timing and with relatively good match of magnitudes in all time scales when compared with the areal precipitation from the gauges.

5 SHyFT

5.1 Genereal

The Statkraft Hydrologic Forecasting Toolbox (SHyFT), is an open source hydrological tool box. The toolbox provides an optimized platform for efficient modelling of hydrologic processes. Financed by Statkraft, the development of the tool was initially with the goal of enhancing hydrological estimates for hydropower scheduling but there is also a great possibility that it might turn into the future hydrological tool kit for hydropower planners and developers (https://github.com/statkraft/shyft).

The recent developments have introduced more physically based and process-level methods, SHyFT follows the paradigm of distributed and lumped models. The code is based on an early initiative for distributed hydrological simulation, called ENKI developed at Sintef (https://github.com/statkraft/shyft). SHyFT is more flexible to input data and is not limited to raster based simulations like ENKI and LANDPINE. Additionally, SHyFT will run faster and calibration results will be available quickly for people dealing with operational hydrology especially hydropower companies which require to run hundreds of simulations to forecast runoff for reservoir operations.

5.2 Requirements and installation procedure

The important requirement and installation procedure are given in **Appendix-4**

5.3 The model setup

The **Figure 5.1**shows a general setup of the model with the arrangement of folders and sub-folders along with the location of input files, program files and the configuration files. The sub folder api contains the python wrappers for the shyft core containing basic data structure, cell-medels, models and algorithms. The orchestration folder contains a basic infrastructure to read the orchestration codes that uses YAML configuration files to define a calibration or a simulation run. SHyFT needs to ingest observed hydro-meteorological data with previous calibration runs to fill its internal data structure and proceed with simulation run. This process of ingestion of data is called orchestration in SHyFT. It also allows users tailor the calibration or simulation by adding their own codes. Repository contains some python code which can read data collected in some arbitrary format and feed it to the SHyFT core. Test contains information of all routines with equations and methods and makes the integral part of operation of SHyFT.

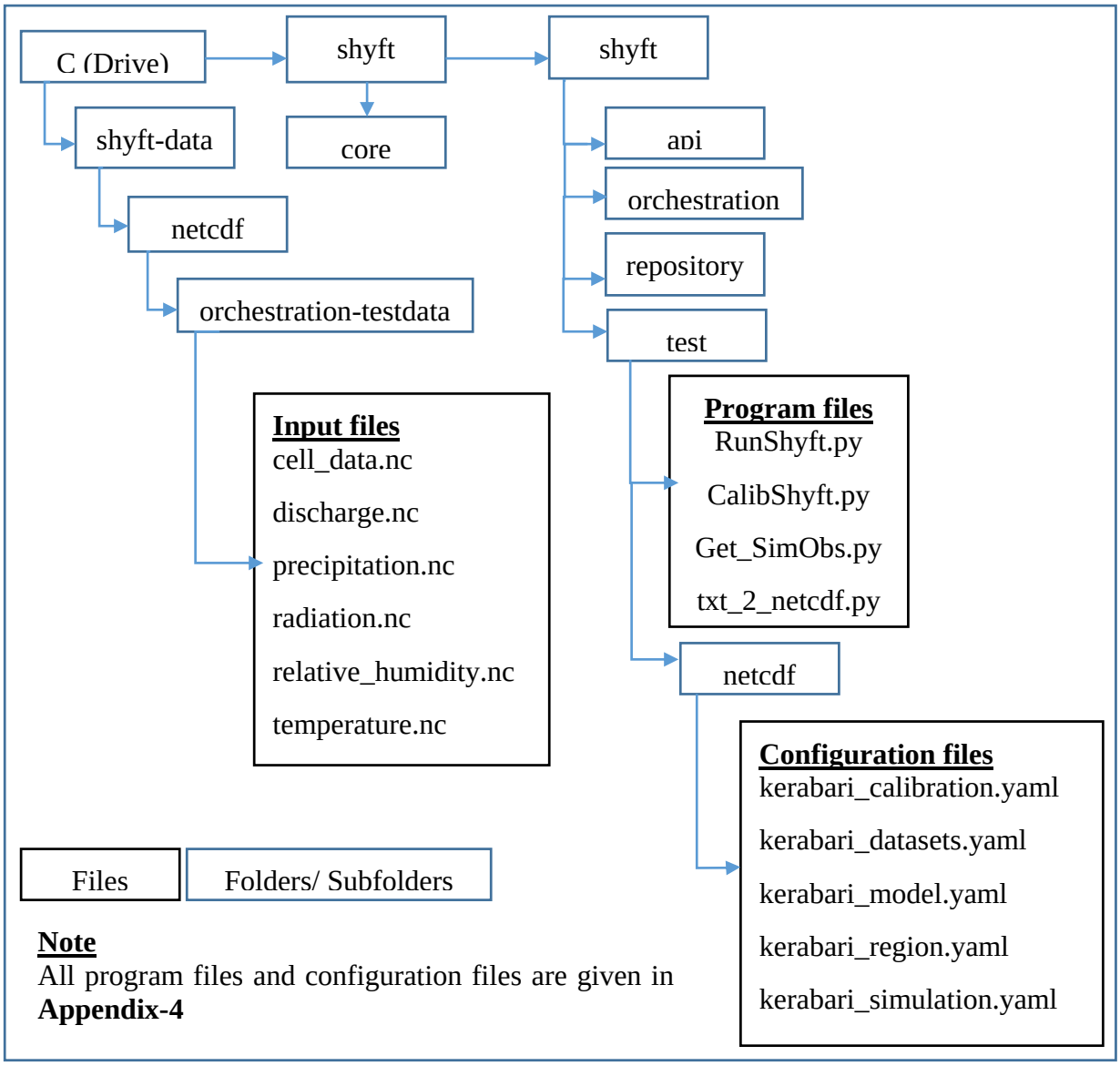


Figure 5.1: The SHyFT Setup

The subfolder netcdf in test contains all the YAML configuration files which are vital to functioning of SHyFT. A brief description on each of these files are given below:

1. **Region.yaml** configures the domain of the area of interest (kerabari in this study) with EPSG code of the region and coordinates of lower left corner of the domain along with number of grids/cells in x and y direction. There could be a single or multiple catchments with some unique catchment indices inside this domain. The region.yaml file connects to the physiographic data (cell_data.nc) available in shyft-data for the specified domain/region.

2. **Datastes.yaml** is responsible to connect with the climate data (precipitation.nc, temperature.nc, wind_speed.nc, relative_humidity.nc and global_radiation.nc) inside shyft-data
3. **Model.yaml** contains the interpolation and model parameter set for the simulation and calibration purpose
4. **Simulation.yaml** calls and returns the region.yaml, dataset.yaml and model.yaml files with user defined function to specify the start time, run time step and number of steps to be used in the simulation and calibration process
5. **Calibration.yaml** calls and uses the simulation.yaml as the simulator using one of the three optimizers (min_bobyqa, sceua and dream) to do an automatic calibration of the model using the input discharge.nc file which is stored in shyft-data folder

RunShyft.py is a program file that performs the simulation by importing the simulation.yaml file wihich is already configured to the area of interest. Likewise CalibShyft.py is another program file that performs the calibration by importing the calibration.yaml file which is also configured to the area of interest. The Get_SimObs.py when executed gives out the time series of observed runoff, simulated runoff, areal temperature and precipitation of the study area in Microsoft Excel format from where one can easily perform some statistical analysis and make plots. The txt_2_netcdf.py converts the tab delimited text file containing physiographic and climatic data into netcdf format which has to be kept inside the folder orchestration-testdata.

5.4 The model structure

SHyFT is a distributed hydrological model which works from the regional level to the cell level by distributing the input parameters in to the individual cells. **Figure 5.2** gives an overview of the model structure in SHyFT. A brief description of the elements in the model are given below.

Region

It is a geographic region associated with some data which describes the properties at cell or grid level for the region. Physiographic data such as area, elevation, forest cover, lake percentage etc.., are termed as static data and observed climate data such as temperature, precipitation, radiation wind speed and snow are variables. The region is also associated to the responses such as observed runoff recorded as one or few cells.

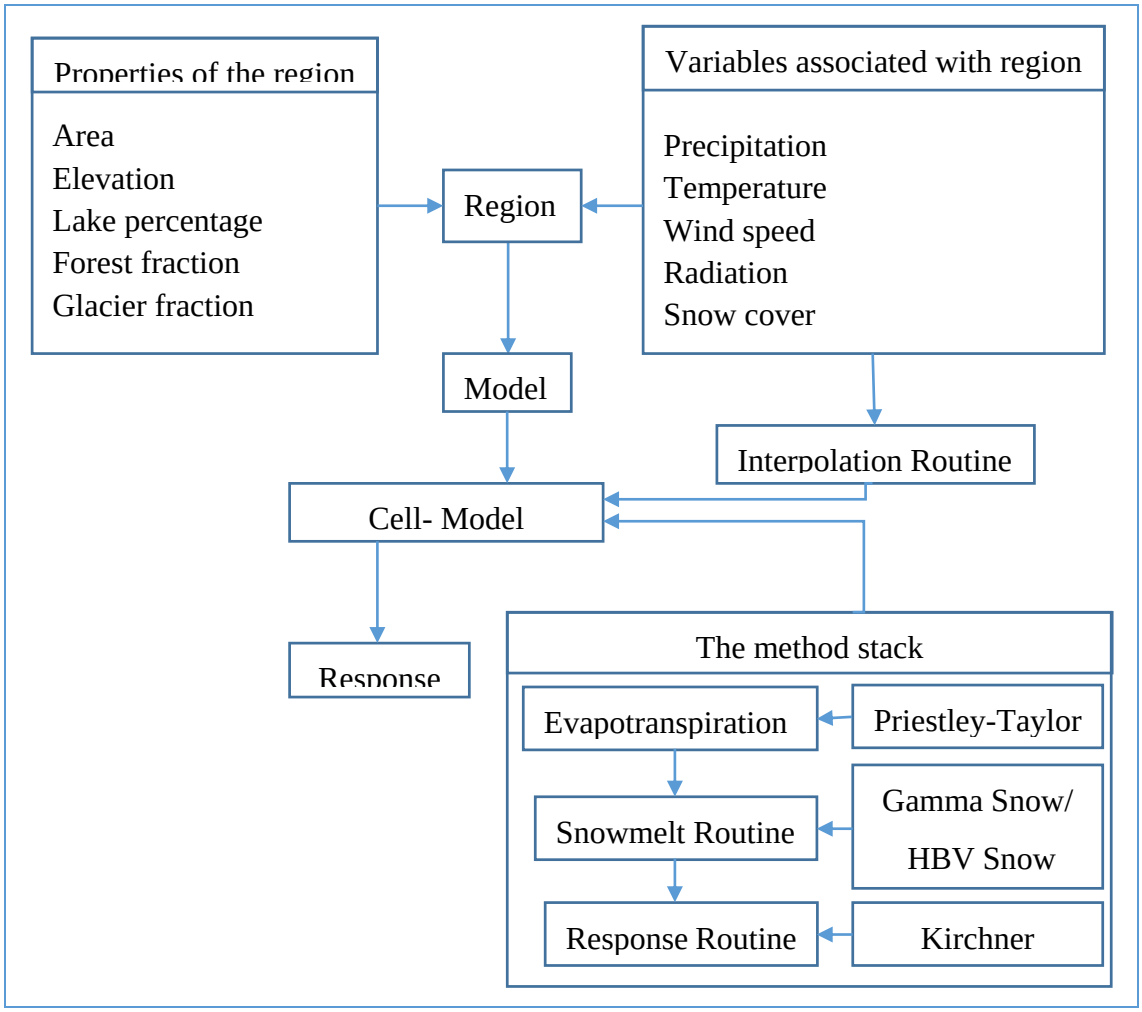


Figure 5.2: The conceptual model

Model

It is a computational model which takes inputs from the region. Given static properties, observed variables and initial state as input the model can compute responses, forecasts for runoff, snow reservoir, new set of state, optimized model parameters etc.., according to the method composition and parameters selected.

Cell -Model

The cell model takes initial cell-state, calibration parameters and cell environment inputs (precipitation, temperature etc.) and computes the response and a new cell – state for each time-step using the layered method stack.

The interpolation routines such as IDW and Bayesian are responsible for projecting input sources forecasted at some location to individual cells.

5.5 The method stacks

Evapotranspiration routine

The evapotranspiration routine uses the Priestley-Taylor's model. The model which is a modification of Penman's equation through empirical approximations to eliminate the need for input data other than radiation data (Kouwen, 1986) is described by equation shown below.

$$PET = \alpha \frac{S(T_a)}{S(T_a) + \gamma}(K_n + L_n)\frac{1}{\rho_w \lambda_v} \quad (17)$$

Where, PET is the Potential evapotranspiration, K_n is the short-wave radiation, L_n is the long-wave radiation, $S(T_a)$ is the slope of saturation –vapor pressure versus the temperature curve, γ is the Psychrometric constant, ρ_w is the mass density of water and λ_v is the latent heat of vaporization and α is the Priestley-Taylor's constant. The α may vary throughout the day and season to season but an average value of 1.26 is generally accepted (Kouwen, 1986).

Snowmelt routine

SHyFT provides an option to choose either of HBV - snow routine or Gamma-snow routine. In this study, the melt release from the saturated snow is worked out in the snowmelt routine based on the Gamma distributed snow depletion curve (SDC) shown below. Snow-rain threshold temperature and snow melt sensitivity to wind speed are free parameters in this routine.

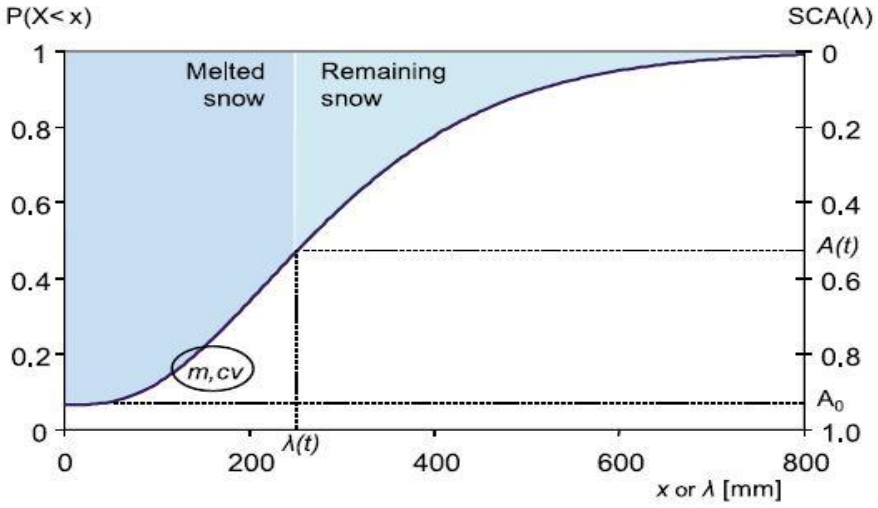


Figure 5.3: The Snow Depletion Curve [Kolberg et al. (2010)]

The Snow depletion curve (SDC) in **Figure 5.3** from Kolberg et al. (2010) is given a three parameters model. Two parameters quantify the snow pack by a Gamma distribution and the third quantify the maximum snow covered area in a cell at melt onset. The full SDC model for a single grid cell is given by:

$$A(t) = A_0.\{1 - F[\lambda(t)]\} \tag{18}$$

$$F[\lambda(t)] = \int_0^{\lambda(t)} p(x; m, cv)dx = \gamma\left(\frac{1}{cv^2}, \frac{1}{cv^2}, \frac{\lambda}{m}\right) \tag{19}$$

Where, p is the probability density function, F is the cumulative probability distribution function which is equal to the incomplete Gamma function (γ) with shape and scale arguments and A(t) is the snow covered area of the grid cell at time t. The variables defining the state of the snow pack in each cell are:

1. The average snow water equivalent m (mm) at the beginning of the melt season
2. The coefficient of variation of snow water equivalent cv which quantifies the heterogeneity of the cells
3. The snow covered area A_0 at the beginning of melt season and
4. The accumulated melt depth since the melt season λ(t)

Response routine

The response routine uses the three-parameter Kirchner's model for computing discharge based on observed precipitation and evapotranspitation data. The model is developed based on the assumptions that:

a. The flow Q solely depends on the amount of water stored (S) in the catchment

$$Q = f(S)\ and\ S = f^{-1}(Q) \tag{20}$$

b. The flow in the catchment is primarily controlled by the release of water from the storage rather than bypassing flow from direct precipitation and
c. The saturated and unsaturated storages are hydraulically connected and the net ground water flow across watershed boundary is zero.

This algorithm for this routine is based on the log transform of the formulation of the time change in discharge as a function of measured precipitation, evapo-transpiration and discharge as shown by the equation below.

$$\ln\big(g(Q)\big) \approx \ln\left(\frac{dQ}{dS}\right) \approx \ln\left(\frac{-dQ/dt}{Q}\middle| P \ll Q, E \ll Q\right) \approx C_1 + C_2 \ln(Q) + C_3(\ln(Q))^2 \qquad (21)$$

Where C_1, C_2 and C_3 are the first, second and third parameter in the Kirchner's model.

6 Preparation of input data

In the current set up of SHyFT, we will require all input data in netcdf format. To do this we require to use ArcGIS, Microsoft Excel and some python scripts as shown by the process diagram in **Figure 6.1.**

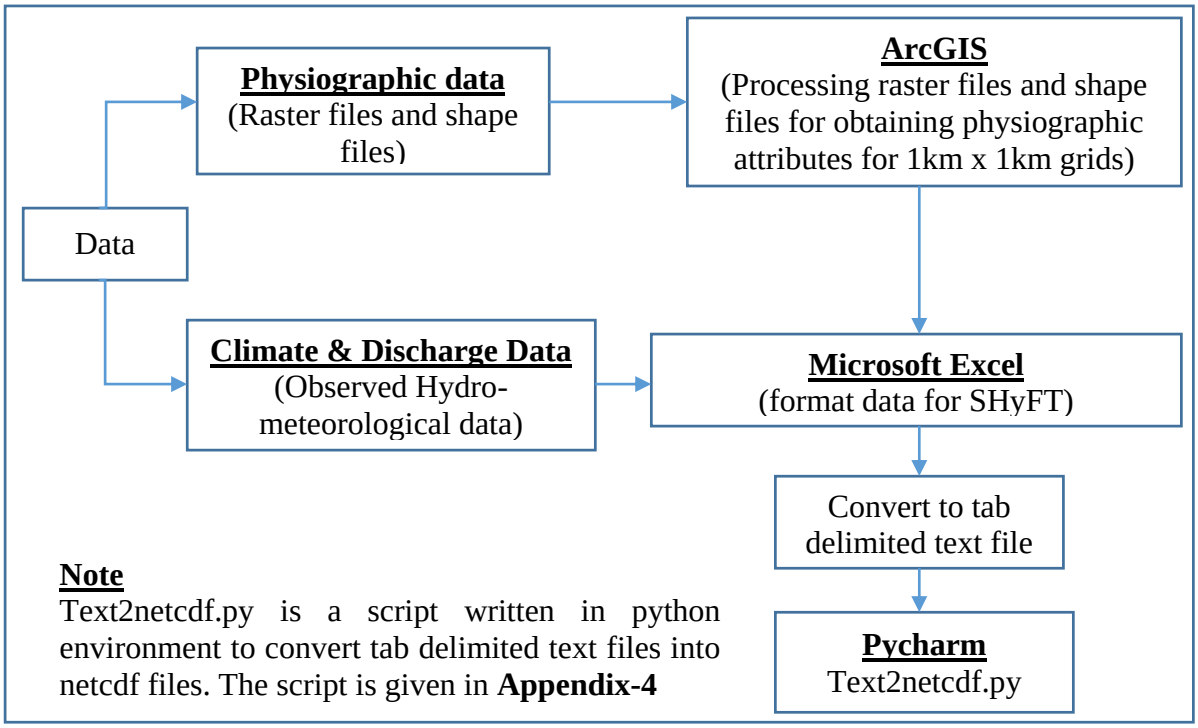


Figure 6.1: Procedure for preparing data for SHyFT

6.1 Physiographic Data required

As shown in **Figure 6.2** the physiographic data of the catchment is gridded and georeferenced in a grid size of 1km x 1km using ArcMap. The physiographic features required for each grid of 1 km2 along with its x and y coordinate of the center of each grid are:

1. Area (m2)
2. Elevation (m.a.s.l)
3. Reservoir fraction (0 to 1)
4. Lake fraction (0 to 1)
5. Forest fraction (0 to 1) and
6. Glacier fraction (0 to 1)

The elevation data is extracted from a 25 m resolution digital elevation model (DEM) provided by the Department of Hydropower and Power Systems, Ministry of Economic Affairs of

Bhutan and other features such as area, reservoir, lake, forest and glacier fractions are extracted from the land cover map provided by the National Soil Service Centre, Ministry of Agriculture and Forests of Bhutan. Both the files were available in raster format and it was possible to extract these data using GIS.

6.2 Format of Physiographic data for SHyFT

All layers of raster and shape files used in the process of making data was projected to UTM zone 46N with WGS 1984 datum for Bhutan. **Figure 6.2** shows the Kerabari catchment which is gridded into 1km x 1km along with the required physiographic data for each grid. Once the catchment physiographic features are gridded for 1 x 1 km2 grid and georeferenced for each feature, the attribute tables are joined together using spatial join tool in GIS. The single attribute table consisting of all physiographic data is then exported to Microsoft Excel and formatted to make a table of data as shown in **Table 15.**

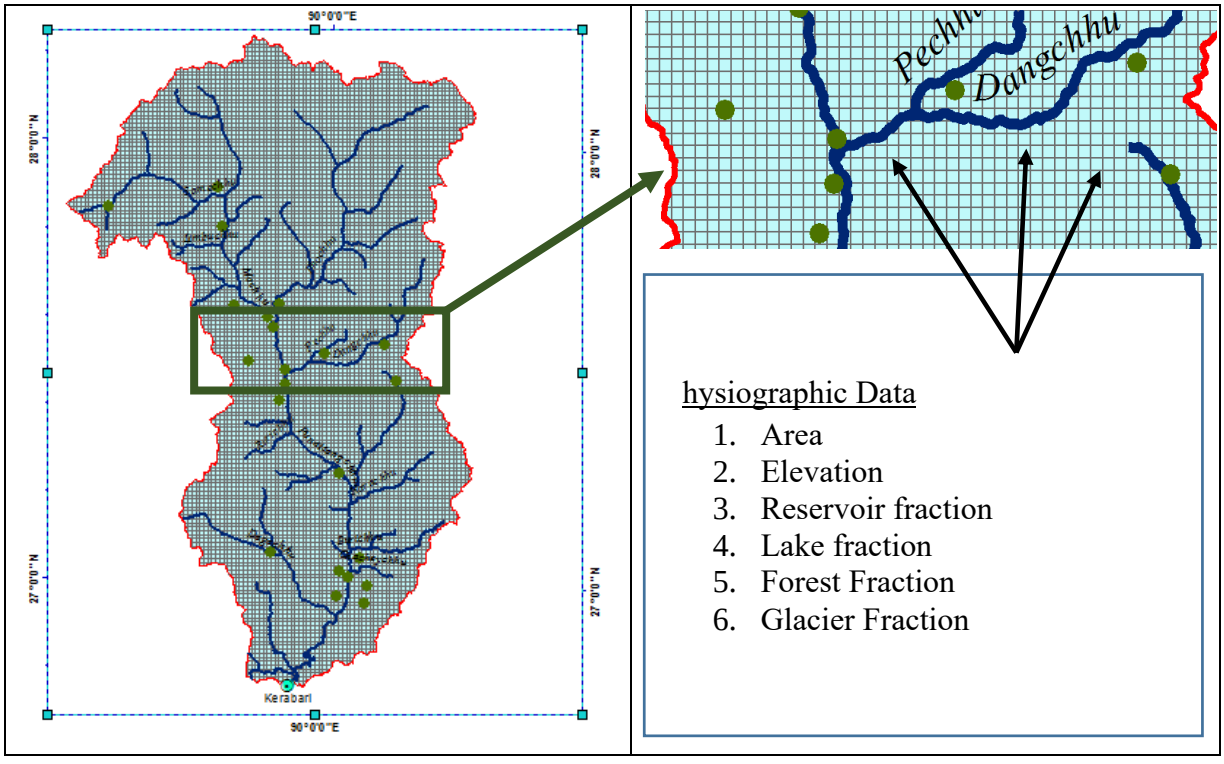


Figure 6.2: Physiographic Data

Table 15: Physiographic Data format in MS Excel

EPSG	**32646**							
Cat ID	**X_Coordinate [m]**	**Y_Coordinate [m]**	**Elevation [m]**	**Area [m2]**	**reservoir_fraction [0-1]**	**lake_fraction [0-1]**	**forest_fraction [0-1]**	**glacier_fraction [0-1]**
0	194316	2964552	305	1	0	0	0.029	0
0	196627	2964617	597	1	0	0	0.002	0
0	197757	2964390	660	1	0	0	0.196	0
0	xxx	xxx	xxx	xxx	xxx	xxx	xxx	xxx
1	184316	2964546	214	1	0	0	0.031	0
1	185206	2953443	222	1	0	0	0.004	0
1	yyy	yyy	yyy	yyy	yyy	yyy	yyy	yyy

Note:

1. EPSG is a geodetic parameter. For Bhutan which uses UTM Zone 46N the EPSG code is 32646
2. The Cat ID (Catchment Identity) is unique for a catchment. Physiographic data for multiple catchments or sub-catchments can be provided by using unique Cat ID for example in the table above, Cat ID 0 and 1 is used to show physiographic data for two catchments.

6.3 Climate and discharge data

The climate and discharge data recorded in the ground hydro-meteorological stations are also prepared in Microsoft excel. The data required is shown in **Figure 6.3** and the format for precipitation data is shown in **Table 16** as an example. Similar to precipitation other data such as temperature, wind speed global radiation and discharge from multiple stations are also prepared taking care of the units of measurement.

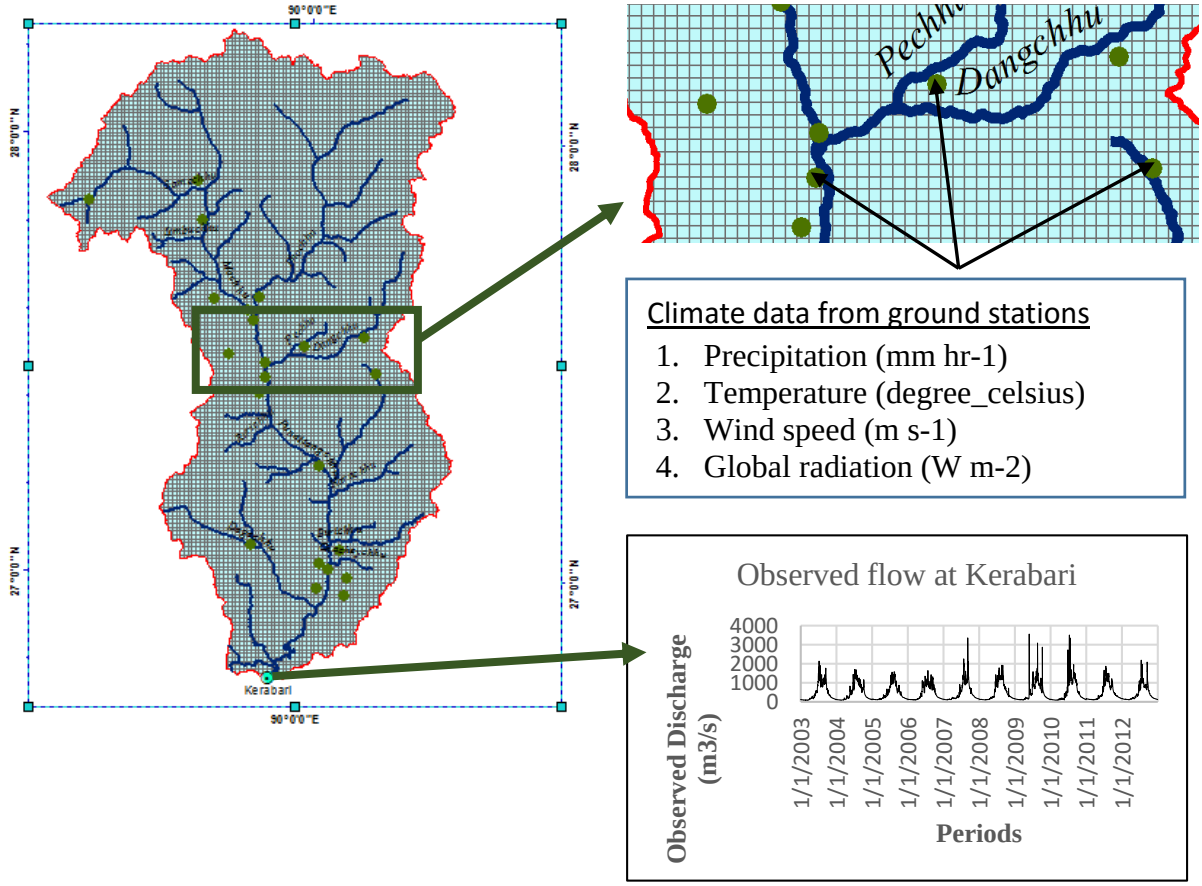


Figure 6.3: Climate and discharge data

Table 16: Input format for precipitation

EPSG	32646			
Missing_value	-999			
Unit	mm hr-1			
Station_Name	**Wangdue**	**Sankosh**	**Punakha**	**xxxx**
X_coordinate [m]	193764	209378	190625	xxxx
Y_coordinate [m]	3044170	2991660	3054780	xxxx
Elevation [m]	1180	410	1236	xxxx
2003.01.01.00:00	0	0.32	0.16	xxxx
2003.01.02.00:00	0	0	0	xxxx
xxxx	xxxx	xxxx	xxxx	xxxx

6.4 Converting text to netcdf

When all the input data is ready in Microsoft excel, the data files are then saved as tab delimited text files with the following file names:

1. precipitation.txt
2. temperature.txt
3. wind_speed.txt
4. relative_humidity.txt
5. global_radiation.txt
6. discharge.txt
7. cell_data.txt

Once the input data in text format is ready, the script **text2netcdf** is executed to convert all the input files into **netcdf** format by properly following the path for input and output file location as mentioned in the script. All these netcdf files are then placed in **C:\shyft-data\netcdf\orchestration-testdata.**

7 Model Calibration and validation

7.1 General

Based on subjective or objective methods, the process of estimating the best set of free parameters that gives the best simulation when compared to observed runoff is called Calibration. The free parameters are those parameters that cannot be measured directly in the field and must be determined though model calibration. Subjective methods are based on comparison of plots of observed and simulated runoff while objective methods are based on the use of numerical criteria, an error function which is derived from the differences between observed and simulated runoff during the calibration period. Example of numerical goodness of fit criteria or the quantitative performance indicators are Nash-Sutcliffe efficiency criteria (R2) and the water balance criteria (Rinde, 2015, lecture notes) as described by the equations below.

R2 criteria:

$$R2 = 1 - \frac{\sum_1^N (Q_s - Q_o)^2}{\sum_1^N Q_o \ - Average\ Q_o)^2} \tag{22}$$

Water balance criteria:

$$Accumulated\ difference = \sum_1^N (Q_S - Q_O) \tag{23}$$

Where N is the number of data, Q_S is the simulated runoff and Q_O is the observed runoff. The basic objective in the process of calibration by using above two criteria is to maximize the R2 value and minimize the accumulated difference.

7.2 Calibration

A brief description of the parameters to be calibrated with their default values are presented in **Table 17.** The model is calibrated to estimate the best set of parameters for the following three input cases:

1. Gauge precipitation as input (Calibration_G)
2. Raw satellite estimates as input (Calibration_RSE)
3. Bias corrected satellite estimates as input (Calibration _BCSE)

Table 17: Parameters to be calibrated

Parameters	Description	Default Values
c_1	1st parameter in Kirchner model	-3.336
c_2	2nd parameter in Kirchner model	0.334
c_3	3rd parameter in Kirchner model	-0.125
ae_scale factor	Actual evapotranspiration scale factor	1.50
TX	Snow rain threshold temperature (°C)	-0.575
wind_scale	Slope in turbulent wind function [m/s]	1.896
max_water	Maximum liquid water content	0.10
wind_const	Intercept in turbulent wind function	1.0
fast_albedo_decay_rate	Albedo decay rate during melt [days]	6.753
slow_albedo_decay_rate	Albedo decay rate in cold conditions [days]	37.173
surface_magnitude	Surface layer magnitude	30.0
max_albedo	Maximum albedo value	0.90
min_albedo	Minimum albedo value	0.60
snowfall_reset_depth	Snowfall required to reset albedo [mm]	5.0
snow_cv	Spatial coefficient variation of fresh snowfall	0.40
glacier_albedo	Glacier ice fixed albedo	0.40
p_corr_scale_factor	Precipitation correction scale factor	1.0
snow_cv_forest_factor	Spatial coefficient of variation of fresh snow with forest	0.0
snow_cv_altitude_factor	Spatial coefficient of variation of fresh snow with altitude	0.0

With an objective to maximize the R2 value, an automatic calibration was performed using SCE-UA and initial parameters as default parameters in SHyFT. As the observed data is available for a period of ten years (2003 to 2012) first five years' data are used for calibration and the remaining five years' data are kept for the purpose of validation of the model.

7.3 Calibrated parameters

The calibrated parameters for each of the three different input cases are given in **Table 18** along with the performance indicators.

Table 18: Result of calibration

	Calibration_G	**Calibration_RSE**	**Calibration_BCSE**
Input precipitation	Gauge (G)	Raw Satellite estimates (RSE)	Bias corrected Satellite estimates (BCSE)
Calibration period	2003 to 2007	2003 to 2007	2003 to 2007
c_1	-8.378	-8.378	-7.769
c_2	-0.895	-0.895	-0.885
c_3	-0.125	-0.125	-0.078
ae_scale factor	1.50	1.50	1.50
TX	-0.280	-0.280	-1.985
wind_scale	5.439	5.439	2.646
max_water	0.10	0.10	0.10
wind_const	1.0	1.0	1.0
fast_albedo_decay_rate	8.438	8.438	9.564
slow_albedo_decay_rate	20.665	20.665	29.557
surface_magnitude	30	30	30
max_albedo	0.90	0.90	0.90
min_albedo	0.60	0.60	0.60
snowfall_reset_depth	5.0	5.0	5.0
snow_cv	0.40	0.40	0.40
glacier_albedo	0.40	0.40	0.40
p_corr_scale_factor	1.116	1.116	1.507
snow_cv_forest_factor	0	0	0.00
snow_cv_altitude_factor	0	0	0.00
NASH-R2	**0.742**	**0.666**	**0.738**

7.4 Calibrated Simulation

The calibrated parameters as presented in **Table 18** were updated in the model and simulation was run again to obtain the simulated runoff for each of the input cases as discussed earlier. The simulated plots on daily and monthly time scales for the calibration period are shown in **Figure 7.1.**

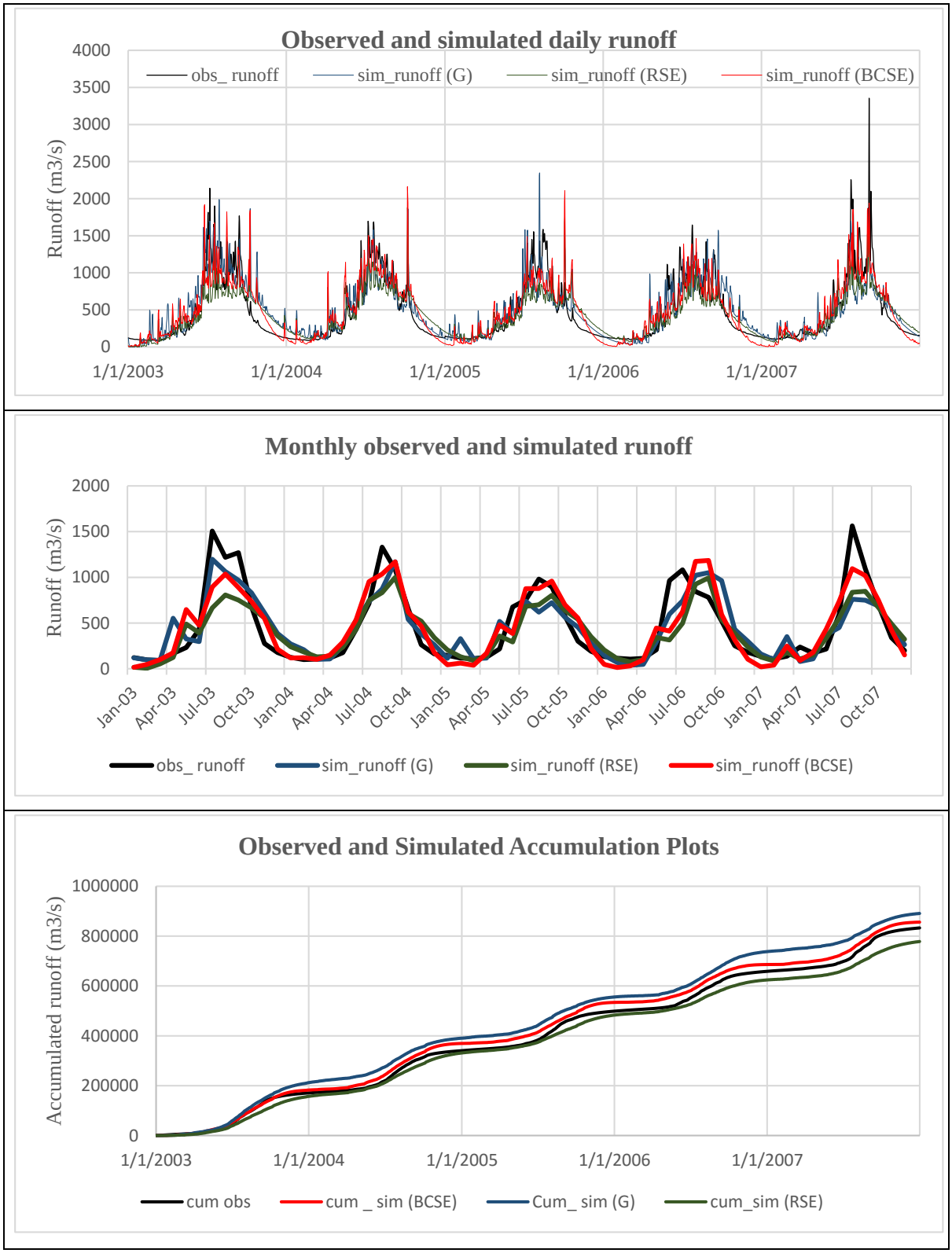


Figure 7.1: Comparison of observed and simulated runoff (calibration period)

7.5 Model validation

A new simulation was run to validate the model for another period (2008 to 2012) that was not used during calibration. A comparison of observed and simulated runoff during this period is shown in **Figure 7.2.**

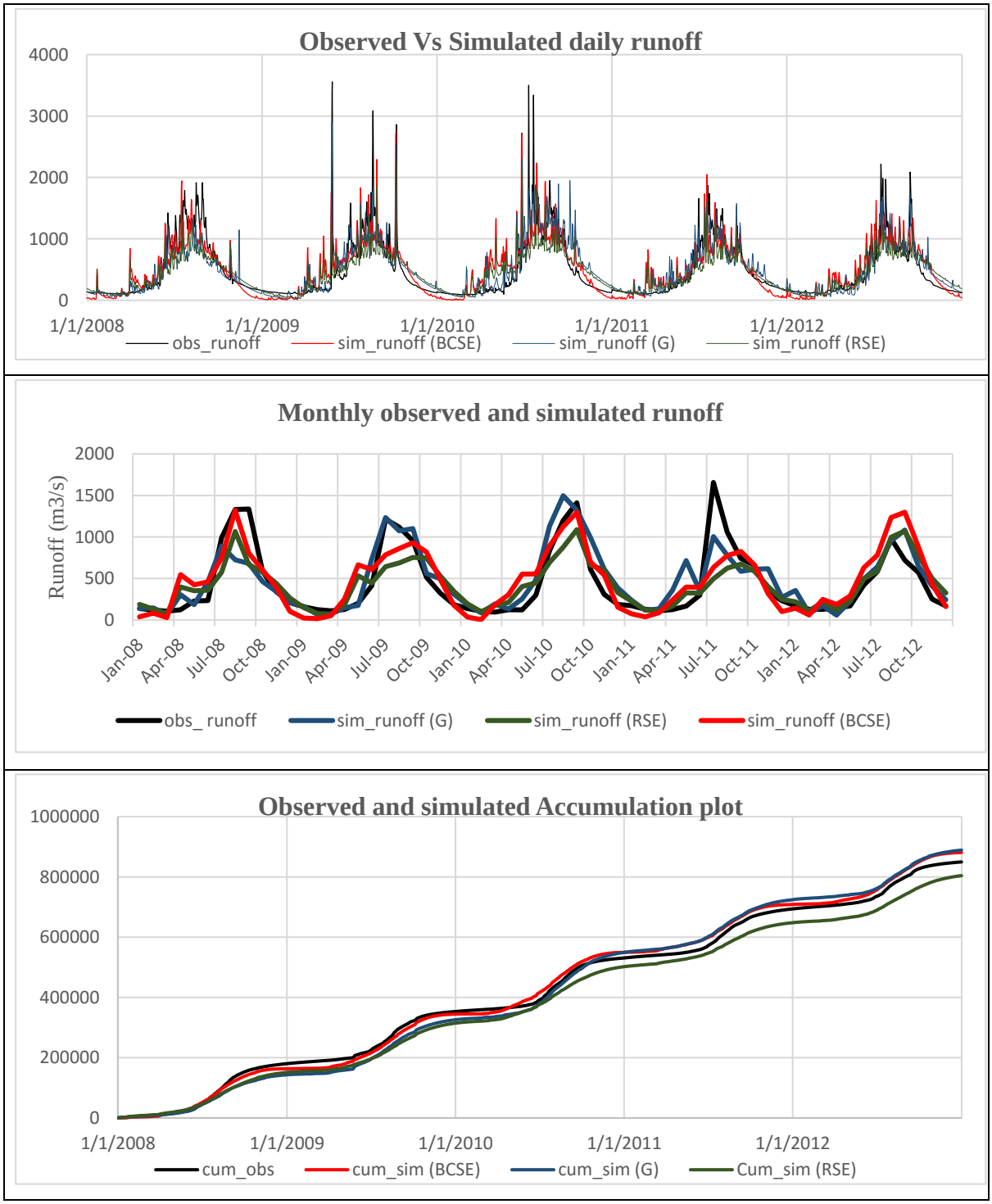


Figure 7.2: Comparison of observed and simulated runoff (validation period)

8 Results and Discussion

8.1 Results

The model parameters are calibrated for three different input cases for a period of five years from January 2003 to December 2007 using the automatic calibration method (SCE-UA) by maximizing the R2 value between observed and simulated runoff. The calibrated model is subsequently validated for another period of five years from January 2008 to December 2012. A comparison plots between simulated and observed runoff have been prepared for both the calibration (**Figure 7.1**) and validation (**Figure 7.2**) periods on daily and monthly time scale and the corresponding result on performance indicators are shown in **Table 19.** In addition **Figure 8.1** and **Figure 8.2** shows another comparison with the flow duration curve in terms of probability of exceedance during both the periods.

Table 19: Result of calibration and validation

Input cases	**Performance indicators**		
	NASH-R2	5 years Accumulated Difference (mm/day)	Correlation coefficient (RR)
Calibration period			
Gauge (G)	0.742	519.59	0.86
Raw Satellite estimates (RSE)	0.666	-490.88	0.81
Bias corrected Satellite estimates (BCSE)	0.738	205.84	0.85
Validation period			
Gauge (G)	0.812	347.78	0.84
Raw Satellite estimates (RSE)	0.632	-410.37	0.77
Bias corrected Satellite estimates (BCSE)	0.701	278.49	0.78

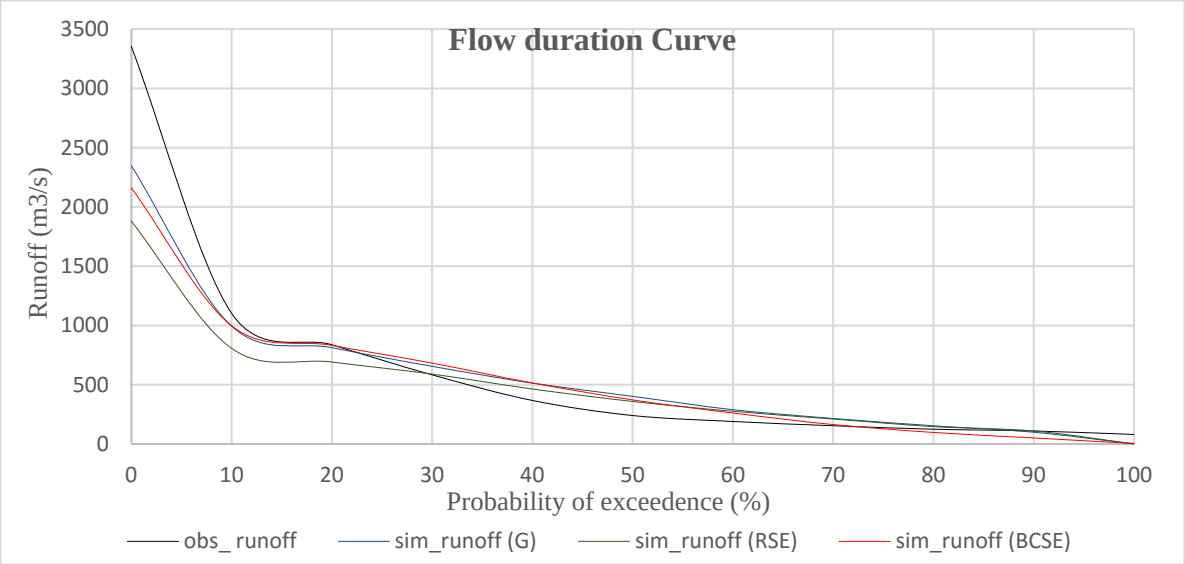


Figure 8.1: Flow duration curve (Calibration period)

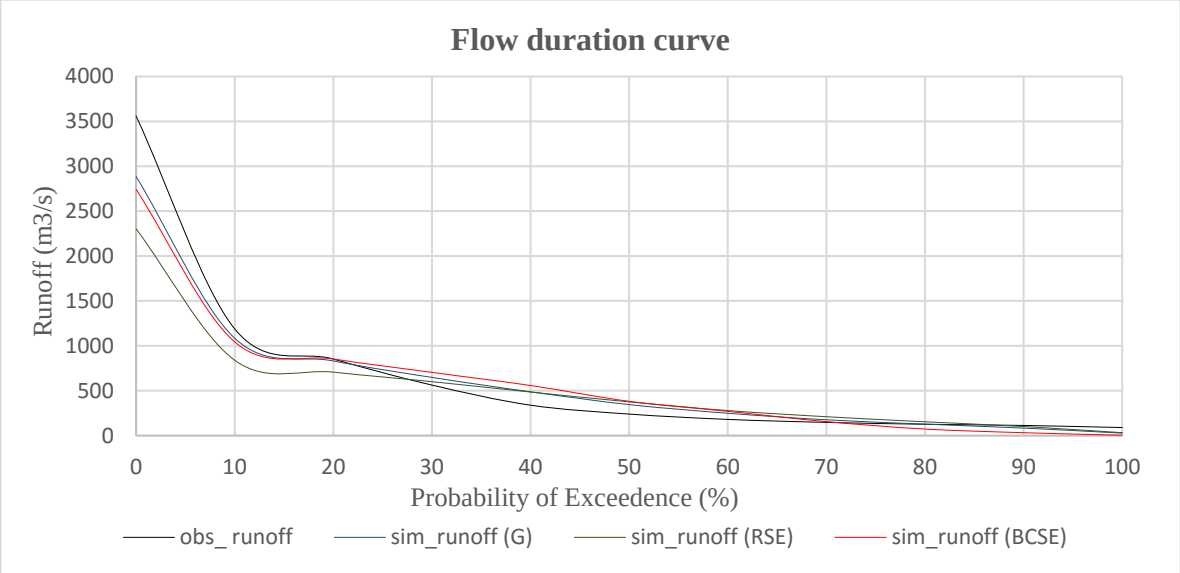


Figure 8.2: Flow duration curve (validation period)

8.2 Discussions

From **Figure 7.1** and **Figure 7.2** we can see that there is a general qualitative match between the observed and simulated runoffs in daily time scale. However, there are varying level of discrepancies between the three input cases in monthly time scale.

8.2.1 Simulation with gauge precipitation (G)

The simulation with gauge precipitation was conducted using 11 gauging stations in the catchment. The observed and simulated hydrographs in **Figure 7.1** and **Figure 7.2** illustrates that the hydrographs are in quite close correspondence with each other in most of the years. There is an exception in the year 2003 of calibration period and year 2011 of validation period that during these years there was significant underestimation in the simulated runoff. However, there is an improvement in the performance of the model in the validation period with increase in R2 from 0.742 to 0.812 and reduction in accumulated difference from 519.59 to 347.78 mm/day with a small reduction in RR from 0.86 to 0.84. The accumulated difference is positive indicating overestimation in simulated runoff. The exceedance probabilities presented by the flow duration curve (**Figure 8.1, Figure 8.2**) of both the periods show that overestimation of flow has occurred mostly during medium flows in both the periods while, there is underestimation of high flows. The simulated hydrograph during low flow periods match relatively well with slight underestimation.

8.2.2 Simulation with raw satellite estimates (RSE) as input

The plots show that the simulated runoff with raw satellite estimates (without bias correction) consistently underestimated the flow in most years during both calibration and validation periods. The negative accumulated difference from the water balance criteria also reveals that there is underestimation of simulated runoff. The R2 value reduced from 0.666 to 0.632 in the validation period with slight improvement in the accumulated difference. Looking at the exceedance probability in the flow duration curves (**Figure 8.1, Figure 8.2**) of both the periods, there is a clear indication that significant underestimation has occurred at high observations while simulations match relatively well at low flows with slight overestimation in the medium ranges. The result is quite consistent to the earlier comparison made between the gauge precipitation and satellite estimates and it is a clear indication that there is a need to adjust the biases in the satellite estimates.

8.2.3 Simulation with bias corrected satellite estimates (BCSE) as input

The biases in the satellite estimates were corrected by the ratios of the annual average sums of satellite estimates and gauge records. The simulated and observed hydrographs presented in **Figure 7.1** and **Figure 7.2** illustrates that the simulation with the BCSE yielded much better correspondence with the observed flows when compared to simulation with RSE. The improvement in the values of performance indicators in **Table 19** with the change in input case

from RSE to BCSE during both calibration and validation periods are also an indication of improved relation between observed and simulated hydrographs. However, the R2 and RR values with BCSE as input, reduced from 0.738 and 0.85 during calibration to 0.701 and 0.78 in validation period respectively. The accumulated difference is positive indicating overestimation in simulated runoff and this value has been increased from 205.84 mm/day during calibration to 278.47 mm/day during validation period. The exceedance probabilities presented by the flow duration curve (**Figure 8.1, Figure 8.2**) of both the periods show that overestimation of flow has occurred mostly during medium flows in both the periods while, there is underestimation of high flows. The simulated hydrograph match relative well with slight underestimation during low flow periods.

8.3 Summary of Result and Discussion

The comparison between simulated and observed hydrographs, flow duration curves, and performance indicators show that there is very good agreement between the simulated and observed runoffs during calibration period and overall best performance was achieved using gauge precipitation. The model performed reasonably well with BCSE and better with gauge precipitation during validation periods also. The calibration and validation results indicate that SHyFT performs well to evaluate the satellite precipitation products for hydrological prediction in the catchment.

9 Runoff series for an ungauged Catchment

9.1 Location of ungauged catchment

As shown in **Figure 9.1** below, Dangchhu is an ungauged catchment with an area of 434 sq.km. It is located in the central east of the basin between elevations 1807 and 5199 m.a.s.l.. The name of the catchment is given according to the name of the river Dangchhu which flows approximately about 60 km from its origin until the confluence with the main river Punatsangchhu. The outlet point as shown in the figure is tentatively about 20 km upstream of this confluence.

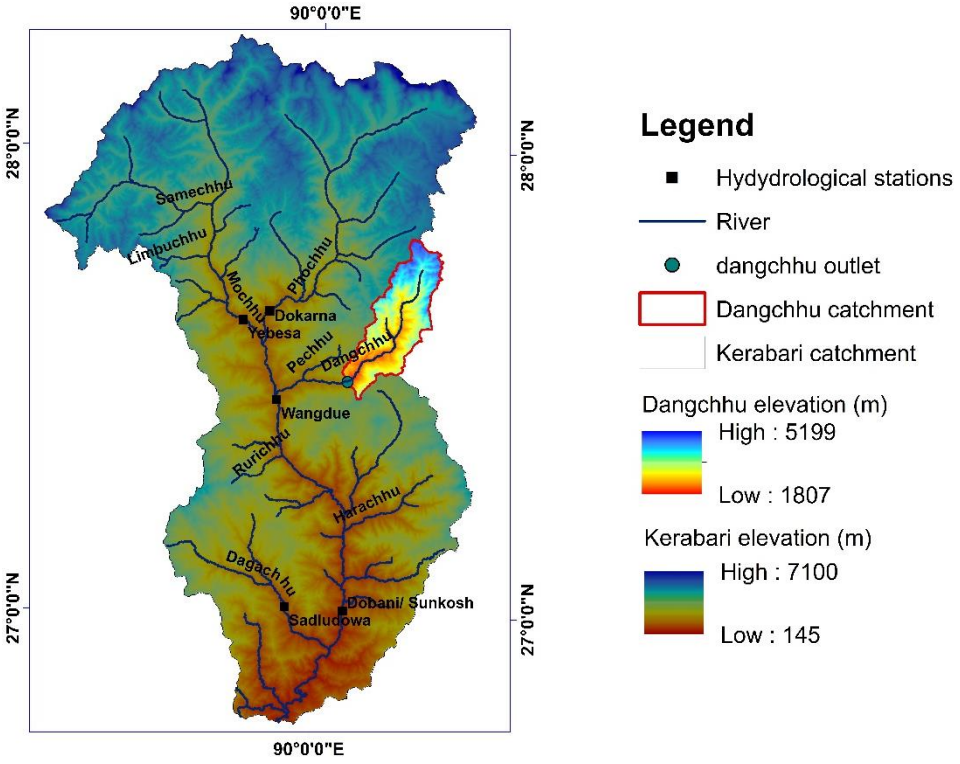


Figure 9.1: Location of Dangchhu Catchment

9.2 Generating runoff series for Dangchhu and Wangdue Catchment

The calibrated model from two of the input cases (gauge precipitation BCSE) was used to simulate the runoff series for a period of five years (2003 to 2007). Subsequent to the simulations, the runoff series from a flow gauging station at Wangdue located just downstream of the confluence of Dangchhu with the main river was transposed to Dangchhu using the ratio of the catchment area to compare with the simulated runoff. In addition, another simulation was performed at the Wangdue flow gauging itself in order to make more realistic comparison.

9.3 Results of simulation

The simulated runoff for an ungauged catchment and at Wangdue flow gauging stations are shown in **Figure 9.2** and **Figure 9.3** in daily and monthly time scale along with flow duration curves. Some statistics on the performance indicators are also given in **Table 20.**

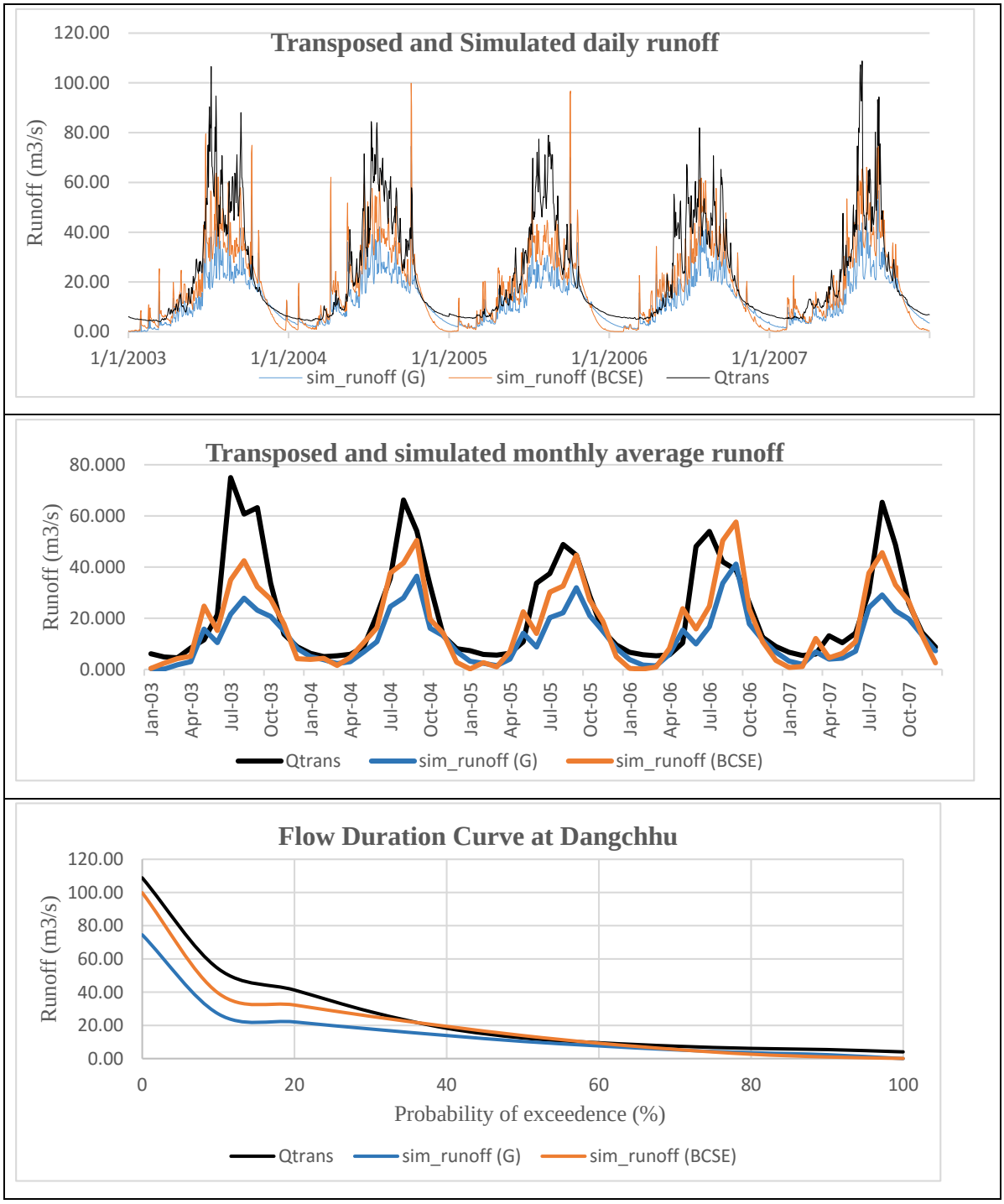


Figure 9.2: simulated hydrographs and Flow duration curve (ungauged catchment)

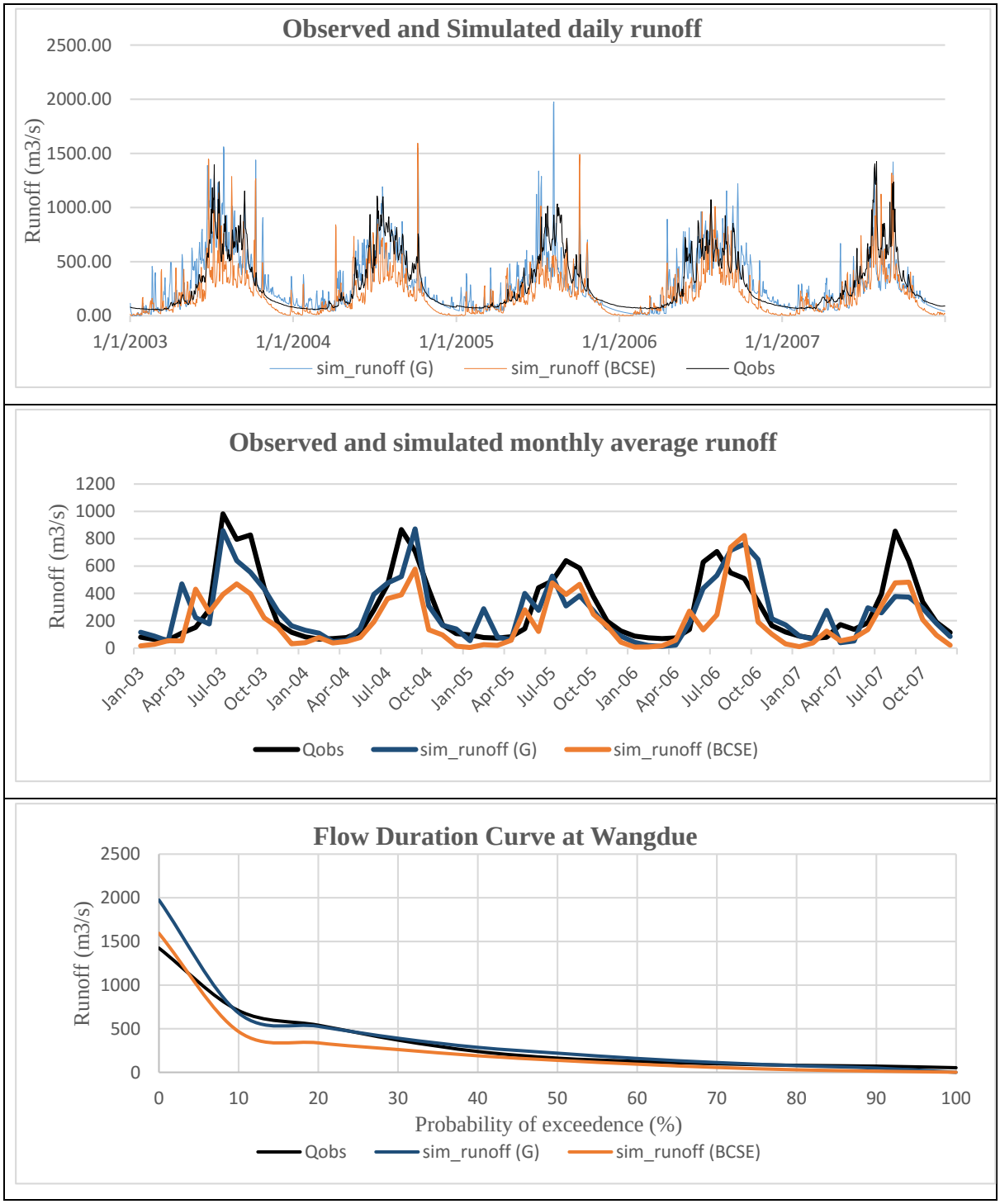


Figure 9.3: Simulated hydrographs and flow duration curve for Wangdue catchment

Table 20: Performance indicators for Dangchhu and Wangdue

Input cases	Performance indicators		
	NASH-R2	5 years Accumulated Difference (mm/day)	Correlation coefficient (RR)
At Dangchhu (ungauged catchment)			
Gauge (G)	-0.4	-3653	0.81
Bias corrected Satellite estimates (BCSE)	0.5	-1804	0.83
G Vs BCSE	0.79	1848	0.98
At Wangdue (gauged catchment)			
Gauge (G)	0.62	390	0.81
Bias corrected Satellite estimates (BCSE)	0.26	-2670	0.73

9.4 Discussion

The calibrated model with three different input cases were subsequently validated and was used to generate runoff series for an ungauged catchment (Dangchhu) for a period of 5 years (2003 to 2007). Flow series from a nearby gauging station at Wangdue was transposed to generate another flow series (Qtrans) for comparision. It is to be noted that transposition of discharge is a traditional approach to predict flow in an ungauged catchment and there are lots of uncertainties associated with spatial heterogeneity (ex. physiographic characteristics of the catchment) between the catchments. However, in the absence of other means of comparison, the BCSE simulated runoff series is compared with the both the gauge simulated hydrograph and transposed series from Wangdue.

Figure 9.2 compares the transposed hydrograph with simulated hydrographs from two input cases on daily and monthly time scale. It can be seen that the simulated hydrograph with gauge precipitation as input consistently underestimated the flow throughout the period whereas, the hydrograph generated with BCSE has good match along the rising limb and recession limb with some underestimation of peaks. The exceedance probability on the flow duration curves also show that both gauge and BCSE simulation have underestimated the high flows which occurred for about 30 to 40% of the time with relatively good match during the medium and low flow period. The above points can be backed by the statics on the performance indicators presented on **Table 20** with high negative accumulated difference. The model performance with BCSE was at an average with R2 value of 0.5 and a good correlation coefficient of 0.83.

Whereas, with gauge precipitation as input, the R2 value of -0.4 indicates that the transposed runoff is better than the simulated runoff. On the other hand, the comparison of BCSE hydrograph with the gauge simulated hydrograph showed a good R2 and RR value of 0.79 and 0.98 but there was overestimation of flow by the satellite data as indicated by the accumulated difference of 1848 mm/day over the five-year period.

Additional comparison was made by simulating flow using the BCSE and gauge precipitation at the Wangdue flow gauging station itself. The performance indicators show that the simulation with gauge precipitation is better with R2 of 0.62, RR of 0.81 and low positive accumulated difference of 390 mm/day. whereas with the BCSE as input, the performance of the model reduced with low R2 (0.26) and high negative accumulated difference (2670 mm/day). This indicates that satellite underestimated flow at wangdue. However, BCSE and gauge simulated hydrographs was matching relatively well for the last three out of five years simulated. The exceedance probabilities on the duration curve show that gauge simulation was able to capture intermediate and low flows occurring for about 90% of the time and did not capture well the high flows which occurred for about 10% of the time. On the other hand, the BCSE simulated duration curve show a clear underestimation of runoff for more than 95% of the time.

10 Conclusions and Recommendations

10.1 Conclusions

The study “Runoff modelling for Bhutan using satellite data” was conducted as a case study for Kerabari catchment in the Punatsangchhu basin of Bhutan. The basin is equipped with meteorological stations mostly in the central and southern regions and flow gauging stations are mainly on the main river Punatsangchhu, and few other major tributaries. The northern regions in the basin has very limited or no meteorological coverage and most of the east west flowing rivers have remained ungauged until today. In regions where ground based measurements are sparse or not available, satellite precipitation products can be of key source for rainfall data. Therefore, the main purpose of this study was to evaluate the gridded satellite precipitation product for runoff modelling applications in data scarce region and to generate a time series of discharge for an ungauged catchment in the basin.

In regard to above, TRMM3B42 version 7 satellite based precipitation estimates was selected based on the validation and comparison studies carried out by previous researchers on the use of such products in Bhutan and in similar Himalayan regions. The TRMM3B42V7 daily estimates with a spatial resolution of 0.25° x 0.25° was compared and evaluated on daily, monthly and annual time scales against the ground precipitation measurements through visual assessment of plots and by using some statistical methods such as Nash-Sutcliffe Coefficient of Efficiency, root mean square difference, estimation bias, correlation coefficients and so on.

On comparison of daily datasets, the R2 value ranged from -2.45 to 0.32 and was negative in most of the pixels which revealed that satellite did not estimate the precipitation well on the daily basis. There was a small trend of reducing R2 value with increasing elevation. Low positive value of RR (0.34 to 0.58) showed week to average correlation of satellite data with the gauged rainfall data. In contrast, when comparison was made on monthly and annual scales, better results were obtained with R2 and RR values ranging from 0.73 to 0.98 and 0.71 to 1 respectively in most pixels compared. The monthly and annual data sets performed equally well with no trend of increasing or decreasing R2 with elevation. The biases in the datasets ranged between -31.54% to as high as 84.12%. Negative biases mainly occurred in elevations lower than 2060 m with high MAD and high RMSD indicating that satellite underestimated high precipitation in the southern regions of the basin. However, an average CPOD_S range of 0.6 to 0.89 on daily data, 0.98 to 1 on monthly and 1 on annual data showed a good detection efficiency of precipitation by satellite. In contrast, the case was opposite in the central and

northern parts of the basin with positive biases, low MAD and low RMSD demonstrating over estimation. Low performance by satellite data on daily timescale can be attributed to difference in the observation time of 4.5 hours between the satellite and gauges. This difference can be brought down to 1.5 hours and some improvements in R2 and RR values can be achieved if three hourly satellite data were used and manually accumulated at 3:00 UTC. The basin does not have ground gauges in northern regions, in the central regions the gauges are inadequately distributed and most of them are located in the valleys. This was a case of poor spatial coverage by the gauges which do not capture well the spatial variability and underestimated the orographically induced precipitations that occurs in mountainous topography. This was the main reason for satellite data showing higher precipitation amounts compared to gauges in the central and northern regions of the basin. Overall, the comparison of satellite data with that of the ground gauges revealed a general qualitative match in terms of timing and with quantitative differences. Therefore, there was a requirement to adjust the biasness in the data before it could be used as input for simulating stream flows.

The daily stream flow simulations were carried out for Kerabari catchment using SHyFT. The model was first calibrated for the period 2003 to 2007 to estimate the best set of free parameters for three input cases: [1] gauge precipitation (G), [2] raw satellite estimates (RSE) and [3] bias corrected satellite estimates (BCSE). The calibrated model was subsequently validated for each of the input case for another five-year period from 2008 to 2012.

Amongst the three input cases, simulation with gauge precipitation performed the best with R2 and RR value of 0.742 and 0.86 with positive accumulated difference of 519.59 mm/day during the calibration period. There was improvement in the model performance with increase in the R2 value to 0.812 and reduction in the accumulated difference to 347.78 mm/day however, with some reduction in the correlation coefficient (RR) to 0.84 during the validation period. Bias corrected simulation also performed equally well with R2 and RR value of 0.738 and 0.85 with an accumulative difference of 205.84 mm/day during the calibration. However, during the validation period, the R2 and RR value reduced to 0.701 and 0.78 and accumulated difference increased to 278.49 mm/day. Lowest performance in the simulation was given by the RSE with R2, RR and accumulated difference of 0.666, 0.81 and -490.88 mm/day respectively during the calibration. There was a little reduction of the accumulated difference to -410.37 but both R2 and RR values reduced to 0.632 and 0.77 correspondingly. There was consistent underestimation of simulated runoff in most years with RSE whereas, the BCSE and gauge simulated hydrographs was able to match relatively well with the observed hydrographs in in

terms of time of occurrence and magnitude in most years during both the calibration and validation periods. The exceedance probabilities revealed that all the input cases underestimated high flows with some overestimation of intermediate flows and had relatively good match with low flow observations. Overall it can be concluded that, the RSE cannot be used to predict flows for ungauged catchments in the basin without bias adjustments.

The calibrated model from two input cases (BCSE and gauge data) was used to generate five years (2003 to 2007) runoff for an ungauged catchment (Dangchhu) in the basin. The simulated hydrograph from BCSE for Dangchhu had higher flow estimates compared to gauge simulation because, Dangchhu is a small catchment which is almost entirely located within a single pixel and has only one ground gauge that it is inadequate to represent the catchments precipitation distribution. The precipitation gauge is close to the outlet point and the nearby gauges are very far in the south compared to the distance between the centers of the nearby pixels on the four sides of the catchment. When comparing simulations from both the inputs with the transposed runoff from Wangdue, the simulated runoff from BCSE was found to be more close with the transposed hydrograph with R2 of 0.5 and significant underestimation was experienced with gauge precipitation with R2 of -0.4. This is an indication that the transposed runoff is better than the simulated runoff though there are physiographical uncertainties associated in the method of transposing discharge from one catchment to another. In case of wangdue flow gauging station, located downstream of Dangchhu, there are more gauges but, the spatial distribution of gauges was not uniform. Most of meteorological stations were clustered in lower altitudes in the valleys close to the populous areas and did not provide representative information on the climatic conditions in the headwater regions. As a result, there was some underestimation of simulated runoff with gauge precipitation as input. Nevertheless, it was able to capture the pattern and the peaks of the observed hydrographs for most of the years. The performance indicators with R2, RR and accumulated difference of 0.62, 0.81 and 390 mm/day respectively, showed that the efficiency of the model in predicting the flow at Wangdue improved significantly compared to that at Dangchhu with gauged inputs. Whereas, the input with BCSE performed well enough to match the peaks in three out of five years but there was underestimation of rising and falling limbs throughout the period when compared with observed hydrograph.

It was observed that the performance of the model greatly depended on the quality and distribution of ground gauges and the spatial resolution of the satellite data. As in the current study catchment, the spatial distribution of gauges was not uniform and most of them were

clustered at central and southern regions and in the valleys of the catchment whereas, the northern region suffered from sparse or no ground network. In such a case, the gauges had the tendency to underestimated the orographically enhanced precipitation occurring in the mountainous terrain in addition to its inability to represent the climatic conditions in the headwater regions of the catchment. This statement can be backed by the underestimation in simulated runoff with gauge precipitation as input for all the three catchments as already discussed above. When it comes to the spatial resolution of satellite data, there was high variability of precipitation within a single pixel and average value for each pixel of size 27.5 x 27.5 km2 provided by the satellite estimates was not a good representation of precipitation distribution for an area of such size especially in mountainous topography. In addition, the number of gauges falling inside each pixel or in the nearby pixel was also an issue because, there was a needed to adjust the biasness in the satellite estimates with reference to the accumulation in the gauges which are already inadequate. Therefore, in areas like Dangchhu and other upstream areas where there are no gauges, the correctness of the adjustments made in bias was a big question and hence the simulated runoffs for Dangchhu with both input case is also doubtful and cannot be used further.

It was also perceived that, the performance by the simulated hydrographs with both gauge and BCSE as input increases with the increase in the size of the catchment and as we move the outlet towards the south. However, in both the input cases there was underestimation of peaks and mostly by the simulated hydrographs with satellite precipitation as input. Maximum performance was achieved at the outlet of Kerabari catchment itself. This is because larger catchments had better coverage of ground meteorological stations especially in the central and southern regions and more pixels were also available inside the catchment in addition to relatively good bias adjustments possible. Furthermore, the choice of interpolation routine (IDW) could be one reason for giving low response for catchments where there are no rain gauges. This is because precipitation needs to be interpolated from the surrounding gauges, but as the distance increases the weights get negligible and hence underestimates precipitation.

Lastly it can be summarized by saying that the performance of the selected product in predicting hydrology in Bhutan was poor especially for catchments located at higher altitudes where the ground network is sparse and it was good for catchments located at lower altitudes where relatively good meteorological network was available. The results were purely dependent on the quality and spatiotemporal resolution of ground gauges used to evaluate satellite estimates and to adjust the biasness in addition to the resolution of the satellite data

itself. Since the product selected for study could not be used independently without bias correction therefore, other satellite data with much finer resolution should be assessed.

10.2 Recommendations for further study

Based on the conclusions made above, following are some of the recommendations suggested for further study.

1. Only one satellite product (TRMM3B42 daily estimates with a spatial resolution of 0.25° x 0.25°) was used in this study. Validation and analysis with other products with much finer time and spatial resolutions are recommended to check if bias adjustments can be avoided.
2. Similar studies can be performed in the adjacent catchments for comparison and validation of results from this study.
3. More realistic method for correction of biasness in the precipitation estimates should be studied for use considering seasonal, local and regional variations in the precipitation pattern. One way of finding the correction factor could be the correction factor calculated by the model during calibration with satellite data without correction of bias as input.
4. The melt release from the saturated snow was worked out in the snowmelt routine based on the Gamma distributed snow depletion curve. This can be changed to HBV - snow routine in SHyFT and the results can be compared.
5. The performance of the model could be checked by changing the interpolation routine from IDW to Bayesian kriging or other methods.

References

Bajracharya, S., Shrestha, M., Mool, P., & Thapa, R. (2010a). *Validation of satellite rainfall estimation in the summer monsoon dominated area of the Hindu Kush-Himalayan region.* Paper presented at the Proceedings of 10th International Symposium on High Mountain Remote Sensing Cartography.

Bajracharya, S., Shrestha, M., Mool, P., & Thapa, R. (2010b). *Validation of satellite rainfall estimation in the summer monsoon dominated area of the Hindu Kush Himalayan region.* Paper presented at the Proceedings of 10th International Symposium on High Mountain Remote Sensing Cartography.

Bajracharya, S. R., Palash, W., Shrestha, M. S., Khadgi, V. R., Duo, C., Das, P. J., & Dorji, C. (2015). Systematic Evaluation of Satellite-Based Rainfall Products over the Brahmaputra Basin for Hydrological Applications. *Advances in Meteorology, 2015*.

Bajracharya, S. R., Shrestha, M. S., & Shrestha, A. B. (2014). Assessment of high-resolution satellite rainfall estimation products in a streamflow model for flood prediction in the Bagmati basin, Nepal. *Journal of Flood Risk Management*, n/a-n/a. doi:10.1111/jfr3.12133

Bárdossy, A. (2007). Calibration of hydrological model parameters for ungauged catchments. *Hydrology and Earth System Sciences Discussions, 11*(2), 703-710.

Beldring, S., & Voksø, A. (2011). <*Climate change impact on flow regimes of rivers in bhutan and possible consequences fro hyddropower development* (4). Retrieved from http://webby.nve.no/publikasjoner/report/2011/report2011_04.pdf

Dingman, S. L. (1994). *Physical hydrology*: Waveland press.

Duncan, J. M., & Biggs, E. M. (2012). Assessing the accuracy and applied use of satellite-derived precipitation estimates over Nepal. *Applied Geography, 34*, 626-638.

Ebert, E. E. (2007). Methods for verifying satellite precipitation estimates *Measuring precipitation from space* (pp. 345-356): Springer.

Ebert, E. E., Janowiak, J. E., & Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bulletin of the American Meteorological Society, 88*(1), 47-64.

Fakhruddin, S. (2015). Development of Flood Forecasting System for the Wangchhu River Basin in Bhutan. *Journal of Geography and Geology, 7*(2), P71.

Gruber, A., & Levizzani, V. (2008). Assessment of Global Precipitation Products A project of the World Climate Research Programme Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel.

Islam, M. N., Das, S., & Uyeda, H. (2010). Calibration of TRMM derived rainfall over Nepal during 1998–2007. *Open Atmos. Sci. J, 4*, 12-23.

Khandu, Awange, J., & Forootan, E. (2015). An evaluation of high-resolution gridded precipitation products over Bhutan (1998–2012). *International Journal of Climatology*.

Kidd, C., Levizzani, V., Turk, J., & Ferraro, R. (2009). Satellite Precipitation Measurements for Water Resource Monitoring. *JAWRA Journal of the American Water Resources Association, 45*(3), 567-579. doi:10.1111/j.1752-1688.2009.00326.x

Krakauer, N. Y., Pradhanang, S. M., Lakhankar, T., & Jha, A. K. (2013). Evaluating satellite products for precipitation estimation in mountain regions: A case study for Nepal. *Remote Sensing, 5*(8), 4107-4123.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The tropical rainfall measuring mission (TRMM) sensor package. *Journal of Atmospheric and Oceanic Technology, 15*(3), 809-817.

McCollum, J. R., Krajewski, W. F., Ferraro, R. R., & Ba, M. B. (2002). Evaluation of Biases of Satellite Rainfall Estimation Algorithms over the Continental United States. *Journal of Applied Meteorology, 41*(11), 1065-1080. doi:doi:10.1175/1520-0450(2002)041<1065:EOBOSR>2.0.CO;2

Shrestha, M., Artan, G. A., Bajracharya, S. R., & Sharma, R. (2008). Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin. *Journal of Flood Risk Management, 1*(2), 89-99.

Shrestha, M., Rajbhandari, R., & Bajracharya, S. R. (2013). Validation of NOAA CPC_RFE: satellite-based rainfall estimates in the Central Himalayas. *ICIMOD Working Paper*(2013/5).

Su, F., Hong, Y., & Lettenmaier, D. P. (2008). Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. *Journal of Hydrometeorology, 9*(4), 622-640.

Tamrakar, B. (2011). Evaluation of precipitation distribution over Nepal using satellite data and its applications in Hydrological Modelling.

Tamrakar, B., & Alfredsen, K. (2012). Applicability of TRMM satellite data in hydropower planning. *Rentech Symposium Compendium, 2*.

Xue, X., Hong, Y., Limaye, A. S., Gourley, J. J., Huffman, G. J., Khan, S. I., . . . Chen, S. (2013). Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? *Journal of Hydrology, 499*, 91-99.

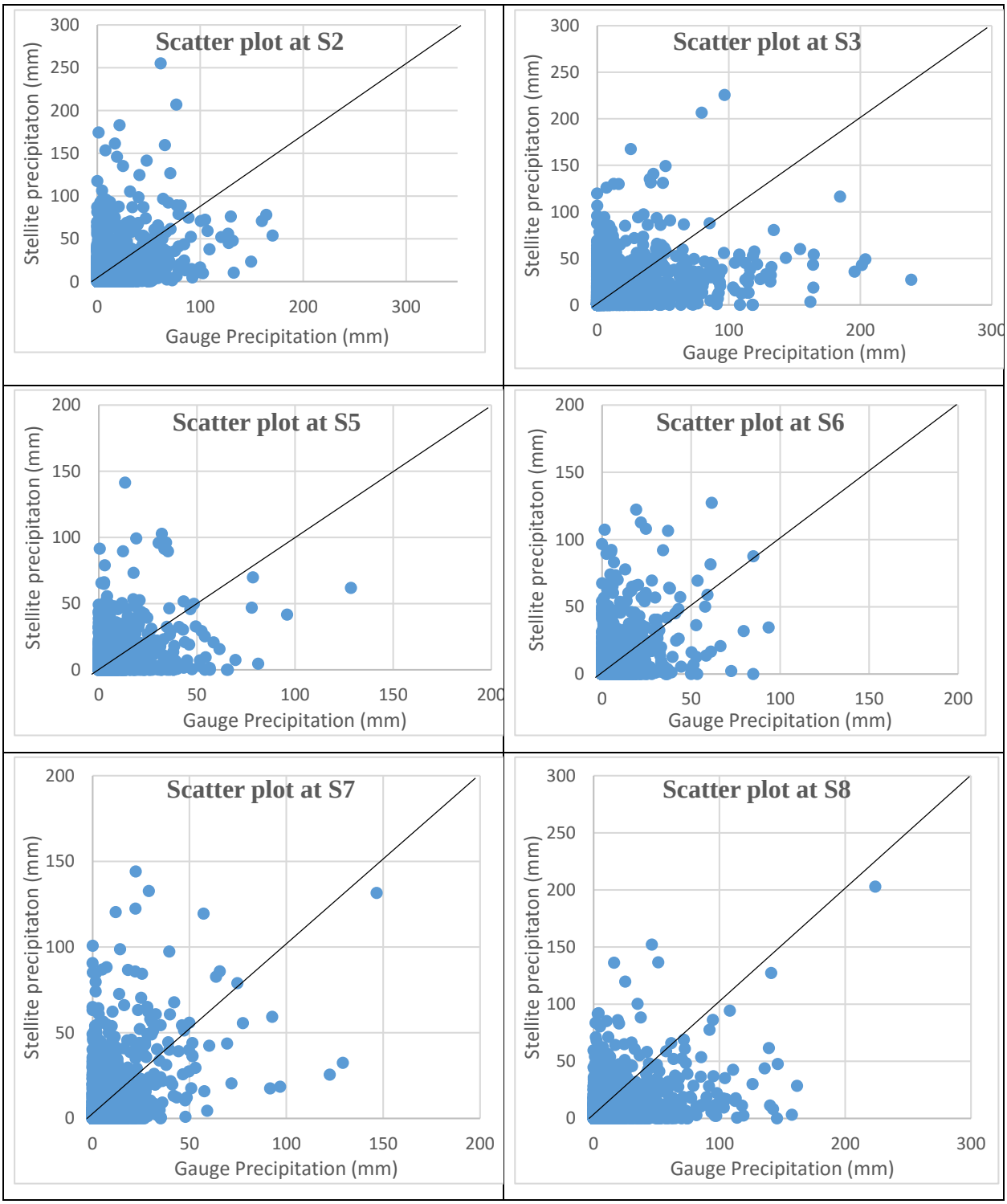
Curriculum and Professional Support Division (CAPSD)1994. A Geography of Bhutan. Course book for class IX and X, Department of School Education, Ministry of Education. Royal Government of Bhutan. Second Edition, 2006

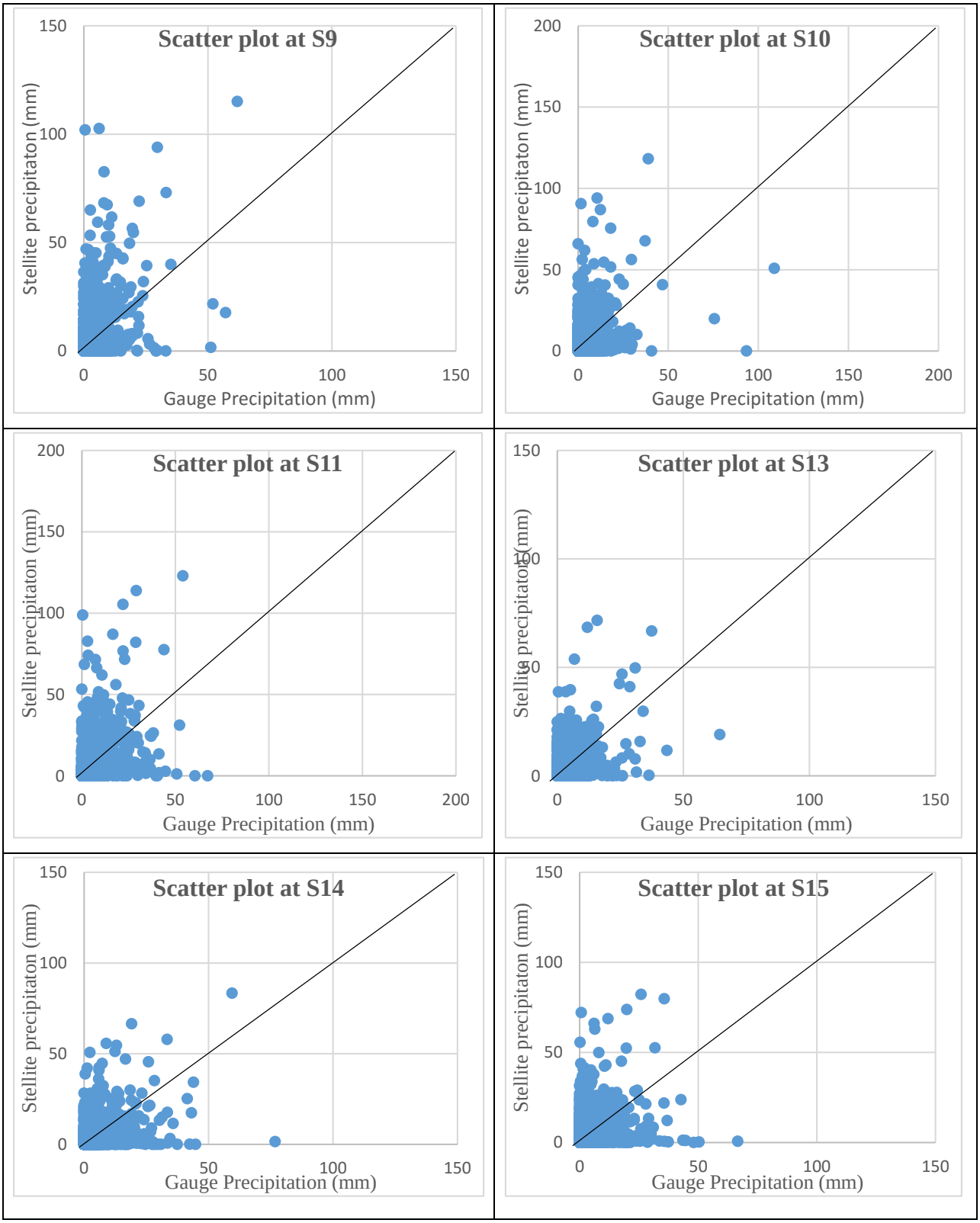
Pearson product-moment correlation coefficient. (2016, April 9). In Wikipedia, The Free Encyclopedia. Retrieved 10:50, April 19, 2016, from https://en.wikipedia.org/w/index.php?title=Pearson_productmoment_correlation_coefficient&oldid=714452201

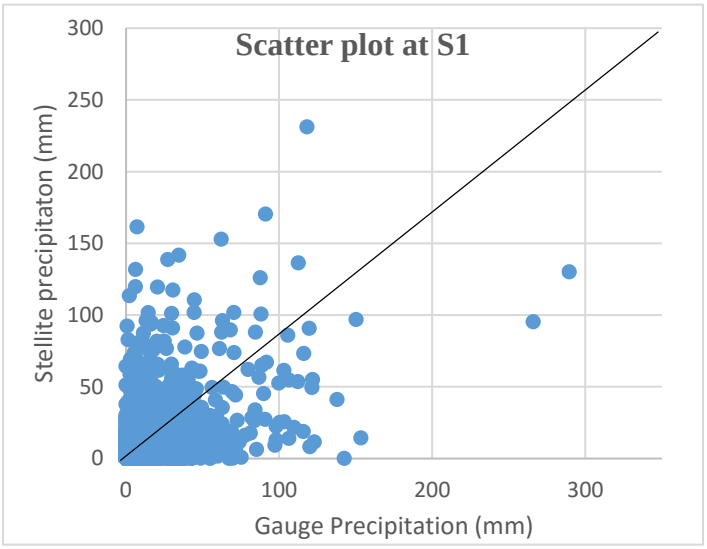
Wikipedia contributors. "Pearson product-moment correlation coefficient." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 9 Apr. 2016. Web. 19 Apr. 2016.

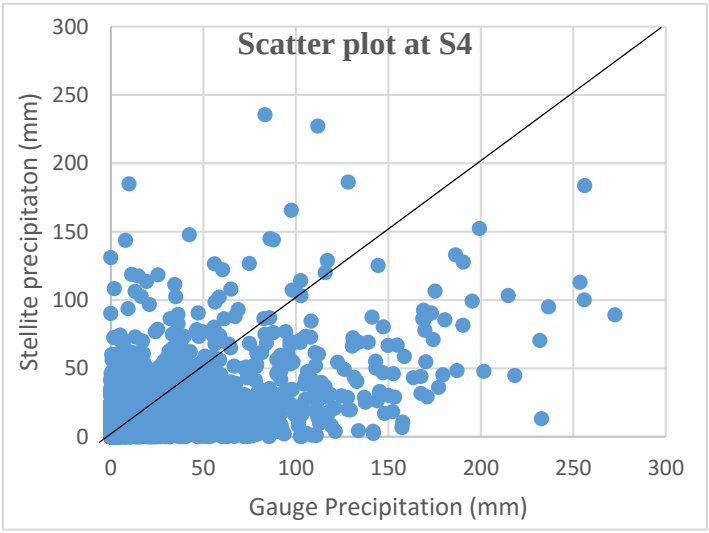
Appendix-1

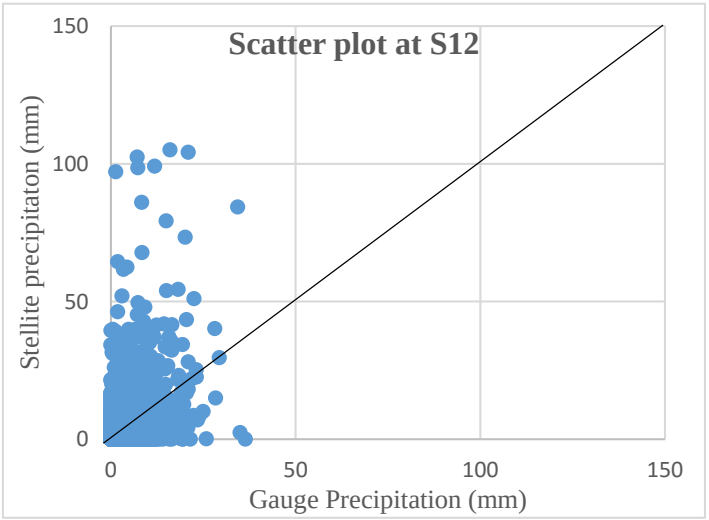
Scatter Plots


Scatter plot at S2
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
200
250
300
Scatter plot at S3
Stellite precipitaton (mm)
Gauge Precipitation (mm)
Scatter plot at S5
Stellite precipitaton (mm)
Gauge Precipitation (mm)
Scatter plot at S6
Stellite precipitaton (mm)
Gauge Precipitation (mm)
Scatter plot at S7
Stellite precipitaton (mm)
Gauge Precipitation (mm)
Scatter plot at S8
Stellite precipitaton (mm)
Gauge Precipitation (mm)


Scatter plot at S9
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
Scatter plot at S10
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
200
Scatter plot at S11
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
200
Scatter plot at S13
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
Scatter plot at S14
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
Scatter plot at S15
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150


Scatter plot at S1
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
200
250
300
0
100
200
300


Scatter plot at S4
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
200
250
300
0
50
100
150
200
250
300


Scatter plot at S12
Stellite precipitaton (mm)
Gauge Precipitation (mm)
0
50
100
150
0
50
100
150

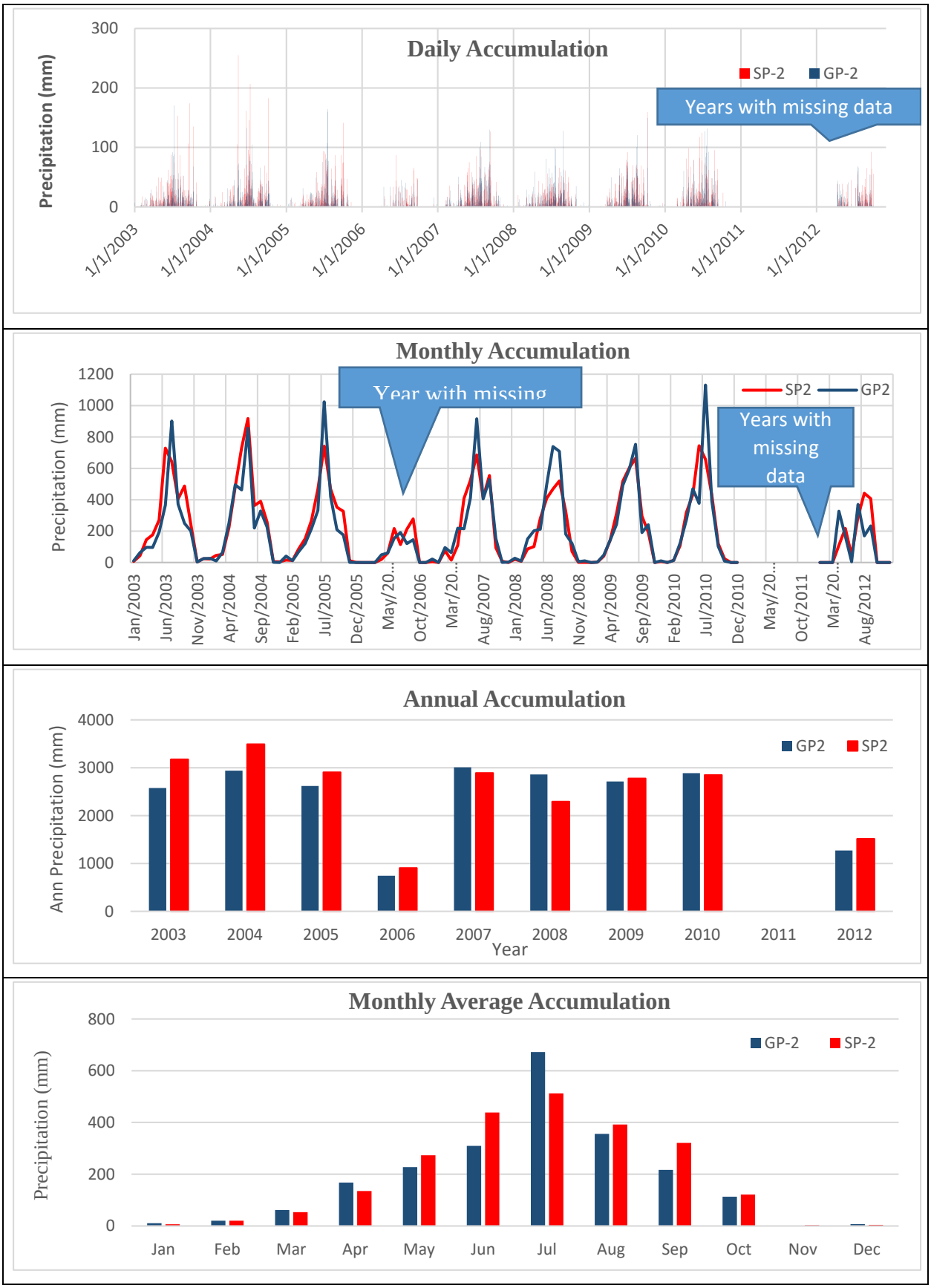
Appendix-2

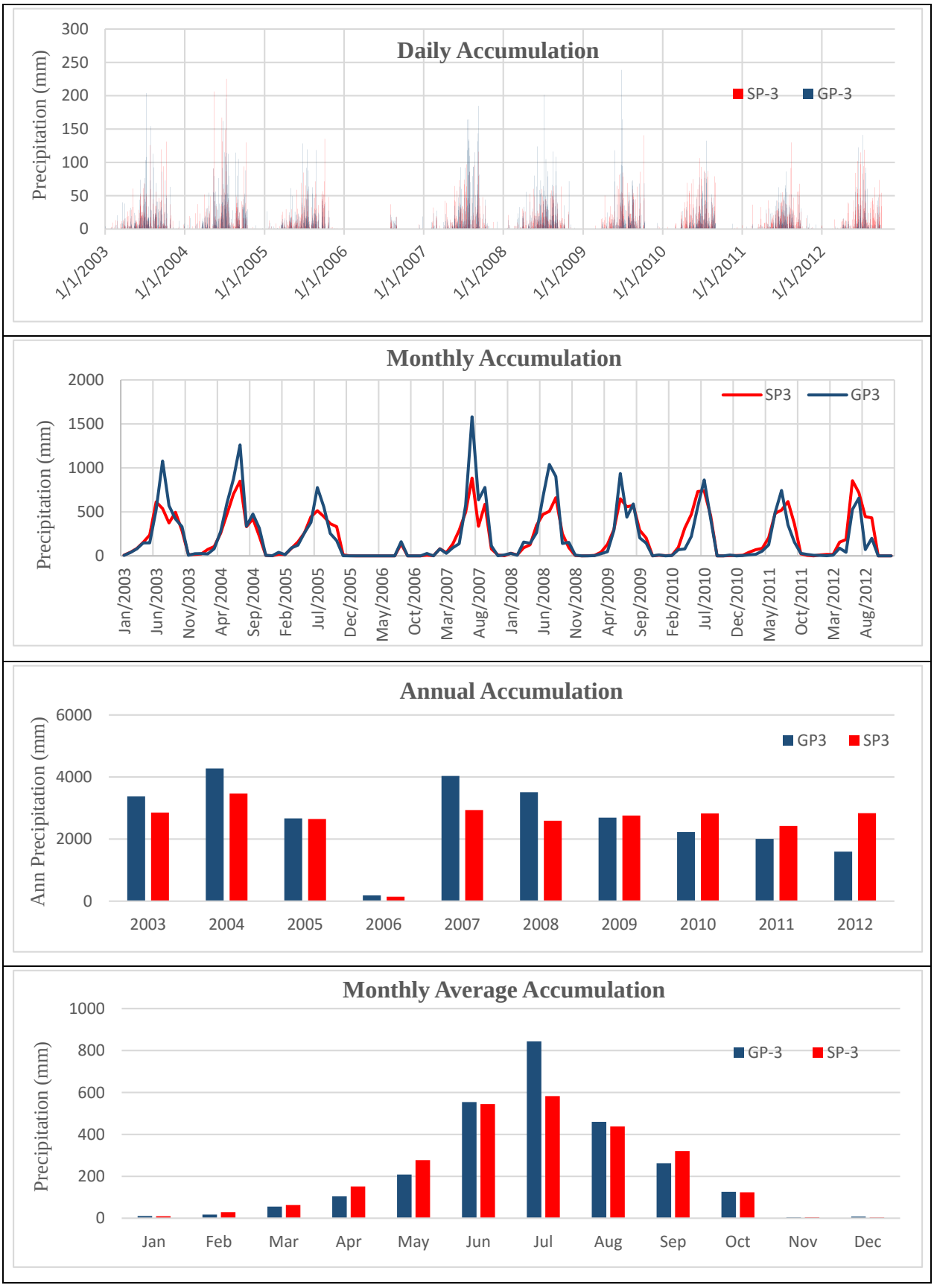
Accumulation Plots

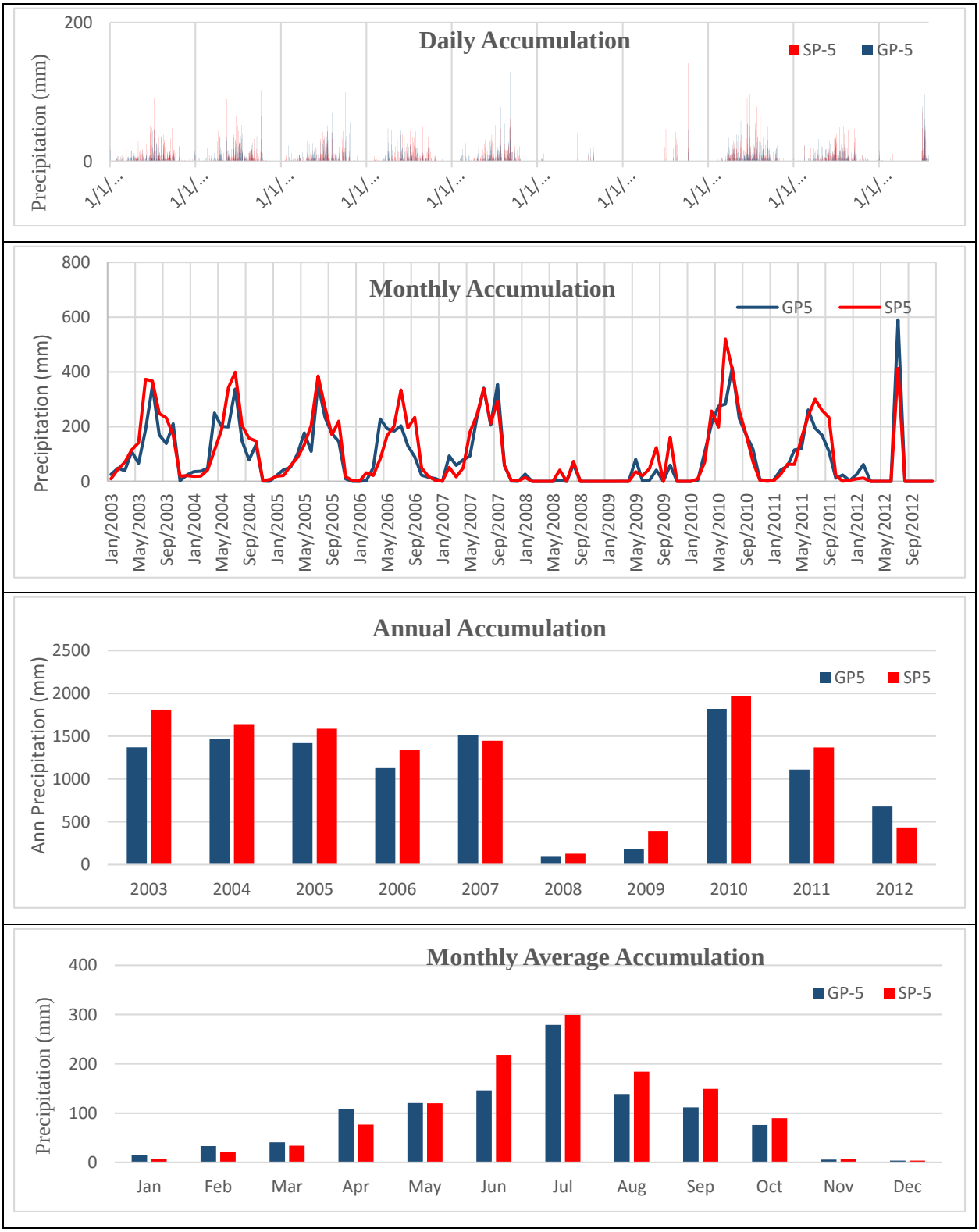
Note:

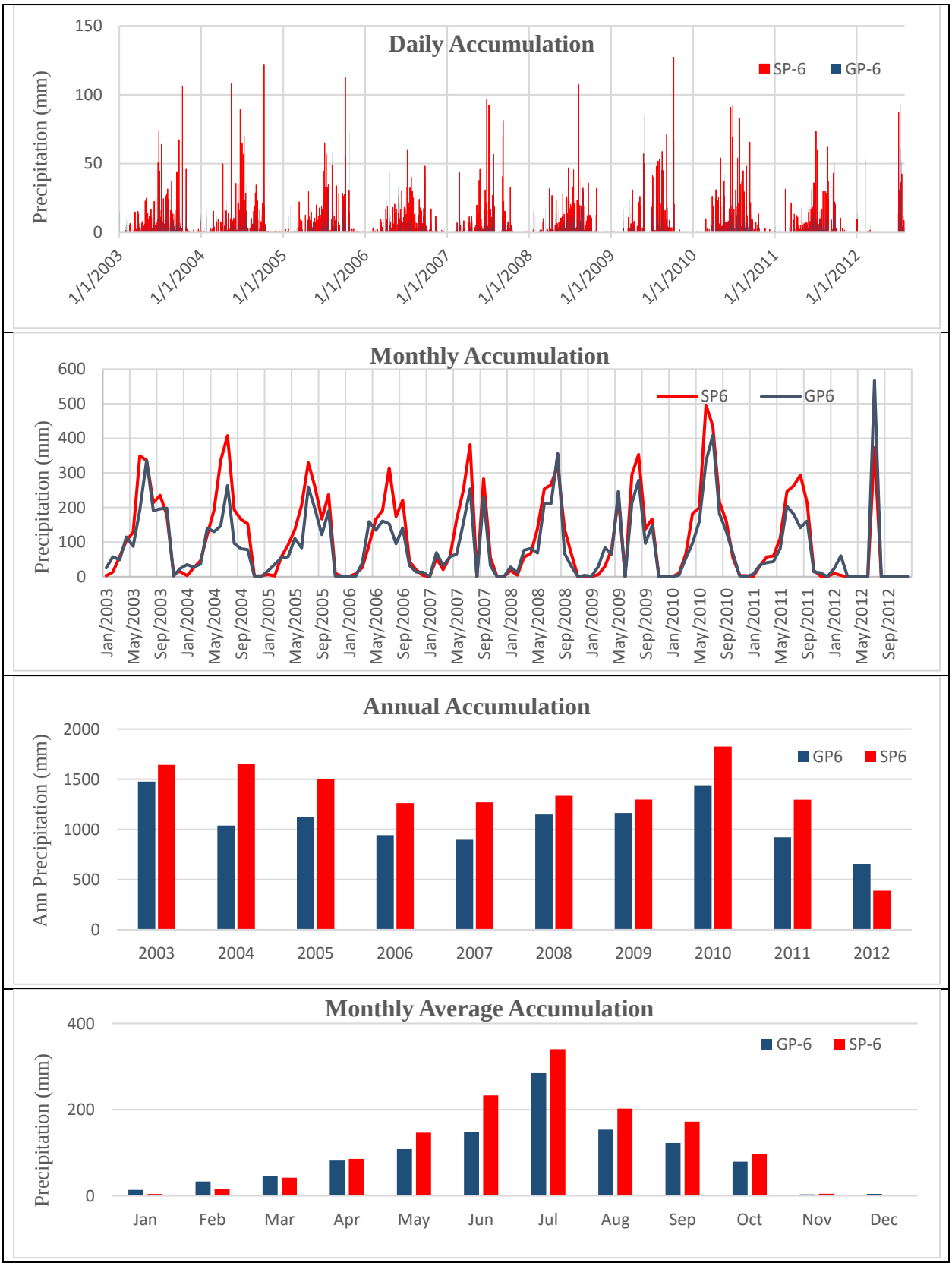
In all the plots presented in this appendix GP is the gauge precipitation, SP is the satellite precipitation estimates from TRMM 3B42 version 7 and the numbers represents the pixel number. For example, GP1 and SP1 represents the accumulation of average gauge and satellite precipitations in pixel number 1.

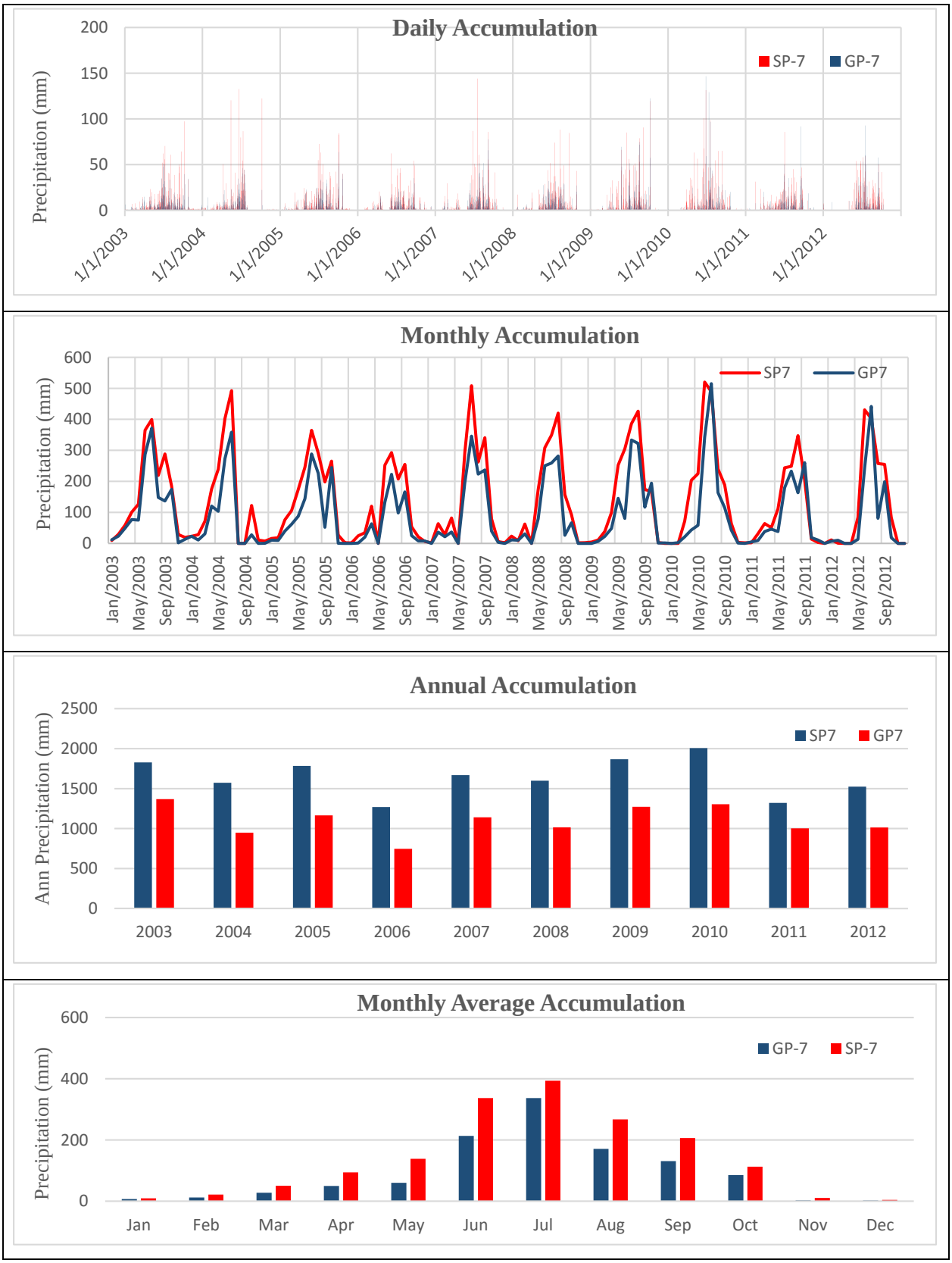
There are some years with missing data and these are clearly visible in the plots. All those years with too much missing data are avoided during the process of bias adjustments but are considered for statistical analysis. This is because we are interested in number of data points so that the ratio of number of days with available from gauge to that of satellite is greater than or equal to 0.8.

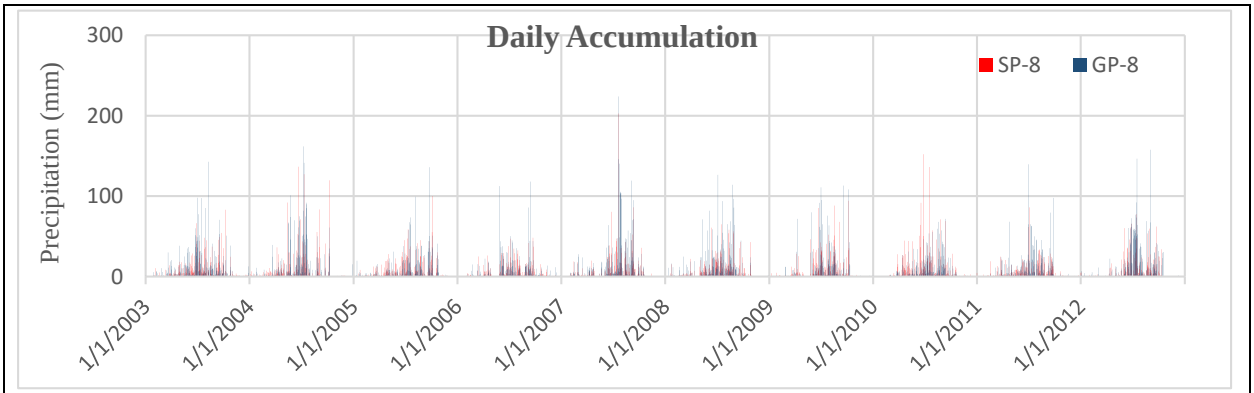

Daily Accumulation
Precipitation (mm)
300
200
100
0
SP-2
GP-2
Years with missing data
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
Precipitation (mm)
1200
1000
800
600
400
200
0
Year with missing
SP2
GP2
Years with missing data
Jan/2003
Jun/2003
Nov/2003
Apr/2004
Sep/2004
Feb/2005
Jul/2005
Dec/2005
May/20...
Oct/2006
Mar/20...
Aug/2007
Jan/2008
Jun/2008
Nov/2008
Apr/2009
Sep/2009
Feb/2010
Jul/2010
Dec/2010
May/20...
Oct/2011
Mar/20...
Aug/2012
Annual Accumulation
Ann Precipitation (mm)
4000
3000
2000
1000
0
GP2
SP2
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Year
Monthly Average Accumulation
Precipitation (mm)
800
600
400
200
0
GP-2
SP-2
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

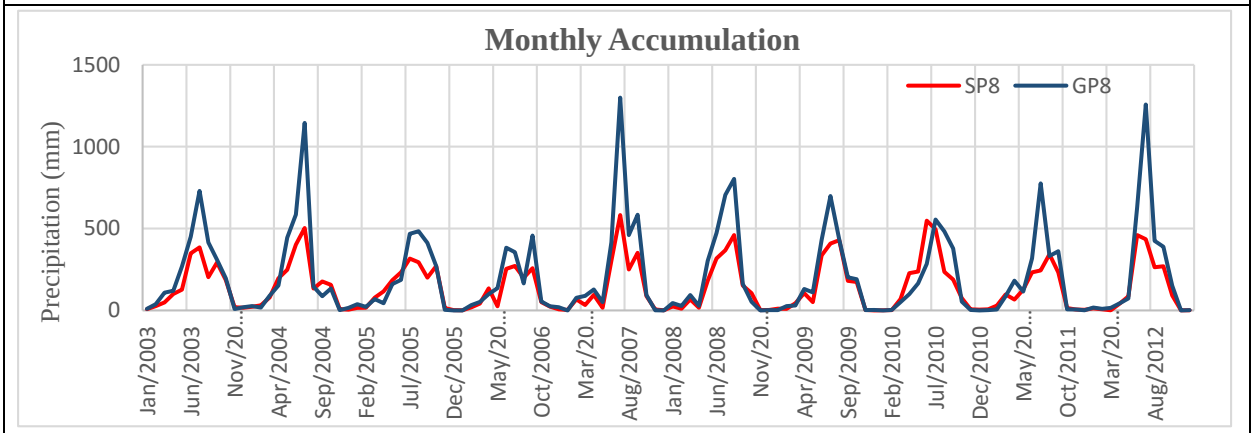

Daily Accumulation
Precipitation (mm)
300
250
200
150
100
50
0
SP-3
GP-3
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
Precipitation (mm)
2000
1500
1000
500
0
SP3
GP3
Jan/2003
Jun/2003
Nov/2003
Apr/2004
Sep/2004
Feb/2005
Jul/2005
Dec/2005
May/2006
Oct/2006
Mar/2007
Aug/2007
Jan/2008
Jun/2008
Nov/2008
Apr/2009
Sep/2009
Feb/2010
Jul/2010
Dec/2010
May/2011
Oct/2011
Mar/2012
Aug/2012
Annual Accumulation
Ann Precipitation (mm)
6000
4000
2000
0
GP3
SP3
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
Precipitation (mm)
1000
800
600
400
200
0
GP-3
SP-3
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

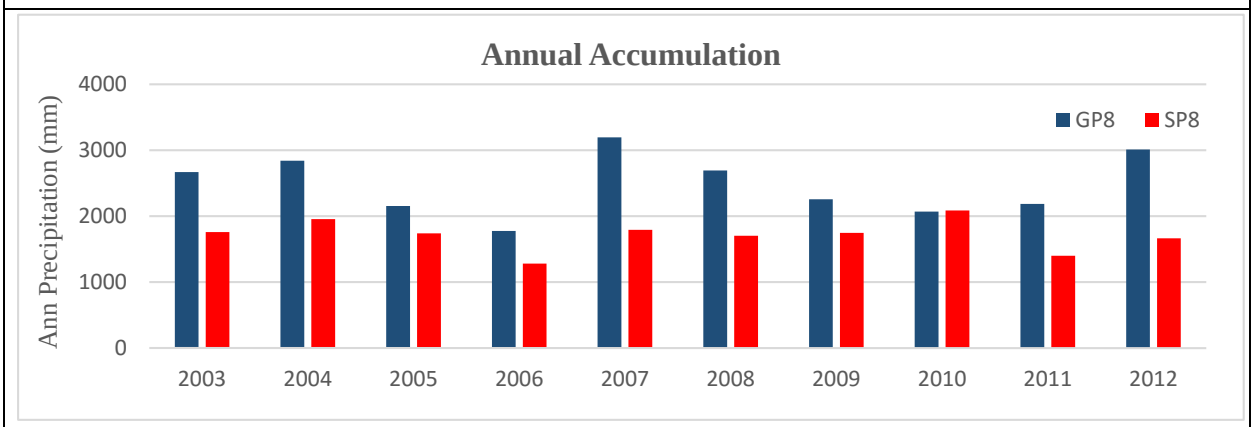

Daily Accumulation
SP-5
GP-5
Precipitation (mm)
200
0
1/1/...
Monthly Accumulation
GP5
SP5
Precipitation (mm)
800
600
400
200
0
Jan/2003
May/2003
Sep/2003
Jan/2004
May/2004
Sep/2004
Jan/2005
May/2005
Sep/2005
Jan/2006
May/2006
Sep/2006
Jan/2007
May/2007
Sep/2007
Jan/2008
May/2008
Sep/2008
Jan/2009
May/2009
Sep/2009
Jan/2010
May/2010
Sep/2010
Jan/2011
May/2011
Sep/2011
Jan/2012
May/2012
Sep/2012
Annual Accumulation
GP5
SP5
Ann Precipitation (mm)
2500
2000
1500
1000
500
0
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
GP-5
SP-5
Precipitation (mm)
400
300
200
100
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

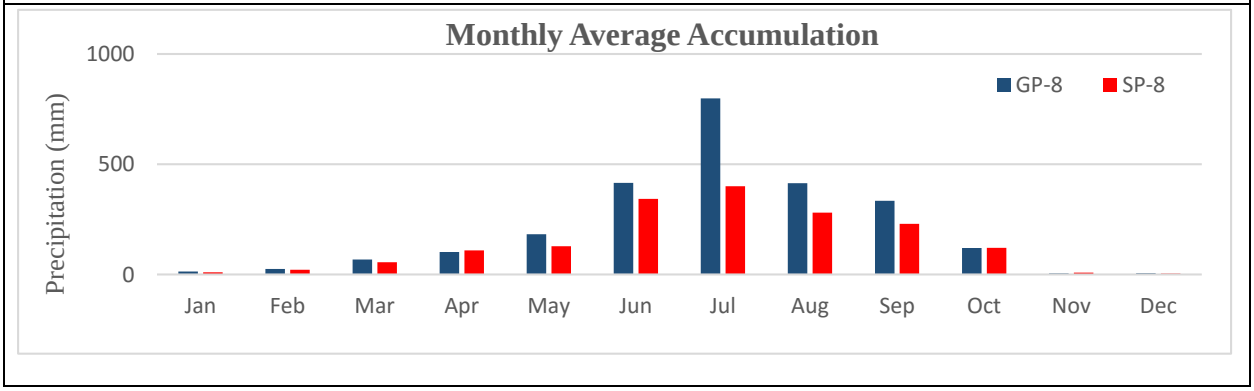

Daily Accumulation
Precipitation (mm)
150
100
50
0
SP-6
GP-6
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
Precipitation (mm)
600
500
400
300
200
100
0
SP6
GP6
Jan/2003
May/2003
Sep/2003
Jan/2004
May/2004
Sep/2004
Jan/2005
May/2005
Sep/2005
Jan/2006
May/2006
Sep/2006
Jan/2007
May/2007
Sep/2007
Jan/2008
May/2008
Sep/2008
Jan/2009
May/2009
Sep/2009
Jan/2010
May/2010
Sep/2010
Jan/2011
May/2011
Sep/2011
Jan/2012
May/2012
Sep/2012
Annual Accumulation
Ann Precipitation (mm)
2000
1500
1000
500
0
GP6
SP6
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
Precipitation (mm)
400
200
0
GP-6
SP-6
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

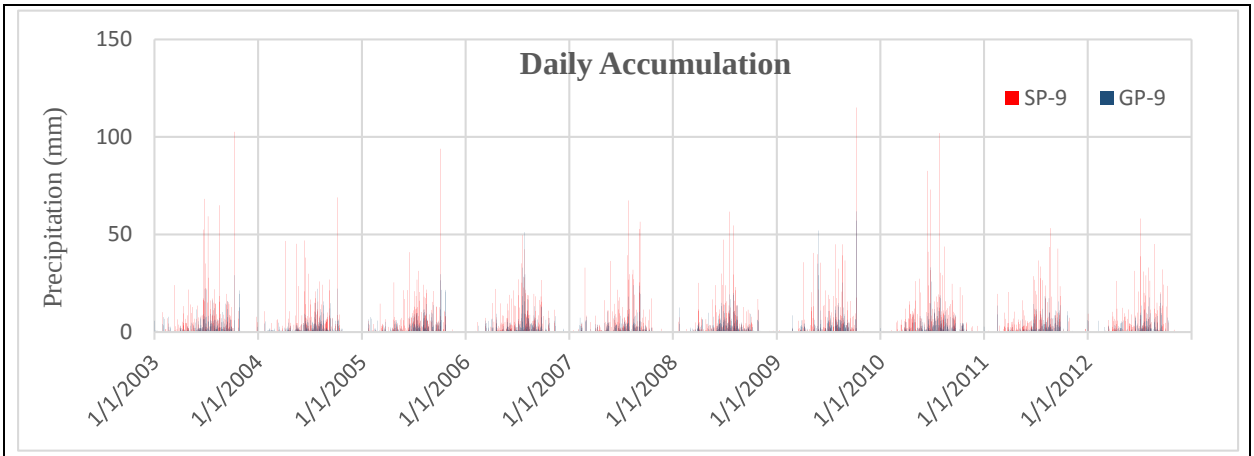

Daily Accumulation
Precipitation (mm)
200
150
100
50
0
SP-7
GP-7
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
Precipitation (mm)
600
500
400
300
200
100
0
SP7
GP7
Jan/2003
May/2003
Sep/2003
Jan/2004
May/2004
Sep/2004
Jan/2005
May/2005
Sep/2005
Jan/2006
May/2006
Sep/2006
Jan/2007
May/2007
Sep/2007
Jan/2008
May/2008
Sep/2008
Jan/2009
May/2009
Sep/2009
Jan/2010
May/2010
Sep/2010
Jan/2011
May/2011
Sep/2011
Jan/2012
May/2012
Sep/2012
Annual Accumulation
Ann Precipitation (mm)
2500
2000
1500
1000
500
0
SP7
GP7
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
Precipitation (mm)
600
400
200
0
GP-7
SP-7
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

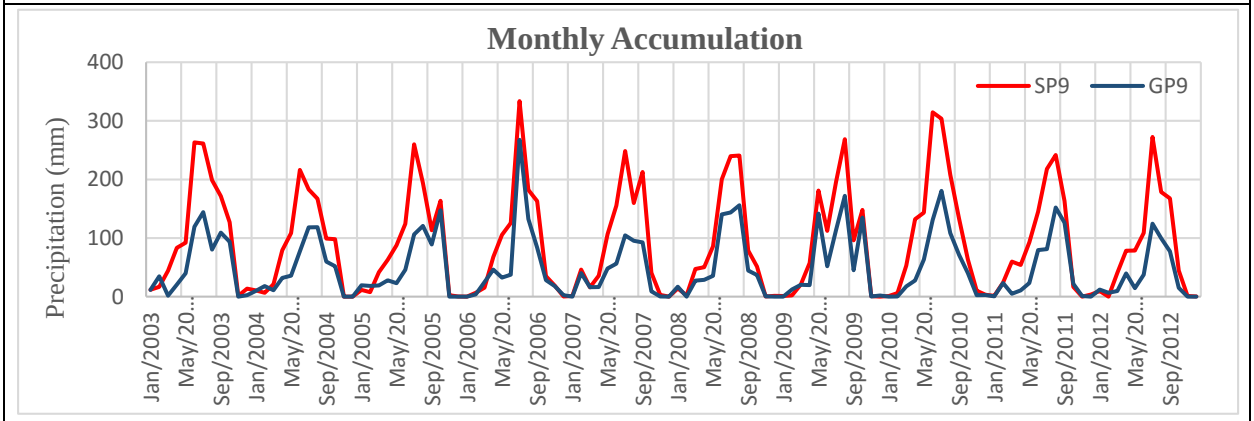

Daily Accumulation
Precipitation (mm)
300
200
100
0
SP-8
GP-8
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012

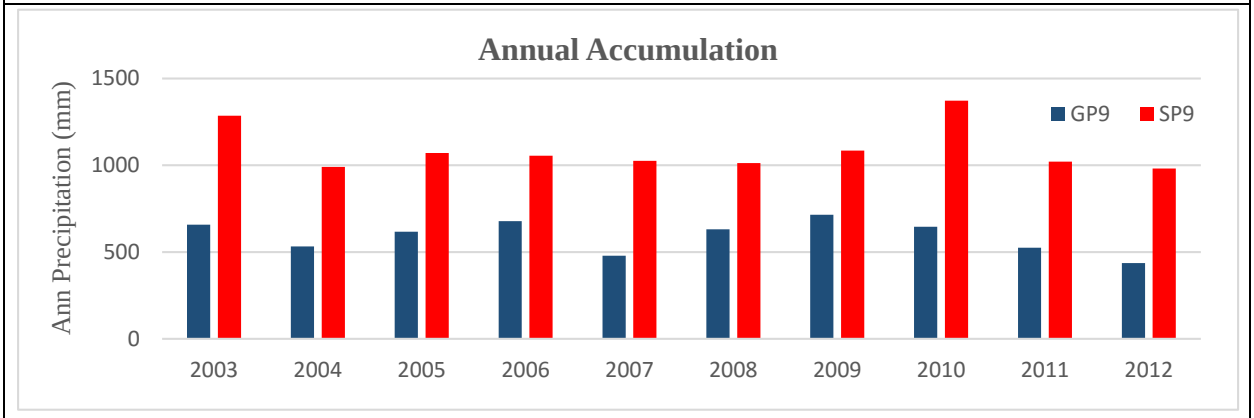

Monthly Accumulation
Precipitation (mm)
1500
1000
500
0
SP8
GP8
Jan/2003
Jun/2003
Nov/20...
Apr/2004
Sep/2004
Feb/2005
Jul/2005
Dec/2005
May/20...
Oct/2006
Mar/20...
Aug/2007
Jan/2008
Jun/2008
Nov/20...
Apr/2009
Sep/2009
Feb/2010
Jul/2010
Dec/2010
May/20...
Oct/2011
Mar/20...
Aug/2012

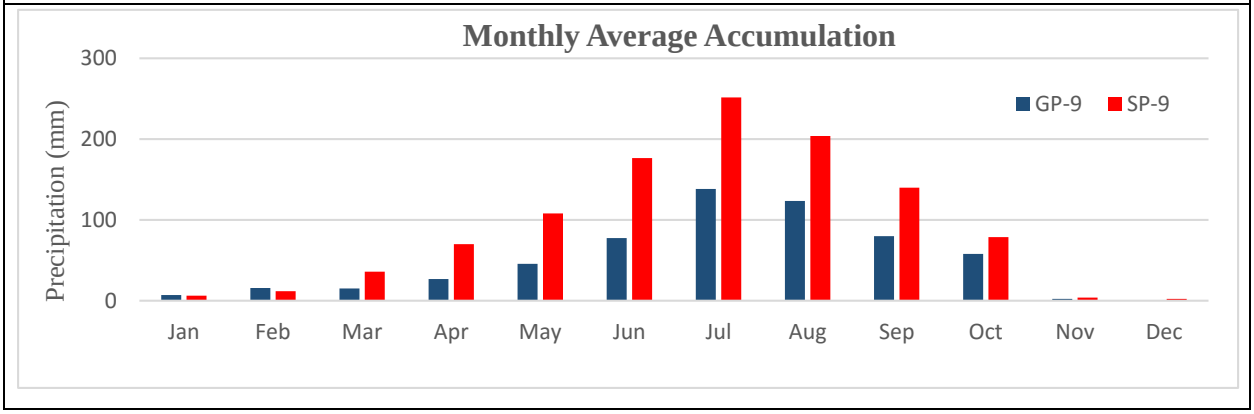

Annual Accumulation
Ann Precipitation (mm)
4000
3000
2000
1000
0
GP8
SP8
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012

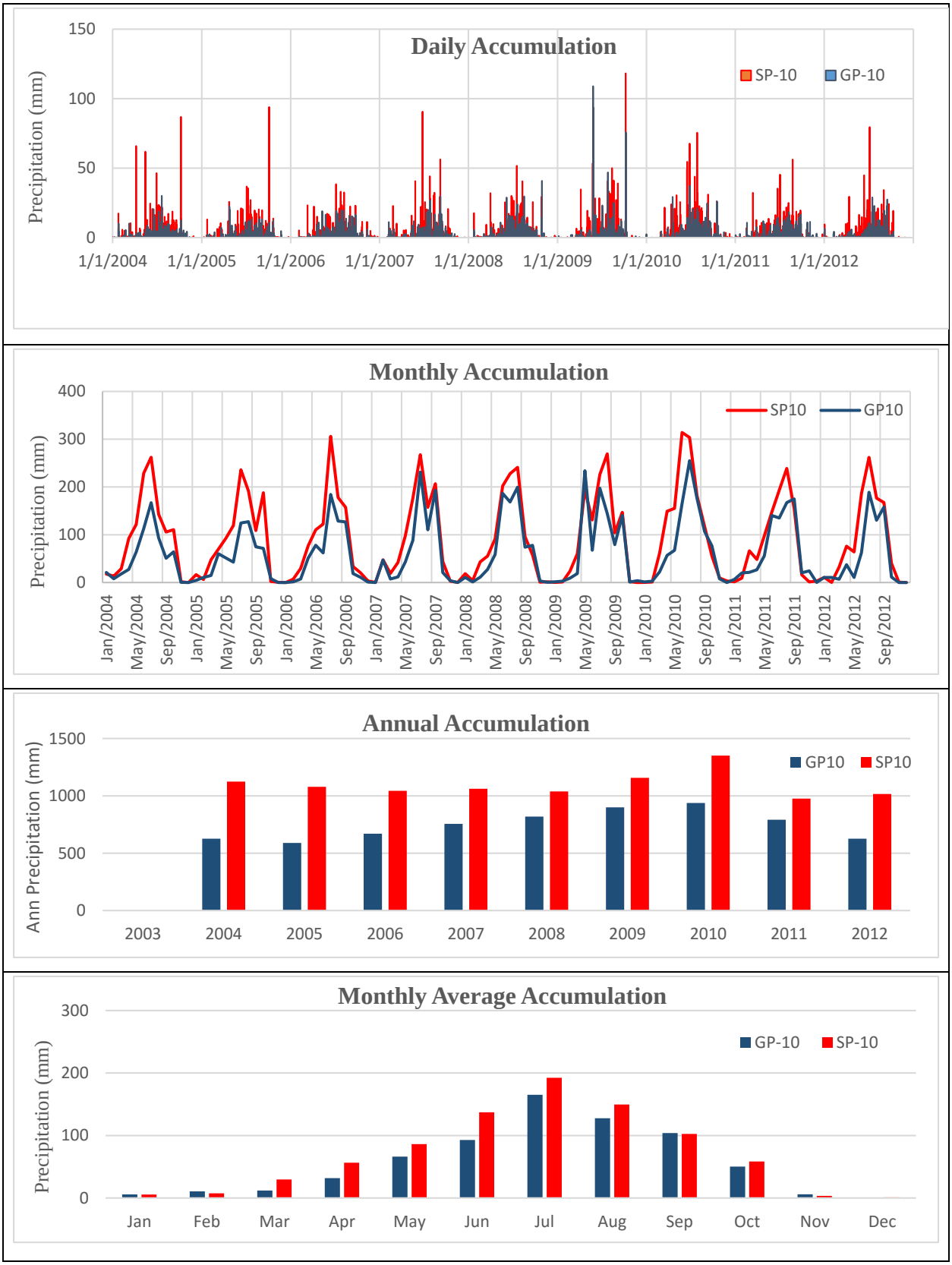

Monthly Average Accumulation
Precipitation (mm)
1000
500
0
GP-8
SP-8
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

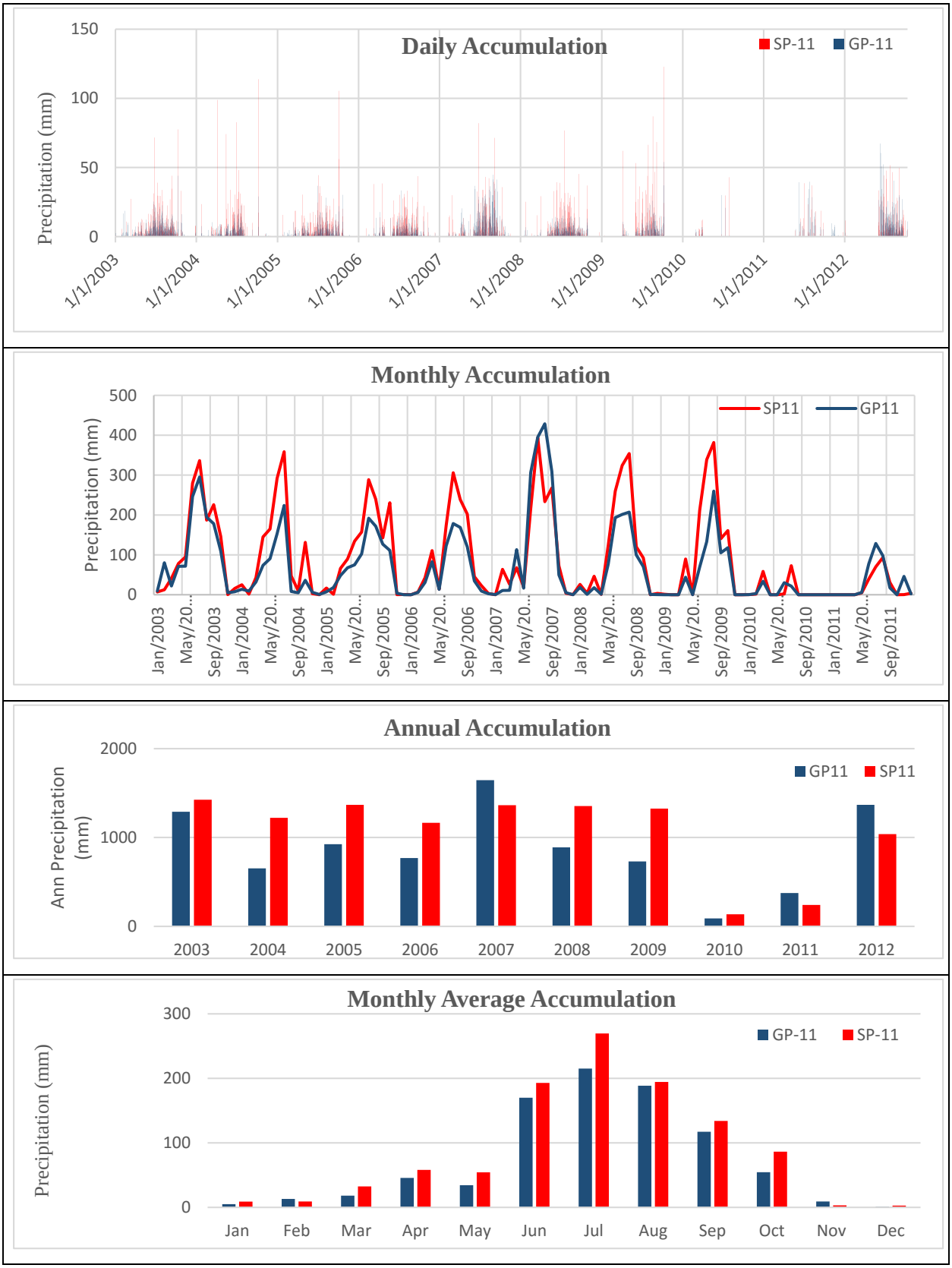

Daily Accumulation
Precipitation (mm)
150
100
50
0
SP-9
GP-9
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012

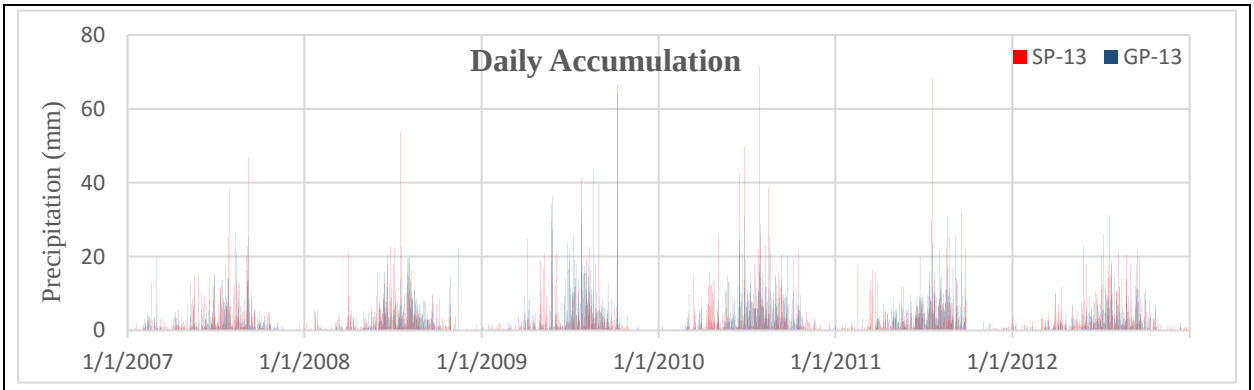

Monthly Accumulation
Precipitation (mm)
400
300
200
100
0
SP9
GP9
Jan/2003
May/20...
Sep/2003
Jan/2004
May/20...
Sep/2004
Jan/2005
May/20...
Sep/2005
Jan/2006
May/20...
Sep/2006
Jan/2007
May/20...
Sep/2007
Jan/2008
May/20...
Sep/2008
Jan/2009
May/20...
Sep/2009
Jan/2010
May/20...
Sep/2010
Jan/2011
May/20...
Sep/2011
Jan/2012
May/20...
Sep/2012

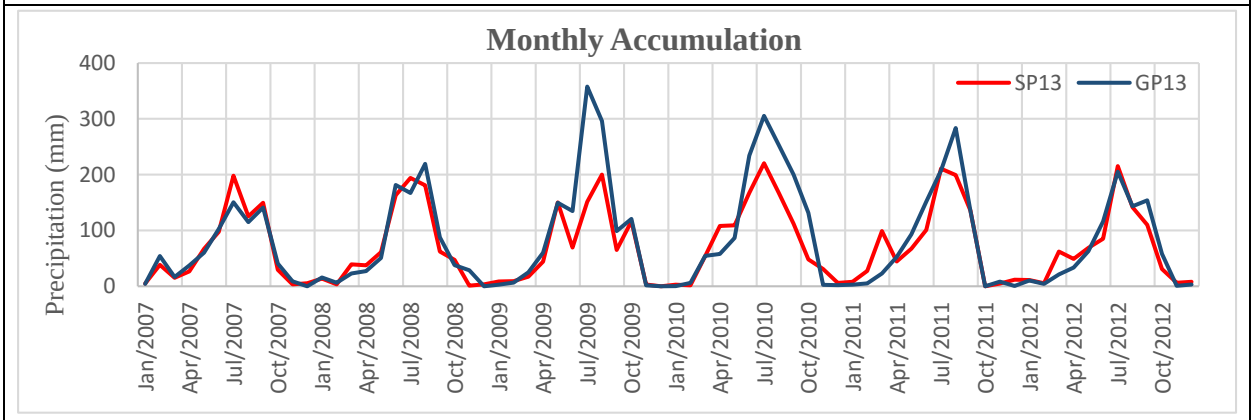

Annual Accumulation
Ann Precipitation (mm)
1500
1000
500
0
GP9
SP9
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012

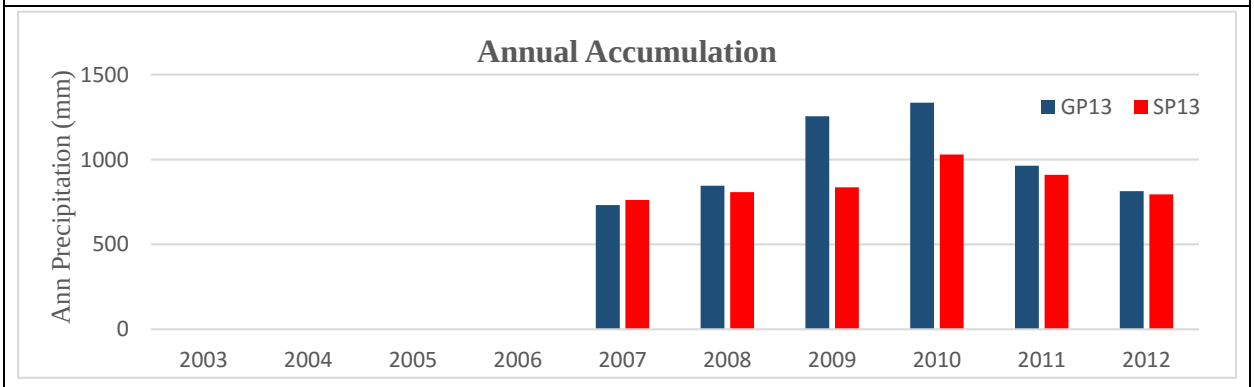

Monthly Average Accumulation
Precipitation (mm)
300
200
100
0
GP-9
SP-9
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

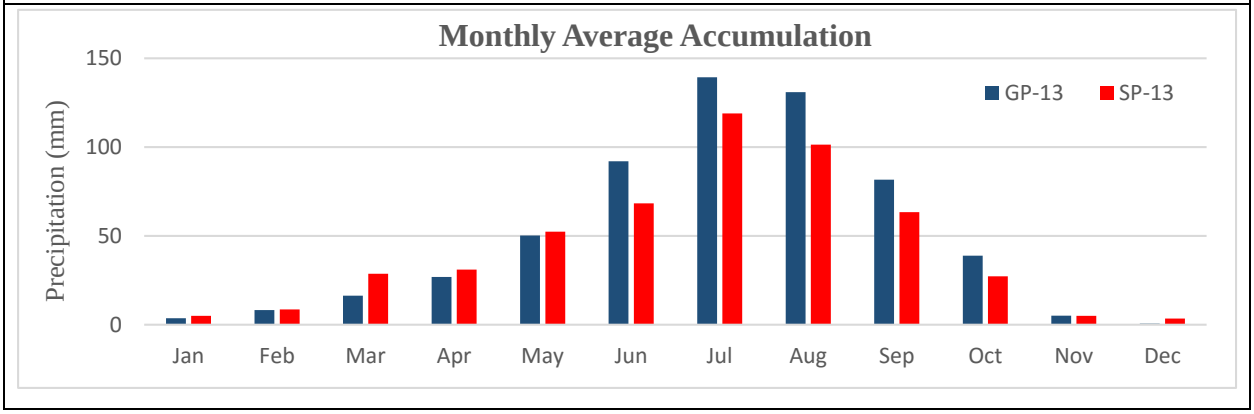

Daily Accumulation
Precipitation (mm)
150
100
50
0
SP-10
GP-10
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
Precipitation (mm)
400
300
200
100
0
SP10
GP10
Jan/2004
May/2004
Sep/2004
Jan/2005
May/2005
Sep/2005
Jan/2006
May/2006
Sep/2006
Jan/2007
May/2007
Sep/2007
Jan/2008
May/2008
Sep/2008
Jan/2009
May/2009
Sep/2009
Jan/2010
May/2010
Sep/2010
Jan/2011
May/2011
Sep/2011
Jan/2012
May/2012
Sep/2012
Annual Accumulation
Ann Precipitation (mm)
1500
1000
500
0
GP10
SP10
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
Precipitation (mm)
300
200
100
0
GP-10
SP-10
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

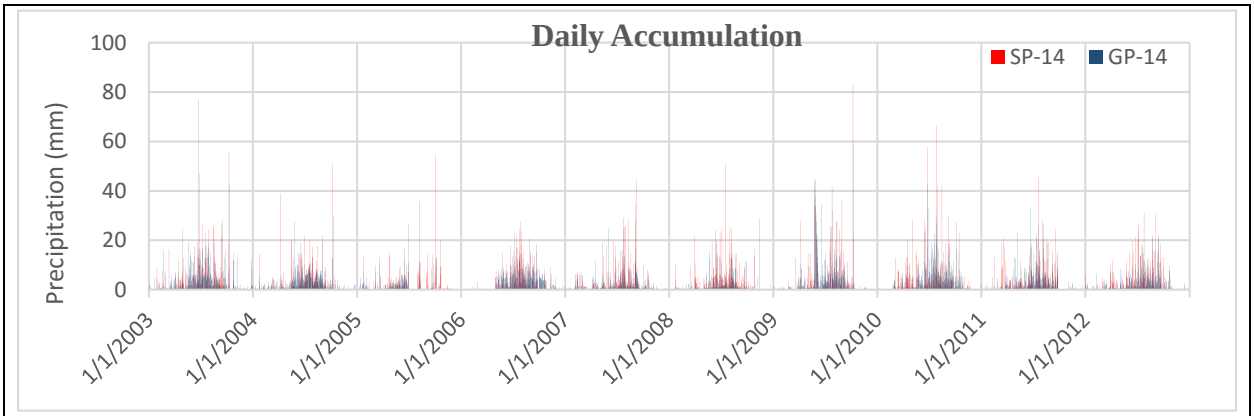

Daily Accumulation
SP-11
GP-11
Precipitation (mm)
150
100
50
0
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012
Monthly Accumulation
SP11
GP11
Precipitation (mm)
500
400
300
200
100
0
Jan/2003
May/20...
Sep/2003
Jan/2004
May/20...
Sep/2004
Jan/2005
May/20...
Sep/2005
Jan/2006
May/20...
Sep/2006
Jan/2007
May/20...
Sep/2007
Jan/2008
May/20...
Sep/2008
Jan/2009
May/20...
Sep/2009
Jan/2010
May/20...
Sep/2010
Jan/2011
May/20...
Sep/2011
Annual Accumulation
GP11
SP11
Ann Precipitation (mm)
2000
1000
0
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
Monthly Average Accumulation
GP-11
SP-11
Precipitation (mm)
300
200
100
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

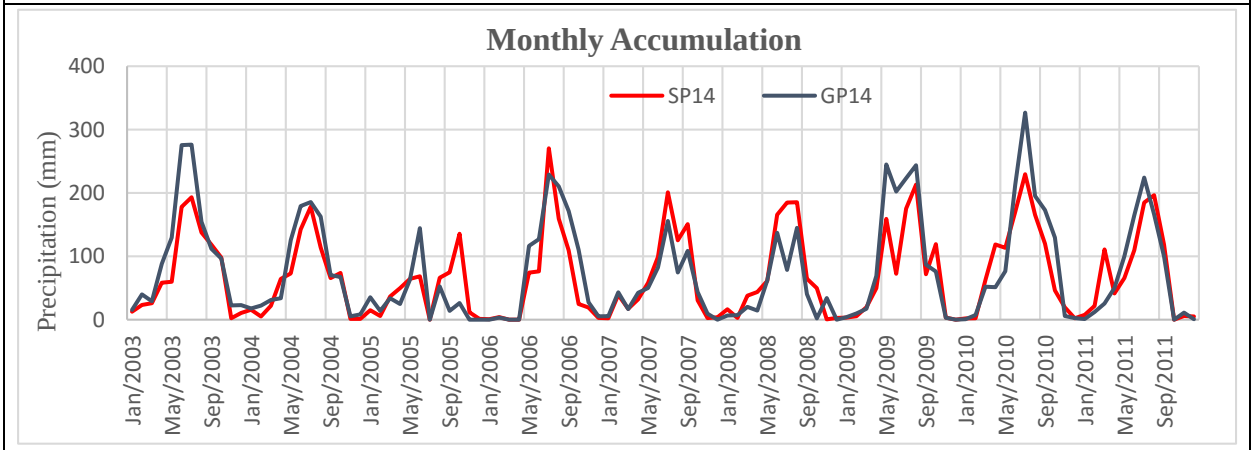

Daily Accumulation
SP-13
GP-13
Precipitation (mm)
80
60
40
20
0
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012

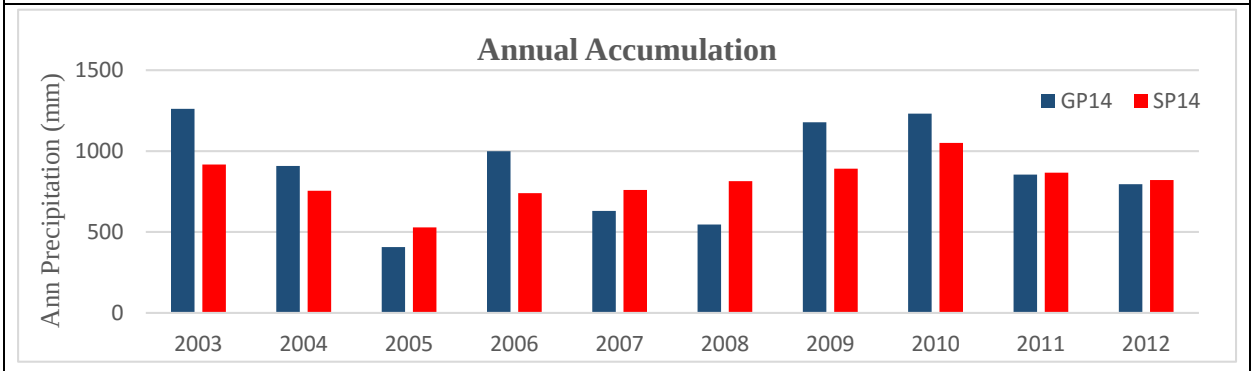

Monthly Accumulation
SP13
GP13
Precipitation (mm)
400
300
200
100
0
Jan/2007
Apr/2007
Jul/2007
Oct/2007
Jan/2008
Apr/2008
Jul/2008
Oct/2008
Jan/2009
Apr/2009
Jul/2009
Oct/2009
Jan/2010
Apr/2010
Jul/2010
Oct/2010
Jan/2011
Apr/2011
Jul/2011
Oct/2011
Jan/2012
Apr/2012
Jul/2012
Oct/2012

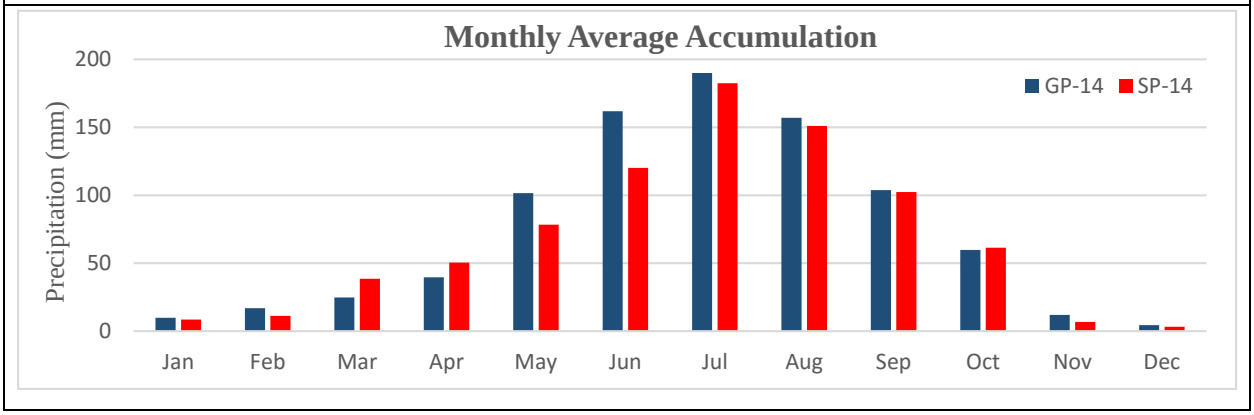

Annual Accumulation
GP13
SP13
Ann Precipitation (mm)
1500
1000
500
0
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012

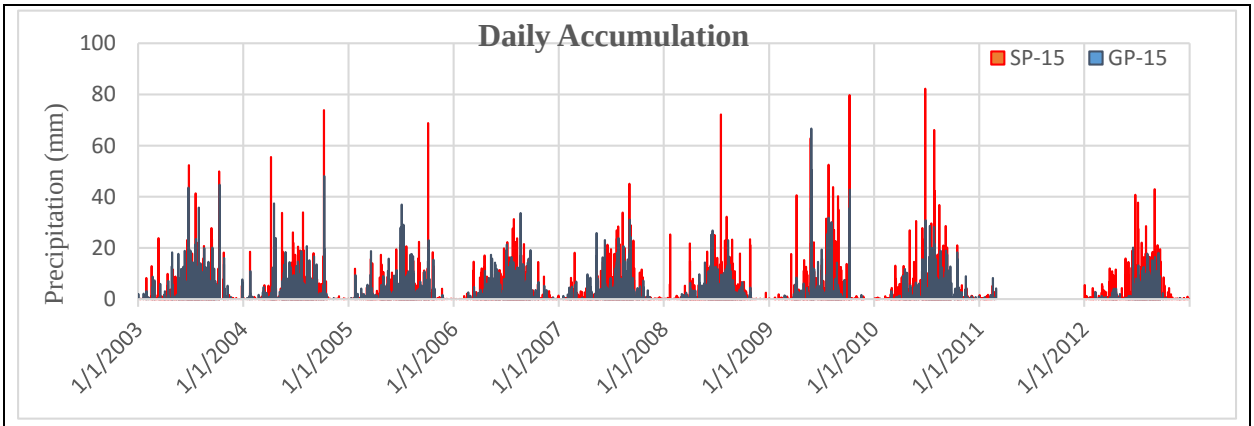

Monthly Average Accumulation
GP-13
SP-13
Precipitation (mm)
150
100
50
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

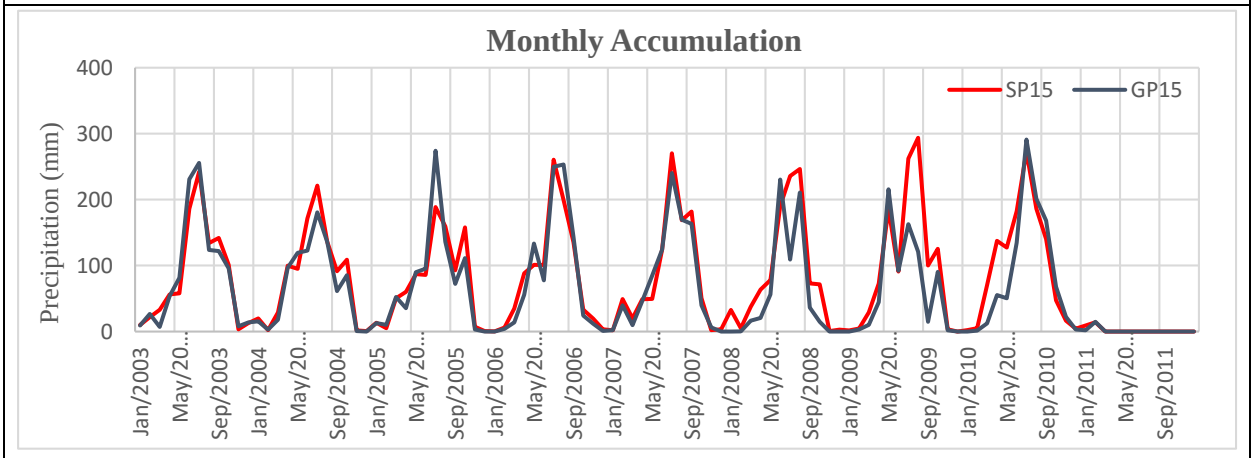

Daily Accumulation
Precipitation (mm)
100
80
60
40
20
0
SP-14
GP-14
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012

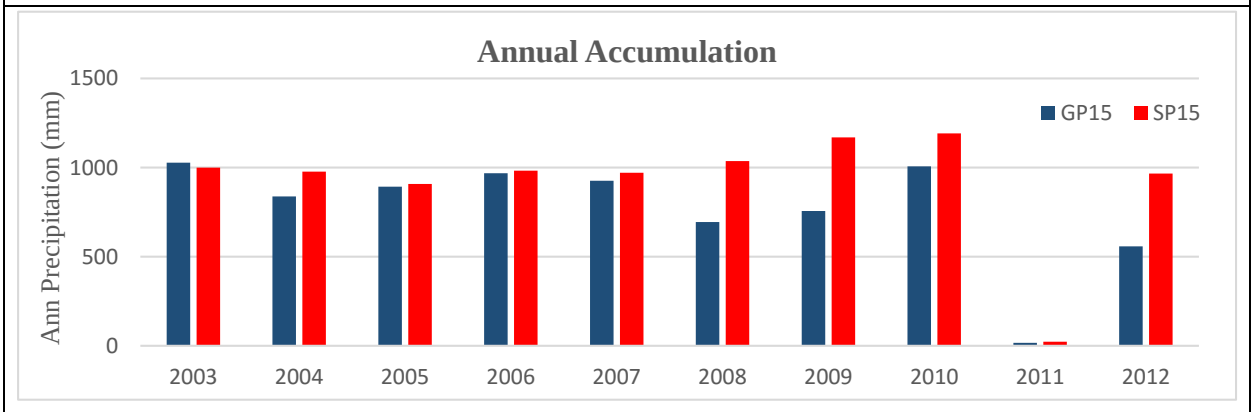

Monthly Accumulation
Precipitation (mm)
400
300
200
100
0
SP14
GP14
Jan/2003
May/2003
Sep/2003
Jan/2004
May/2004
Sep/2004
Jan/2005
May/2005
Sep/2005
Jan/2006
May/2006
Sep/2006
Jan/2007
May/2007
Sep/2007
Jan/2008
May/2008
Sep/2008
Jan/2009
May/2009
Sep/2009
Jan/2010
May/2010
Sep/2010
Jan/2011
May/2011
Sep/2011

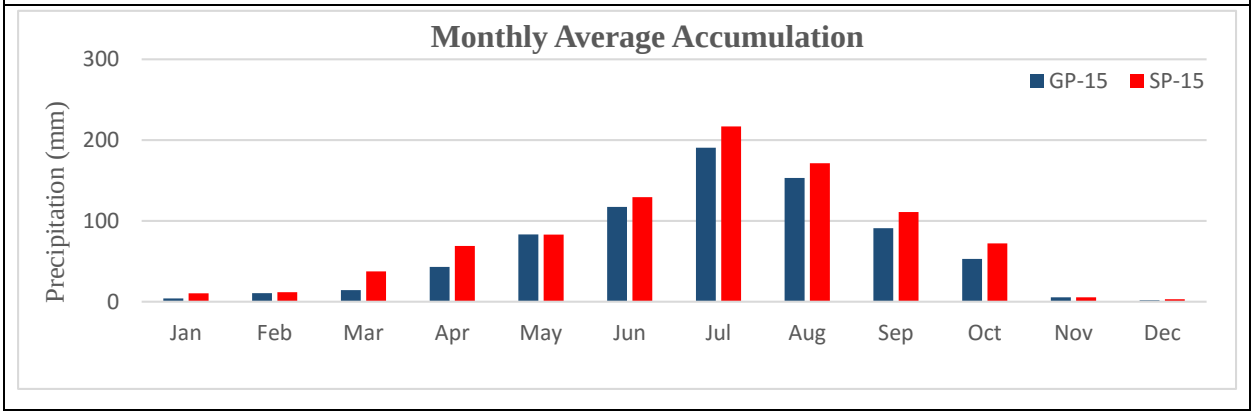

Annual Accumulation
Ann Precipitation (mm)
1500
1000
500
0
GP14
SP14
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012


Monthly Average Accumulation
Precipitation (mm)
200
150
100
50
0
GP-14
SP-14
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec


Daily Accumulation
Precipitation (mm)
100
80
60
40
20
0
SP-15
GP-15
1/1/2003
1/1/2004
1/1/2005
1/1/2006
1/1/2007
1/1/2008
1/1/2009
1/1/2010
1/1/2011
1/1/2012


Monthly Accumulation
Precipitation (mm)
400
300
200
100
0
SP15
GP15
Jan/2003
May/20...
Sep/2003
Jan/2004
May/20...
Sep/2004
Jan/2005
May/20...
Sep/2005
Jan/2006
May/20...
Sep/2006
Jan/2007
May/20...
Sep/2007
Jan/2008
May/20...
Sep/2008
Jan/2009
May/20...
Sep/2009
Jan/2010
May/20...
Sep/2010
Jan/2011
May/20...
Sep/2011


Annual Accumulation
Ann Precipitation (mm)
1500
1000
500
0
GP15
SP15
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012


Monthly Average Accumulation
Precipitation (mm)
300
200
100
0
GP-15
SP-15
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec

Appendix-3

Results of Statistical analysis

Table 21: Result of Statistical Analysis on Daily Time Scale

Pixel center-elevation	Statistical Parameters								
	NASH (R2)	NAD (%)	RMSD	MAD	MRAD	RR	EB (%)	CPOD_S	CPOD_G
S1-845	0.04	-11.47	18.37	1.07	4.0E-05	0.50	-11.47	0.82	0.75
S2-1505	-0.30	5.45	19.39	0.44	2.0E-05	0.42	5.45	0.82	0.74
S3-1305	-0.03	-4.07	22.42	0.38	1.4E-05	0.42	-4.07	0.78	0.68
S4-87	0.32	-31.02	28.28	5.40	1.1E-04	0.58	-31.54	0.88	0.82
S5-2600	-0.54	12.26	11.32	0.50	4.6E-05	0.35	12.26	0.79	0.61
S6-2700	-0.94	24.72	11.30	0.94	8.6E-05	0.41	24.72	0.60	0.78
S7-1650	-0.56	49.80	12.01	1.92	1.7E-04	0.52	49.80	0.74	0.64
S8-2060	0.07	-31.08	18.31	2.71	1.1E-04	0.41	-31.08	0.79	0.73
S9-3110	-2.43	84.12	8.01	1.75	3.0E-04	0.47	84.12	0.76	0.70
S10-3330	-1.20	46.60	7.83	1.10	1.6E-04	0.41	46.60	0.73	0.73
S11-3020	-1.05	21.82	9.12	0.74	8.5E-05	0.45	21.82	0.83	0.86
S12-2140	-2.45	28.06	8.50	0.92	9.9E-05	0.43	28.06	0.77	0.85
S13-4161	-0.33	-13.54	5.65	0.38	6.4E-05	0.41	-13.54	0.89	0.70
S14-1887	-0.35	-7.63	6.02	0.24	2.7E-05	0.40	-7.58	0.77	0.69
S15-4287	-0.78	20.04	7.15	0.54	7.0E-05	0.34	20.04	0.85	0.63

Table 22: Result of Statistical Analysis on Monthly Time Scale

Pixel center-elevation	Statistical Parameters								
	NASH (R2)	NAD (%)	RMSD	MAD	MRAD	RR	EB (%)	CPOD_S	CPOD_G
S1-845	0.91	-11.47	105.34	25.41	1E-03	0.93	-11.47	1.00	0.97
S2-1505	0.88	5.45	107.96	9.81	5E-04	0.89	5.45	1.00	0.98
S3-1305	0.86	-4.07	145.80	9.02	3E-04	0.89	-4.07	0.99	0.95
S4-87	0.89	-31.54	280.05	130.28	0.003	0.93	-31.54	1.00	0.95
S5-2600	0.75	12.26	56.20	11.02	0.001	0.89	12.26	1.00	0.95
S6-2700	0.73	24.72	56.45	22.24	0.002	0.91	24.72	0.98	0.95
S7-1650	0.76	49.80	72.19	45.54	0.004	0.93	49.80	0.98	0.93
S8-2060	0.77	-31.08	162.13	64.38	0.003	0.89	-31.08	1.00	0.94
S9-3110	0.29	84.12	60.95	41.50	0.007	0.93	84.12	0.99	0.90
S10-3330	0.74	46.60	46.96	29.00	0.004	0.92	46.60	0.99	0.95
S11-3020	0.80	21.82	56.23	15.88	0.002	0.88	21.82	1.00	0.97
S12-2140	0.80	28.06	49.31	21.75	0.002	0.93	28.06	1.00	0.99
S13-4161	0.88	-13.54	41.70	11.18	0.002	0.90	-13.54	1.00	0.97
S14-1887	0.87	-7.58	38.38	5.56	6E-04	0.87	-7.58	1.00	0.96
S15-4287	0.87	20.04	35.79	12.84	0.002	0.91	20.04	1.00	0.91

Table 23: Result of Statistical Analysis on Annual Time Scale

Pixel center-elevation	Statistical Parameters								
	NASH (R2)	NAD (%)	RMSD	MAD	MRAD	RR	EB (%)	CPOD_S	CPOD_G
S1-845	0.98	-10.11	341.71	268.64	0.01	1.00	-11.47	1.00	1.00
S2-1505	0.98	2.67	285.27	57.73	0.00	0.95	5.45	1.00	1.00
S3-1305	0.93	-2.13	692.24	56.71	0.00	0.80	-4.07	1.00	1.00
S4-87	0.90	-27.61	1540.52	1368.47	0.03	0.90	-31.54	1.00	1.00
S5-2600	0.98	8.19	171.78	88.25	0.01	0.96	12.26	1.00	1.00
S6-2700	0.88	23.17	341.88	250.18	0.02	0.82	24.72	1.00	1.00
S7-1650	0.73	45.61	536.28	500.52	0.05	0.90	49.80	1.00	1.00
S8-2060	0.88	-27.41	827.08	681.33	0.03	0.36	-31.08	1.00	1.00
S9-3110	0.30	73.52	469.83	435.23	0.07	0.51	84.12	1.00	1.00
S10-3330	0.77	46.60	345.28	313.16	0.05	0.93	46.60	1.00	1.00
S11-3020	0.82	20.28	381.53	177.03	0.02	0.72	21.82	1.00	1.00
S12-2140	0.88	26.60	312.85	247.40	0.03	0.91	28.06	1.00	1.00
S13-4161	0.96	-13.54	165.58	80.51	0.01	0.97	-13.54	1.00	1.00
S14-1887	0.96	-3.67	176.07	32.36	0.00	0.78	-7.58	1.00	1.00
S15-4287	0.91	20.41	225.99	156.85	0.02	0.85	20.04	1.00	1.00

Appendix – 4

SHyFT

A. Requirements

B. Installation procedure

C. Configuration files

1. kerabari_region.yaml
2. kerabari_datasets.yaml
3. kerabari_model.yaml
4. kerabari_simulation.yaml
5. kerabari_calibration.yaml

D. Program files

1. txt_2_netcdf.py
2. RunShyft.py
3. CalibShyft.py
4. Get_SimObs.py

A. Requirements

For compiling and running SHyFT in Windows, following programs are required to be installed:

a. A C++ Compiler
b. The BLAS and LAPACK libraries (development Packages)
c. GIT bash and GIT CMD
d. A Python3 (3.4 or higher) interpreter
e. The SWIG wrapping tool (>= 3.0.5)
f. The NumPy package (>= 1.8.0)
g. The netCDF4 package (>= 1.2.1)
h. nose (1.3.7)
i. pyproj (1.9.5.1)
j. matplotlib (1.5.1)
k. shapley
l. gdal and
m. The CMake building tool (2.8.7 or higher)

B. Installation Procedure

The step-wise procedure to install SHyFt is given below:

a. Shyft is at: https://github.com/statkraft/shyft
b. First install GIT from https://git-scm.com
c. Start GIT, change the path to the location where you want SHyFT (inside the GIT window, use “cd" to change directory).
d. To get the shyft programs, type "git clone https://github.com/statkraft/shyft.git” in the GIT window. The files will be copied to your disk.
e. To get the shyft data: type “git clone https://github.com/statkraft/shyft-data.git” in the GIT window and keep the same directory as for shyft
f. Go to shyft-data/distro and unpack the windows archive found there. Copy all the files from the resulting directory and put them in /shyft/shyft/api.
g. Then install Anaconda 3.4, SWIG and PyCharm

Below are some additional steps to be carried out in order to get rid from common errors with gdal and python. Go to the GIT command prompt and:

a. Update conda

   Type “conda update conda”

b. Downgrade python to version 3.4

   Type “conda install python=3.4”

c. Install some packages

   Type “conda install netcdf4=1.2.1 nose=1.3.7 pyproj=1.9.4 matplotlib=1.5.1”. If pyproj 1.9.4 does not work, then install pyproj 1.9.5.1 which is easily available in the internet

d. Install shapely 1.5.13

   Type “conda install --channel https://conda.anaconda.org/IOOS shapely”

Finally, open Pycharm and run the test

C. Configuration files

1	kerabari_region.yaml

```
---
repository:
  class:
!!python/name:shyft.repository.netcdf.cf_region_model_repository.CFRegionModelReposito
ry
  data_file: netcdf/orchestration-testdata/cell_data.nc
domain:
  EPSG: 32646
  nx: 103
  ny: 165
  step_x: 1000
  step_y: 1000
  lower_left_x: 130000
  lower_left_y: 2955000

catchment_indices:
    - 0
    - 1
    - 2
    - 3
    - 4
    - 5
```

2	kerabari_datasets.yaml

```
---
sources:
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
    params:
      stations_met: netcdf/orchestration-testdata/precipitation.nc
      selection_criteria: null
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
    params:
      stations_met: netcdf/orchestration-testdata/temperature.nc
      selection_criteria: null
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
    params:
      stations_met: netcdf/orchestration-testdata/wind_speed.nc
      selection_criteria: null
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
    params:
      stations_met: netcdf/orchestration-testdata/relative_humidity.nc
      selection_criteria: null
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
    params:
      stations_met: netcdf/orchestration-testdata/radiation.nc
      selection_criteria: null
...
```

3	kerabari_model.yaml

```yaml
model_t: pt_gs_k.PTGSKModel  # model to construct
parameters:
  interpolation:
    btk:
      gradient: -0.6
      gradient_sd: 0.25
      nugget: 0.5
      range: 200000.0
      sill: 25.0
      zscale: 20.0
    idw:
      max_distance: 600000.0
      max_members: 10
      precipitation_gradient: 2.0
  model:
    actual_evapotranspiration:
      scale_factor: 1.5
    data:
      constant_relative_humidity: 0.8
      constant_wind_speed: 2.0
    gamma_snow:
      calculate_iso_pot_energy: false
      fast_albedo_decay_rate: 8.438395
      glacier_albedo: 0.4
      initial_bare_ground_fraction: 0.04
      max_albedo: 0.9
      max_water: 0.1
      min_albedo: 0.6
      slow_albedo_decay_rate: 20.665371
      snow_cv: 0.4
      tx: -0.279719
      snowfall_reset_depth: 5.0
      surface_magnitude: 30.0
      wind_const: 1.0
      wind_scale: 5.438631
      winter_end_day_of_year: 100
    hbv_snow:
        foo: 10.0
    kirchner:
      c1: -8.378072
      c2: -0.894523
      c3: -0.125215
    p_corr_scale_factor: 1.116462
    priestley_taylor:
      albedo: 0.2
      alpha: 1.26
```

4	kerabari_simulation.yaml

```yaml
---
kerabari:
  region_config_file: kerabari_region.yaml
  model_config_file: kerabari_model.yaml
  datasets_config_file: kerabari_datasets.yaml
  start_datetime: 2003-01-01T00:00:00
  run_time_step: 86400  # 1 hour time step
  number_of_steps: 3653  # 1 year
  #region_id: 1          # this is optional (default 0)
  #interpolation_id: 2   # this is optional (default 0)
  epsg: 32646
...
```

5	kerabari_calibration.yaml

```yaml
kerabari:
  model_config_file: kerabari_simulation.yaml
  calibrated_model_file: calibrated_model.yaml  # file where the calibrated parameters
will go
  optimization_method:
    # name: min_bobyqa # can be 'min_bobyqa', 'dream' or 'sceua'
    # params:
      # max_n_evaluations: 1500
      # tr_start: 0.1
      # tr_stop: 1.0e-5
    # name: dream
    # params:
      # max_n_evaluations: 1500
    name: sceua
    params:
      max_n_evaluations: 10
      x_eps: 0.0001
      y_eps: 1.0e-5
  target:
  - repository: !!python/name:shyft.repository.netcdf.cf_ts_repository.CFTsRepository
    params:
      file: netcdf/orchestration-testdata/discharge.nc
      var_type: discharge
    1D_timeseries:
    - catch_id: [0]
      uid: yebesa
      start_datetime: 2003-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 1826 # 26280
      weight: 1.0
      obj_func:
        name: NSE # Nash-Sutcliffe efficiency (NSE) or Kling-Gupta efficiency (KGE)
        scaling_factors:
          s_corr: 1.0
          s_var: 1.0
          s_bias: 1.0
     - catch_id: [1]
      uid: Dokana
      start_datetime: 2003-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 1826 # 26280
      weight: 1.0
      obj_func:
        name: NSE # Nash-Sutcliffe efficiency (NSE) or Kling-Gupta efficiency (KGE)
        scaling_factors:
          s_corr: 1.0
          s_var: 1.0
          s_bias: 1.0
    - catch_id: [0,1,2]
      uid: wangdue
      start_datetime: 2003-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 1826 # 26280
      weight: 1.0
      obj_func:
        name: NSE # Nash-Sutcliffe efficiency (NSE) or Kling-Gupta efficiency (KGE)
        scaling_factors:
          s_corr: 1.0
          s_var: 1.0
```

```
        s_bias: 1.0
    - catch_id: [0,1,2,3]
      uid: sunkosh
      start_datetime: 2003-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 1826 # 26280
      weight: 1.0
      obj_func:
        name: NSE # Nash-Sutcliffe efficiency (NSE) or Kling-Gupta efficiency (KGE)
        scaling_factors:
          s_corr: 1.0
          s_var: 1.0
          s_bias: 1.0
    - catch_id: [0,1,2,3,4]
      uid: kerabari
      start_datetime: 2003-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 1826 # 26280
      weight: 1.0
      obj_func:
        name: NSE # Nash-Sutcliffe efficiency (NSE) or Kling-Gupta efficiency (KGE)
        scaling_factors:
          s_corr: 1.0
          s_var: 1.0
          s_bias: 1.0
  overrides:
    model:
      model_t: pt_gs_k.PTGSKOptModel
calibration_parameters:
    c1:
      min: -12.0 # -3.0
      max: 0.0 # 2.0
    c2:
      min: -1.0 # 0.8
      max: 1.2 # 1.2
    c3:
      min: -0.15
      max: -0.05
    ae_scale_factor:
      min: 1.5
      max: 1.5
    TX:
      min: -3.0
      max: 2.0
    wind_scale:
      min: 1.0
      max: 6.0
    max_water:
      min: 0.1
      max: 0.1
    wind_const:
      min: 1.0
      max: 1.0
    fast_albedo_decay_rate:
      min: 5.0 # 5.0
      max: 15.0 # 15.0
    slow_albedo_decay_rate:
      min: 20.0 # 20.0
      max: 40.0 # 40.0
    surface_magnitude:
      min: 30.0
      max: 30.0
    max_albedo:
```

```
  min: 0.9
  max: 0.9
min_albedo:
  min: 0.6
  max: 0.6
snowfall_reset_depth:
  min: 5.0
  max: 5.0
snow_cv:
  min: 0.4
  max: 0.4
snow_cv_forest_factor:
  min: 0.0
  max: 0.0
snow_cv_altitude_factor:
  min: 0.0
  max: 0.0
glacier_albedo:
  min: 0.4
  max: 0.4
p_corr_scale_factor:
  min: 0.75
  max: 3.0
```

D. Program files

1	txt_2_netcdf.py

```
# -*- coding: utf-8 -*-

import os
from dateutil.parser import parse
from datetime import datetime, timedelta, timezone
import numpy as np
import netCDF4

def write_nc(ncfilename,mode,dim_name,dim_size,Vars):
    req = ['name','dtype','dims','values']
    f = netCDF4.Dataset(ncfilename, mode)
    if(mode=='w'):
        for i in range(len(dim_name)):
            f.createDimension(dim_name[i], dim_size[i])
    for j in range(len(Vars)):
        print (Vars[j]['name'])
        #print (Vars[j].get('values','None'))
        if(Vars[j]['dtype']=='vlen_t'):
            vlen_t = f.createVLType(np.int32, 'vlen')
            var = f.createVariable(Vars[j]['name'],vlen_t,Vars[j]['dims'])
        else:
            var =
f.createVariable(Vars[j]['name'],Vars[j]['dtype'],Vars[j]['dims'],zlib=True,complevel=
4,fill_value=Vars[j].get('missing_val',None))
        #var.units = Var_units[j]
        var.setncatts({k:Vars[j][k] for k in Vars[j].keys() if k not in req })
        if(len(Vars[j]['dims'])>0):
            if(Vars[j]['dtype']=='vlen_t'):
                #print(Vars[j]['values'][0:-1])
                var[:] = Vars[j]['values'][0:-1]
            else:
                var[:] = Vars[j]['values']#[:]
    f.close()

def write_hydclim_to_nc(var_name, epsg, missing_val, var_unit, dim_size, t_v, name_v,
x_v, y_v, z_v, var_v, ids, f):
    dim_name = ('time','station')
    proj4_str = '+proj=utm +zone={} +ellps=WGS84 +datum=WGS84 +units=m
+no_defs'.format(epsg[-2:])
    Vars = [{'name':'time','dtype':'f8','dims':('time',),'units':'seconds since 1970-
01-01 00:00:00 +00:00','values':t_v},

#{'name':'station_name','dtype':'S2','dims':('station','strlen',),'cf_role':'timeserie
s_id','values':netCDF4.stringtochar(np.array([str(i)+str(i) for i in range(3)]))},

{'name':'series_name','dtype':str,'dims':('station',),'cf_role':'timeseries_id','value
s':name_v},

{'name':'x','dtype':'f8','dims':('station',),'units':'m','axis':'X','standard_name':'p
rojection_x_coordinate','values':x_v},

{'name':'y','dtype':'f8','dims':('station',),'units':'m','axis':'Y','standard_name':'p
rojection_y_coordinate','values':y_v},

{'name':'z','dtype':'f8','dims':('station',),'units':'m','axis':'Z','standard_name':'h
eight','long_name':"height above mean sea level",'values':z_v},
```

```
{'name':'crs','dtype':'i4','dims':(),'grid_mapping_name':'transverse_mercator','proj4'
:proj4_str,'epsg_code':"EPSG:" + epsg},

{'name':var_name,'dtype':'f8','dims':('time','station',),'units':var_unit,'coordinates
':'y x z','grid_mapping':'crs','missing_val':missing_val,'values':var_v}]
    if ids is not None:

Vars.append({'name':'catchment_id','dtype':'vlen_t','dims':('station',),'values':ids})
    write_nc(f,'w',dim_name,dim_size,Vars)

def write_cellinfo_to_nc(epsg, dim_size, x_v, y_v, z_v, area, ff, rf, lf, gf ,c_id,
f):
    dim_name = ('cell',)
    proj4_str = '+proj=utm +zone={} +ellps=WGS84 +datum=WGS84 +units=m
+no_defs'.format(epsg[-2:])
    Vars =
[{'name':'x','dtype':'f8','dims':('cell',),'units':'m','axis':'X','standard_name':'pro
jection_x_coordinate','values':x_v},

{'name':'y','dtype':'f8','dims':('cell',),'units':'m','axis':'Y','standard_name':'proj
ection_y_coordinate','values':y_v},

{'name':'z','dtype':'f8','dims':('cell',),'units':'m','axis':'Z','standard_name':'heig
ht','long_name':"height above mean sea level",'values':z_v},

{'name':'crs','dtype':'i4','dims':(),'grid_mapping_name':'transverse_mercator','proj4'
:proj4_str,'epsg_code':"EPSG:" + epsg},
            {'name':'area','dtype':'f8','dims':('cell',),'units':'m2','coordinates':'y
x z','grid_mapping':'crs','values':area},
            {'name':'forest-fraction','dtype':'f8','dims':('cell',),'units':'-
','coordinates':'y x z','grid_mapping':'crs','values':ff},
            {'name':'reservoir-fraction','dtype':'f8','dims':('cell',),'units':'-
','coordinates':'y x z','grid_mapping':'crs','values':rf},
            {'name':'lake-fraction','dtype':'f8','dims':('cell',),'units':'-
','coordinates':'y x z','grid_mapping':'crs','values':lf},
            {'name':'glacier-fraction','dtype':'f8','dims':('cell',),'units':'-
','coordinates':'y x z','grid_mapping':'crs','values':gf},
            {'name':'catchment_id','dtype':'i4','dims':('cell',),'units':'-
','coordinates':'y x z','grid_mapping':'crs','values':c_id}]
    write_nc(f,'w',dim_name,dim_size,Vars)

def read_cellinfo_from_txt(f_in,abbrv):
    with open(f_in) as f:
        epsg = f.readline().strip().split('\t')[1]
        col_headers = f.readline().strip().split('\t')
        data = np.loadtxt(f, delimiter='\t')
    return (epsg,(data.shape[0],),data[:,1],data[:,2],data[:,3],data[:,4],
            data[:,7],data[:,5],data[:,6],data[:,8],data[:,0])

def read_hydclim_from_txt(f_in,abbrv):
    var_name={'prec':'precipitation',
              'temp':'temperature',
              'wind':'wind_speed',
              'rh':'relative_humidity',
              'rad':'global_radiation',
              'q':'discharge'}
    var_unit={'prec':'mm hr-1',
              'temp':'degree_celsius',
              'wind':'m s-1',
              'rh':'0-1',
              'rad':'W m-2',
              'q':'m3 s-1'}
```

```
    # Fcns for unit conversion go here in the future
    ext_vct={'prec': lambda v: v,
             'temp': lambda v: v,
             'wind': lambda v: v,
             'rh': lambda v: v,
             'rad': lambda v: v,
             'q': lambda v: v}
    dt_2_foat = lambda dt_str:
parse(dt_str).replace(tzinfo=timezone(timedelta(seconds=0))).timestamp()
    num_header = 7
    catch_id = None
    if abbrv == 'q':
        num_header = 8
    with open(f_in) as f:
        header = [f.readline().strip().split('\t') for i in range(num_header)]
        epsg = header[0][1]
        missing_val = float(header[1][1])
        unit = header[2][1] # TODO: check if unit in txt file matches with
var_unit['abbrv']
        series_names = np.array(header[3][1:], dtype='O')
        x = np.array([float(str) for str in header[4][1:]], dtype = float)
        y = np.array([float(str) for str in header[5][1:]], dtype = float)
        z = np.array([float(str) for str in header[6][1:]], dtype = float)
        if num_header == 8:
            catch_id = np.array([np.array([int(s) for s in
str.split(';')],dtype='int32') for str in header[7][1:]+['0;1']], dtype=object)
        data = np.loadtxt(f, delimiter='\t', converters={0:dt_2_foat})
    t = data[:,0]
    v = data[:,1:]
    dim = v.shape
    nc_var_name = var_name[abbrv]
    return nc_var_name, epsg, missing_val, unit, dim, t, series_names, x, y, z, v,
catch_id

if __name__=='__main__':

    # Change the folder path with your path
    txt_folder = 'D:/Kbariconvert/text' # path to folder with input text files
    netcdf_folder = 'D:/Kbariconvert/netcdf' # path to folder with output netcdf files

    var_to_process = ['precipitation',
                      'temperature',
                      'wind_speed',
                      'relative_humidity',
                      'radiation',
                      'discharge',
                      'cell_data']

    var_dict =
{'precipitation':{'abbrv':'prec','fil_prefix':'precipitation','read_fcn':read_hydclim_
from_txt,'write_fcn':write_hydclim_to_nc},

'temperature':{'abbrv':'temp','fil_prefix':'temperature','read_fcn':read_hydclim_from_
txt,'write_fcn':write_hydclim_to_nc},

'wind_speed':{'abbrv':'wind','fil_prefix':'wind_speed','read_fcn':read_hydclim_from_tx
t,'write_fcn':write_hydclim_to_nc},

'relative_humidity':{'abbrv':'rh','fil_prefix':'relative_humidity','read_fcn':read_hyd
clim_from_txt,'write_fcn':write_hydclim_to_nc},

'radiation':{'abbrv':'rad','fil_prefix':'global_radiation','read_fcn':read_hydclim_fro
m_txt,'write_fcn':write_hydclim_to_nc},
```

```
'discharge':{'abbrv':'q','fil_prefix':'discharge','read_fcn':read_hydclim_from_txt,'wr
ite_fcn':write_hydclim_to_nc},

'cell_data':{'abbrv':'cell_info','fil_prefix':'cell_data','read_fcn':read_cellinfo_fro
m_txt,'write_fcn':write_cellinfo_to_nc}
                }

    for var in var_to_process:
        f_in = os.path.join(txt_folder,var_dict[var]['fil_prefix']+'.txt')
        f_out = os.path.join(netcdf_folder,var_dict[var]['fil_prefix']+'.nc')

        res = var_dict[var]['read_fcn'](f_in,var_dict[var]['abbrv'])
        var_dict[var]['write_fcn'](*res+(f_out,))
```

2	RunShyft.py

```
from os import path
import unittest
from shyft import shyftdata_dir
from shyft.repository.default_state_repository import DefaultStateRepository
from shyft.orchestration.configuration import yaml_configs
from shyft.orchestration.simulators.config_simulator import ConfigSimulator

config_dir = 'C:\\shyft\\shyft\\tests\\netcdf\\kerabari_simulation.yaml'
#config_dir = path.join(path.dirname(__file__), "netcdf")
#cfg = yaml_configs.YAMLSimConfig("kerabari_simulation.yaml",
"sagelva",config_dir=config_dir, data_dir=shyftdata_dir)
cfg = yaml_configs.YAMLSimConfig(config_dir, "kerabari")

# get a simulator
simulator = ConfigSimulator(cfg)
#n_cells = simulator.region_model.size()
#state_repos = DefaultStateRepository(cfg.model_t, n_cells)
n_cells = simulator.region_model.size()
state_repos = DefaultStateRepository(simulator.region_model.__class__, n_cells)
simulator.run(cfg.time_axis, state_repos.get_state(0))

from matplotlib import pylab as plt
from shyft import api
from datetime import datetime

m=simulator.region_model
ts = m.statistics.discharge([0])
cal=api.Calendar()
sim=[value for value in ts.v]
sim_time=[ts.time(i) for i in range(len(sim))]
sim_date=[datetime.utcfromtimestamp(t) for t in sim_time]

print("print")
plt.figure(figsize=(14,7))
plt.plot(sim_date,sim)
plt.ylabel("Simulated discharge [m$^3$/s]", fontsize=12)
plt.show()
```

3	CalibShyft.py

```
import sys
sys.path.insert(0,'C:\\shyft')
from shyft.repository.default_state_repository import DefaultStateRepository
#from shyft.orchestration.configuration.yaml_configs import YAMLCalibConfig
from shyft.orchestration.configuration import yaml_configs
#from shyft.orchestration.simulators.config_simulator import ConfigCalibrator
from shyft.orchestration.simulators import config_simulator

cfg =
yaml_configs.YAMLCalibConfig("C:\\shyft\\shyft\\tests\\netcdf\\kerabari_calibration.ya
ml", "kerabari")#cfg = yaml_configs.……"Kerabari")all in single line
calib = config_simulator.ConfigCalibrator(cfg)
n_cells = calib.region_model.size()
state_repos = DefaultStateRepository(calib.region_model.__class__, n_cells)

calib.init()
res = calib.calibrate(cfg.sim_config.time_axis, state_repos.get_state(0),
cfg.optimization_method['name'],cfg.optimization_method['params'], p_vec=None)
#res = calib. ……=None)all in single line
print ('Done Calibrating')
optim_params_vct=[res.get(i) for i in range(res.size())]
R2=1-calib.optimizer.calculate_goal_function(optim_params_vct)
for i in range(0, res.size()):
    print('Param %s = %f' % (res.get_name(i), res.get(i)))
```

4	Get_SimObs.py

```
# load modules we need for post-processing
from netCDF4 import Dataset
import os
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
from datetime import datetime

# load modules from shyft
from shyft import shyftdata_dir
from shyft.repository.default_state_repository import DefaultStateRepository
from shyft.orchestration.configuration import yaml_configs
from shyft.orchestration.simulators.config_simulator import ConfigSimulator

# Part 1 - Run the simulation after calibration
# set up configuration
config_dir = 'C:\\shyft\\shyft\\tests\\netcdf\\kerabari_simulation.yaml'
config_section = "kerabari"
cfg = yaml_configs.YAMLSimConfig(config_dir,config_section)

# set up the simulator
simulator = ConfigSimulator(cfg)

# create a time axis # build the model # RUN THE MODEL
n_cells = simulator.region_model.size()
state_repos = DefaultStateRepository(simulator.region_model.__class__, n_cells)
simulator.run(cfg.time_axis, state_repos.get_state(0))

m=simulator.region_model
ts = m.statistics.discharge([0])
sim=[value for value in ts.v]
sim_time=[ts.time(i) for i in range(len(sim))]
sim_date=[datetime.utcfromtimestamp(t) for t in sim_time]

#load the discharge observations from netcdf file
data_path = os.path.abspath(os.path.realpath('C:\\shyft-data\\netcdf\\orchestration-
testdata'))
fn_met = "discharge.nc" #your precipitation file name
file_path = os.path.join(data_path,fn_met)

with Dataset(file_path) as dset:
    obs = dset.variables['discharge'][0:3653,4] # 4 is the catchment number
    obs_date = [datetime.utcfromtimestamp(t) for t in dset.variables['time'][0:3653]]


# Part 2 - Copy values to Excel
#Simulated
sim_ts = pd.Series(sim, index=sim_date)
df_sim = pd.DataFrame({'sim':sim_ts})
writer = pd.ExcelWriter('Sim_discharge.xlsx', engine='xlsxwriter')
df_sim.to_excel(writer, sheet_name='Simulated')
```

```
writer.save()

#Observed
obs_ts = pd.Series(obs, index=obs_date)
df_obs = pd.DataFrame({'obs':obs_ts})
writer = pd.ExcelWriter('Obs_discharge.xlsx', engine='xlsxwriter')
df_obs.to_excel(writer, sheet_name='Observed')
writer.save()
print('ok there')

#Part 3 - Get Areal Precipitation & Temperature
#Precipitation
precip = simulator.region_model.statistics.precipitation([0])
arealprecip =[value for value in precip.v]
precip_ts = pd.Series(arealprecip, index=sim_date)
df_precip = pd.DataFrame({'arealprecip':precip_ts})
writer = pd.ExcelWriter('Areal_Precipitation.xlsx', engine='xlsxwriter')
df_precip.to_excel(writer, sheet_name='Areal_Precip')
writer.save()
print('ok here')

#Temperature
temp = simulator.region_model.statistics.temperature([0])
arealtemp =[value for value in temp.v]
temp_ts = pd.Series(arealtemp, index=sim_date)
df_temp = pd.DataFrame({'arealtemp':temp_ts})
writer = pd.ExcelWriter('Areal_Temperature.xlsx', engine='xlsxwriter')
df_temp.to_excel(writer, sheet_name='Areal_Temp')
writer.save()
print('ok')
```