NTNU
Norwegian University of
Science and Technology

Regional modelling of the Narayani basin in Nepal.

Anugya Sapkota

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

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

Anugya Sapkota

Regional Modelling in Narayani Basin in Nepal

Master's Thesis in Hydropower Development

Trondheim, June 2016

Supervisor: Knut Alfredsen

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

NTNU
Norwegian University of Science and Technology

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: Anugya Sapkota

Title: Regional modelling of the Narayani basin in Nepal.

1 BACKGROUND

Computing inflow from ungauged basins is a challenge in practical hydrology since many planning assignments is located in areas where no data is available. A number of methods are developed to handle ungauged basins with variable success. One method that could be useful is regional calibration of hydrological models where the idea is that a parameter set for model simultaneously calibrated for several catchments will be valid also for other catchments in the same region. The purpose of this MSc thesis is to investigate this approach for the Narayani basin in Nepal using the Statkraft Hydrological Forecasting Toolbox (SHyFT) as a basis for the work.

2 MAIN QUESTIONS FOR THE THESIS

1. Prepare climate, runoff and catchment data for the Narayani basin for use in the SHyFT hydrological model. It should be paid particularly attention to runoff data for calibration and evaluation of the model to get a distribution of sub catchment that covers the catchment well. It is also important to document the quality of all data used as input for the model.
2. Set up and calibrate the model for the Narayani region. Evaluate the calibration and the performance of the model, paying particular attention to the simulation of snow accumulation and melt in addition to runoff.

3. Study strategies for regionalization of the parameter sets previously used in literature, and evaluate it´s goodness and uncertainty for the Narayani basin setup.
4. Test the model for a number of catchments not included in the calibration on the ability to predict the duration curve, time series of flow and extremes.
5. Do a test run for a mountainous unmeasured catchment and compare the regional prediction with traditional scaling approaches.

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 11th of January 2016

____

Knut Alfredsen

Professor

Preface

This thesis report entitled “Regional modelling of Narayani Basin in Nepal” is submitted to the Department of Hydraulic and Environmental Engineering at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway as a partial fulfillment of the M.Sc. degree requirements in Hydropower Development.

This report is prepared as an outcome of the study carried out from January to June, 2016 at NTNU under the supervision of Prof. Knut Alfredsen. Required data for the study were collected during the summer field visit to Nepal in 2015.

Finally, I hereby declare that the work presented here is my own and all outside contributions have been acknowledged.

Anugya Sapkota

June 2016

Trondheim, Norway

Acknowledgement

First of all, I would like to express my deepest gratitude to my supervisor Professor Knut Alfredsen whose continual presence and guidance made it possible for the successful completion of this thesis report.

I am indebted to Mr. Yisak Sultan Abdella and Mr. Kuganeshan Sivasubramaniam, PhD at the Department of Hydraulic and Environmental Engineering for their assistance while using SHyFT.

Special thanks to Mr. Abebe Girmay for his assistance throughout the study, especially while using Arc-GIS.

I would like to thank all my professors for sharing their knowledge and experience and making the study at Norwegian University of Science and Technology valuable.

My acknowledgement to Mr. Sagar Ghimirey for suggesting ideas and useful information during the study period. Many thanks to my friends who made my two years stay in Trondheim beautiful.

I am grateful to my husband Mr. Vishal Bista for his moral support, encouragement and help while preparing this report.

I am very thankful to my father and mother for continuously encouraging me to strive for excellence.

Abstract

The two key factors of hydropower generation are discharge and head. Intake location at the higher elevation (head) in a perennial snow-fed river (discharge) will give the best scenario of power production. Nepal is therefore, a perfect hub for hydropower generation. The Himalayan catchment in Nepal which forms an eternal source of the snow-fed rivers also have a steep profile. However, the lack of runoff gauging stations in these locations have constrained the utilization of this colossal source. The flow from these ungauged basins can be achieved by using regional rainfall-runoff model. Hence, the main objective of this study was to create a calibrated parameter set using the regional modelling principle for the ungauged catchments in Narayani basin of Nepal using SHyFT as a modelling tool.

Three cases were examined for different inputs in the model. In Case I, the station data from Nepal was used as input data. The regional modelling was carried out for four main catchments having east-west coverage in Narayani Basin. The average Nash Sutcliffe Efficiency, R^2 of this calibration was -1.34, a poor fit. A high precipitation accumulation was observed for two stations from the nearby Pokhara Airport. Hence, it was expected that the change in the precipitation input would remove this problem. This was checked in the next cases.

In Case II, the ERA-Interim dataset was used as the input in the model. The result from this calibration reduced the precipitation accumulation in the catchment and also improved the R^2 value to 0.39 but this calibration generated a very low areal precipitation which further made the input precipitation data doubtful.

Case III was carried out using the station data for precipitation by eliminating the data which contributed to the local precipitation effect. The number of catchments was reduced to three to get a better fit with the available data. This improved the R^2 value to 0.66 which was the best fit obtained. The validation of the model could also generate the R^2 of 0.71 and the results simulated well with the recession curve.

Altogether 7 sets of free parameters were obtained from the model. The parameter set of Case III was further used to generate the runoff from the two ungauged catchments with varying properties. This model was checked for the ability to generate the time series of flow, flow duration curve, extreme events and snow-melt property of the ungauged catchments.

The obtained model was able to simulate the low flow periods in both the catchments. However, the flow duration curve and extreme events obtained from the model was not able to recreate the high flows. Use of the traditional scaling approach in the high flows was able to fix this problem. Snow-melt obtained from the model was good for the Himalayan catchment. But

further improvement in the parameter value is expected to generate a better snow-melt outflow from the catchment.

The model was promising tool for the generation of the runoff from ungauged basin. However, the uncertainties can be reduced with good quality data and further research on the catchment characteristics.

Table of contents

List of figures

List of tables

Abbreviations

API	Application Programming Interface
BOB	Bay of Bengal
BTK	Bayesian Temperature Kriging
DEM	Digital Elevation Model
DHM	Department of Hydrology and Meteorology
ECMWF	European Centre for Medium Range Weather Forecast
ENKI	Dynamic Environmental Model Framework
ERA	European Reanalysis
ESRI	Environmental System Research Institute
EPSG	European Petroleum Survey Group
GIS	Geographic Information System
HBV	Hydrologiska Byråns Vattenbalansavdelning
ICIMOD	International Centre for Integrated Mountain Development
IDW	Inverse Distance Weighting
KGE	Kling Gupta Efficiency
masl	meters above sea level
NSE	Nash Sutcliffe Efficiency
NetCDF	Network Common Data Frame
SCA	Snow Covered Area
SDC	Snow Depletion Curve
Sintef	Stiftelsen for industriell og teknisk forskning
UTM	Universal Transverse Mercator
YAML	Yet Another Markup Language

1. Introduction

1.1 Background

Finding the outflow from the ungauged basin is a major challenge in practical hydrology. It is even more challenging in the Himalayan catchments where the gauging stations are not present homogeneously. The gauging stations are concentrated in the lower elevation zone which completely makes the upstream process invisible. The permanent ice-reserves of the Himalayas which forms a perennial source of the snow-fed rivers makes it impossible to ignore the upstream feature in these river catchments. The traditional approach used currently is unable to incorporate these characteristics of the catchment. Therefore, it is necessary to regionalize the process.

The head demanding projects such as hydropower development is even more sensitive to the generation of runoff from an ungauged basin. The two keys of Hydropower generation: Head and Runoff depends on the location of intake. The best scenario is when the intakes are located at the upstream of a snow-fed river. But the few gauging stations in the upstream area relates again to greater challenge of acquiring a quality data. Therefore, it is essential to obtain more accurate results for an ungauged intake basin with higher degree of reliability. In order to achieve that, regionalization of the parameters can be applied by setting up a model calibration using several catchments in the region so that the runoff generated will be valid for the other ungauged catchments in the same region. Therefore, the purpose of this study is to generate a reliable runoff in the ungauged basin using the regional rainfall runoff model for Narayani Basin in Nepal using the Statkraft Hydrological Forecasting Toolbox (SHyFT).

SHyFT is an open source tool that is equipped for regional modelling setup and calibration using climate data (precipitation, temperature, relative humidity, radiation and wind speed), physiographic data (reservoir, lake, forest and glacier fraction) and calibration data (discharge).

1.2 Objectives

The main objective of this study is

- To generate reliable runoff from an ungauged basin of Narayani River using SHyFT as a hydrological modelling tool.
- To check the fit of observed runoff against the simulated runoff.
- To further use the model to generate the runoff from other ungauged basin of the same area.

1.3 Scope

The model is set and calibrated for Narayani River Basin in Nepal. The study aims to fulfil the systematic task through the following steps listed in order:

1. Literature review on regionalization of parameters used previously that will fit with the purpose of this study.
2. Preparation of input data for SHyFT using,
   a) Maps including digital elevation models and land use parameters (reservoir, lake, forest and glacier cover) for Narayani basin.
   b) Quality controlled discharge data and selection of the outlet stations based on the same.
   c) Filling up of missing data and quality control for climate data (precipitation, temperature, radiation, wind speed and relative humidity)
   d) Conversion of data into 'netcdf' format using the SHyFT model routine.
3. Calibration of the model for different number of sub-catchments for a time series and validation of the same sub-catchments for different periods.
4. Evaluation of the performance of the model for snow accumulation and snow melt.
5. Running a test for a mountainous ungauged catchment and compare the result with traditional scaling approach for different time series and extreme flow cases.

1.4 Structure of thesis

This thesis report has been prepared on the same order as the tasks performed during the study. Following breakdown gives a brief overview of contents of each chapter.

Chapter 1:

Introduction to the study, methodology adopted and limitations experienced.

Chapter 2:

Discussion on the hydrological theories used in the study.

Chapter 3:

Introduction to the study area in with regards to its location features and hydro-meteorology.

Chapter 4:

Collection and processing of the input data for SHyFT.

Chapter 5:

Use of Arc-GIS in the study.

Chapter 6:

Use of modelling tool SHyFT in the study.

Chapter 7:

Selection of the calibration cases, model validation and comparison between those cases.

Chapter 8:

Illustration of the results obtained from the model.

Chapter 9:

Analysis of the result and further recommendation on the model usage.

1.5 Limitations

Due to the number of factors there are some limitations experienced during the study. Table below gives the summary of those limitations and their consequence on the study.

Table 1-1: Limitations experienced during study

Limitations	**Impact**
Limited number of hydro-meteorological stations on the leeward side	Process still missing complete regionalization feature
Limited knowledge of Himalayan-snow melt process (snow depth, reflectivity)	Gamma-snow parameter used was adopted from the Norwegian catchment, which might have altered the output result by some degree
Lack of documentation for SHyFT model	Challenges in the selection and alteration of the calibration parameters and in extraction of results

2. Literature Review

In this chapter the theories used in the thesis will be reviewed.

2.1 Runoff from an ungauged basin

Predicting the runoff from an ungauged station has always been a major challenge. Different approaches are used for this purpose depending on the sensitivity of the task undertaken. The core idea however is to take the hydrological data from the nearby gauging station and link it to the ungauged catchment by using some equations and processes. Spatial interpolation was one of the earliest methods used for this purpose. This process involves the inverse distance weighting interpolation method to directly estimate the runoff by transferring the runoff from the donor catchment to the ungauged catchment (Zhang et al., 2014). For completing the process, the distance between the neighboring gauging stations are assigned with certain weightage, higher weightage for the nearest station. This is a very rudimentary approach with a narrow scope, therefore for a big catchment with highest variability, hydrological modelling tools have evolved as a very promising and powerful tool.

2.2 Hydrological modeling

Hydrological modelling is the process of the rainfall-runoff modelling using some hydrological regimes and interpolation approach. Basically, the purpose of this process is to extrapolate the available parameters (measurement) in both time and space. Since, the measurement technique is already very limited for the hydrological cycles, the use of the rudimentary approach will further constrain this in space and time especially for the ungauged catchment, where the data availability is already limited in the first place (Beven, 2011).Nonetheless, the hydrological modelling tools can provide the reasonable extrapolation in this situation also. Therefore, it can be helpful in estimating the hydrological dynamics, the water balance and the statistics of the hydrological variables at the ungauged sites. One of such tools is regional runoff-rainfall models.

2.3 Regional-modelling concept and approach to deal with uncertainty

In all the regional models, the basic concept is the parameterization of the hydrological processes with an approach to transfer the parameterization from gauged to ungauged catchment. While doing so, there are some uncertainties associated with the quality of data (inputs and calibration) and model specifics (model parameters and model structures) which is always challenging. Hence, it is important to justify these uncertainties in a regional calibration by using some approaches like multiple regression, spatial interpolation of the regional parameters, proxy basin test etc. It is also necessary to know that these models are never perfect. A regional hydrological model can be trusted with some confidence within a region only when the regionalization parameter sets are identical to all the catchments which again connects to justification of the uncertainties (Engeland et al., 2006). The performance of these approaches is also necessary to evaluate the calibrated parameters. This can be achieved using some evaluation criteria such as Nash Sutcliffe efficiency (NSE), Kling Gupta efficiency (KGE).

2.4 Previous studies

Previous studies have been carried out in the proposed study area using some hydrological modelling tools like ENKI. ENKI uses Bayesian method to estimate the probability distribution of parameters. This method considers the uncertainty in the choice of the parameter value. The model is executed for the several numbers of catchments and iterated for possible range of several free parameters. These calibrated parameters are judged by Nash efficiency.

This study will be carried out using SHyFT. SHyFT is flexible between the use of Bayesian method and Inverse Distance weighting method along with evaluation criteria to choose between NSE and KGE. More about the hydrological tool is discussed separately in a single chapter later.

2.5 Choice of study area

The study area has a range of climatic variations and big tributaries. The origin of these tributaries are mostly in Nepal which makes data acquisition easy. Also, some of these tributaries are snow-fed and some are monsoon generated, which make them very distinct. This area is also the hydropower hub of Nepal, hence the result from study would actually be useful in the study area. Also, it is the first time that SHyFT has been used in the proposed study area, therefore, it is expected to highlight some ability of the model along with the catchment characteristics.

3. Study Area

In this chapter, the location, climatic conditions and hydro-meteorological characteristics of the study area will be discussed.

3.1 Background

Nepal is a Himalayan country stretching between 26° 12’ and 30° 27’ North in latitude and 80°04’ and 88°12’ East in longitude. It is bordered by China on the North and India on East, West and South as seen in Figure 3-1.

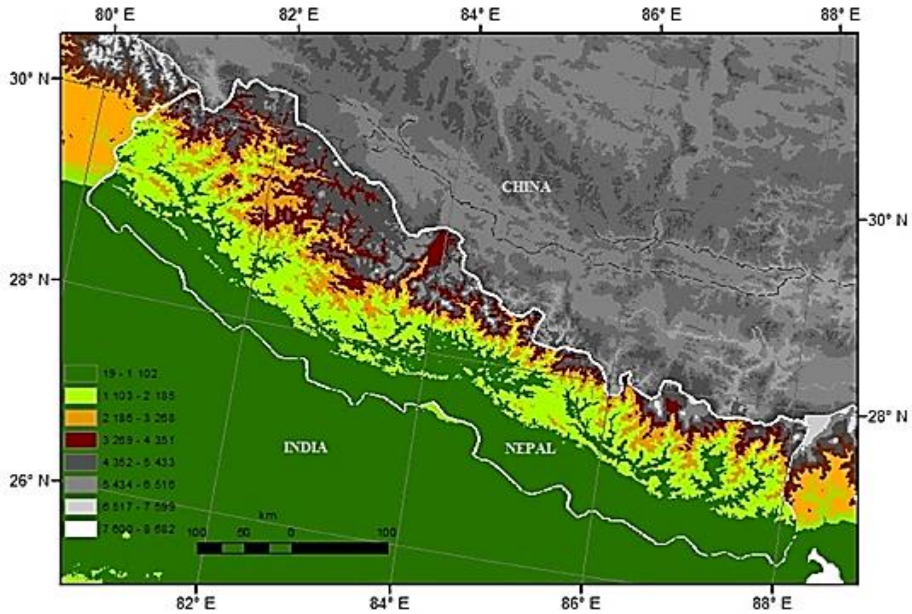


Figure 3-1: Geo-physical map of Nepal

Nepal has more than 6000 rivers (Pande, 1987) most of which originate from the high mountains and flow down all the way to the Ganges in India.

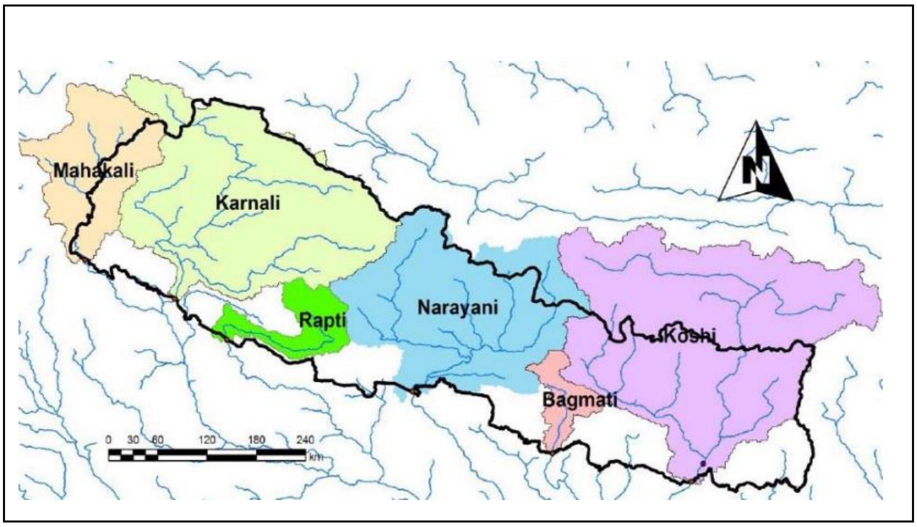


Figure 3-2: Major River Basins of Nepal (Shrestha et al., 2012)

These river systems have been classified into seven major river basins for hydrological study as seen from Figure 3-2 above. They are Kankai Mai River Basin, Sapta-Koshi River Basin, Bagmati River Basin, Narayani River Basin, Karnali River Basin, West-Rapti River Basin and Mahakali River Basin. Kankai Mai River basin is very small river basin on the eastern side of Nepal. Narayani Basin is the area of study.

3.2 Narayani River Basin

3.2.1 Location

Narayani River Basin lies in the central part of Nepal between latitude 29°17′0″N to 25°39′9″N and longitude 85°50′5″E to 85°11′4″E. It is a partly snow-fed river system with the elevation ranging from 8000 masl in the north to 60 masl in the south. Like other river systems in Nepal, Narayani basin also flows into the Ganges in India. The total catchment area of Narayani basin is about 35,000 sq kms (Jeeban et al., 2015) of which 26,300 sq kms falls in Nepal (Pande, 1987)

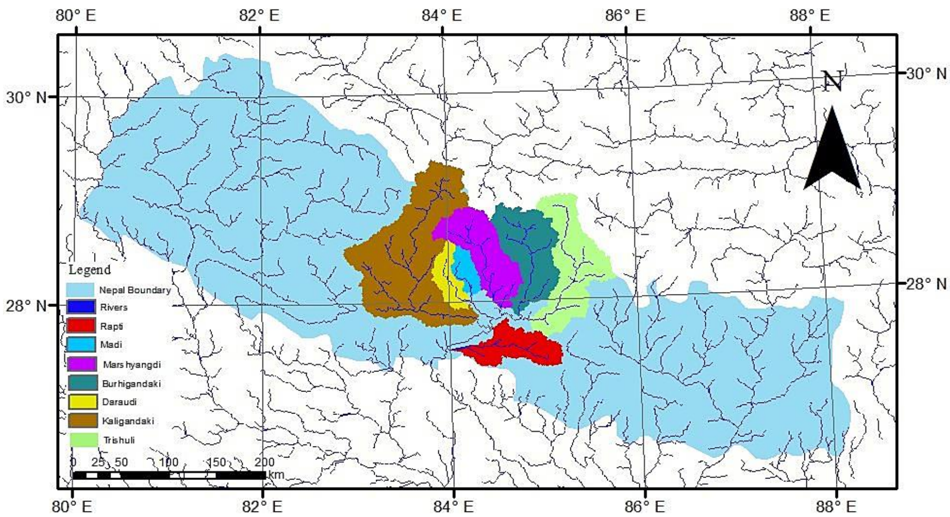


Figure 3-3: Seven tributaries of Narayani Basin

As seen in the figure 3-3, Narayan Basin has seven major tributaries; Kali Gandaki, Burhi Gandaki, Madi, Marshyangdi, Trishuli, Rapti and Daraudi. Of these, Trishuli and Burhi Gandaki originate from Tibet. The largest tributaries with respect to the catchment area is Kali Gandaki (9,300 sq kms) and the smallest is Daraudi (900 sq kms).

Narayani Basin is also the hydropower development capital of Nepal. Of the total electricity generated in Nepal, 44% of electricity is developed from this river basin. The number of lakes and glaciers in this river basin is 338 and 1025 respectively (Bajracharya et al., 2011). It is also one of the major ice reserves in the country 191.39 cubic kms and has the largest area covered by the glaciers.

The summary of the tributaries with their origin and area is given in the Table 3-1 below.

Table 3-1: Information about the tributaries in Narayani Basin

Tributaries	Source	Catchment Area [Sq. kms]	Confluence
Trishuli	Tibet	4000	Devghat, Chitwan
Burhi Gandaki	Tibet	3700	Benighat, Dhading
Daraudi	Gorkha	900	Kharahani, Tanahu
Marshyangdi	Damodar Himal	4600	Mugline, Tanahu
Madi	Annapurna Himal	1100	Damauli, Tanahu
Seti Gandaki	Annapurna Himal	2700	Kandrang, Tanahu
Kali Gandaki	Mustang	9300	Devghat, Chitwan

3.3 Hydro-meteorological stations

The location of all the hydro-meteorological stations (precipitation, temperature and runoff) used within the study area is presented in Figure 3-4 below. All the hydro-meteorological stations used in the figure are assigned by the Department of Hydrology and Meteorology in Nepal.

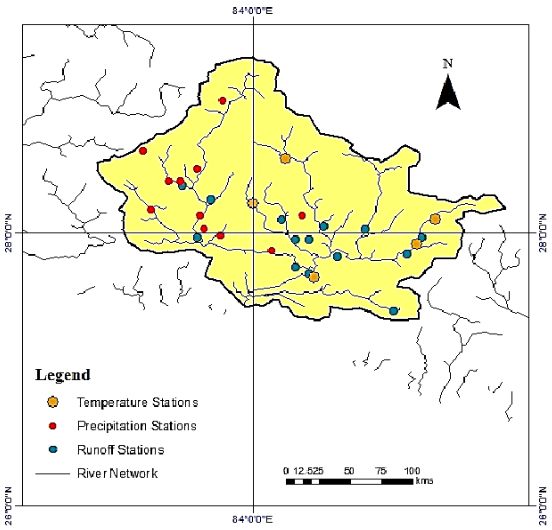


Figure 3-4: Hydro-meteorological stations in study area

3.4 Climate

The study area has been divided into five physiographic regions based on elevation to account the climatic variations. These are listed below in Table 3-2 and Figure 3-5.

Table 3-2: Physiographic regions of Nepal (Shrestha and Aryal, 2011)

Physiographic zone	Dominating Climatic zone	Elevation variation [masl]
Terai plain	Tropical	100 – 200
Siwalik range	Sub-tropical	700 – 1500
Middle mountains	Sub-tropical and Temperate	1500 – 2700
High mountain	Temperate and Subalpine	2000 – 4000
High Himalayas	Alpine and Nival	4000 – 8848

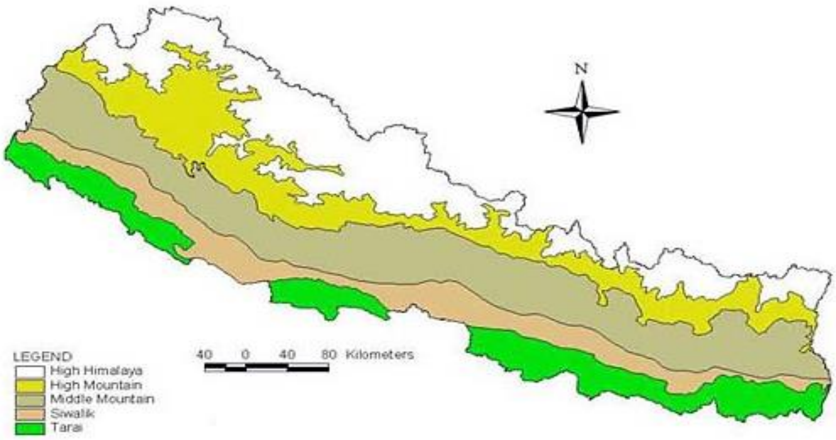


Figure 3-5: Physiographic division of Nepal (Pariyar, 2008)

Climatic variation in the study area occurs due to the combined effect of radiation, wind speed and elevation difference. The effect of radiation and elevation difference makes the northern slope become more humid than the southern slope. The wind speed creates an orographic effect, which further influences the rainfall regime in the area. Details on the effect due to radiation and wind speed is discussed separately below.

3.5 Precipitation

Precipitation is dominated by the south-easterly monsoon associated with the change in direction of the seasonal winds during the summer months (June to September). Due to the topographical variation, precipitation varies significantly from place to place as seen in Figure 3-6. The Siwalik range first intercepts the approaching monsoon wind and receives the heavier rainfall. This rainfall increases with altitude on the windward side and abruptly decreases on the leeward side of the Tibetan plateau. Monsoon precipitation occurs in the solid form (snow/ice) in higher altitude (above 5000 masl) and is the major source of summer accumulation-type glaciers (Shrestha and Aryal, 2011). On the other hand, winter precipitation (between February and April) caused by westerly weather system originating from the Mediterranean Sea plays a secondary role in mass balance of the glaciers.

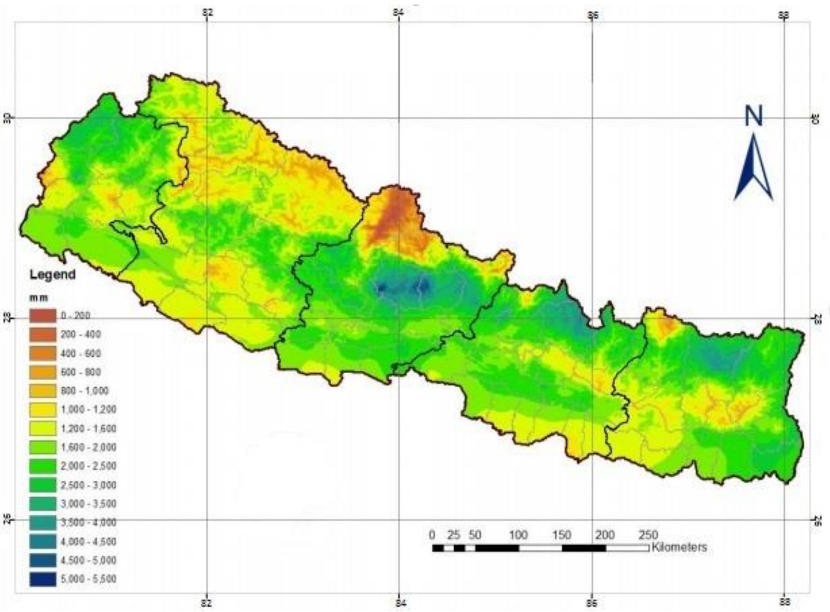


Figure 3-6: Spatial distribution of average annual rainfall

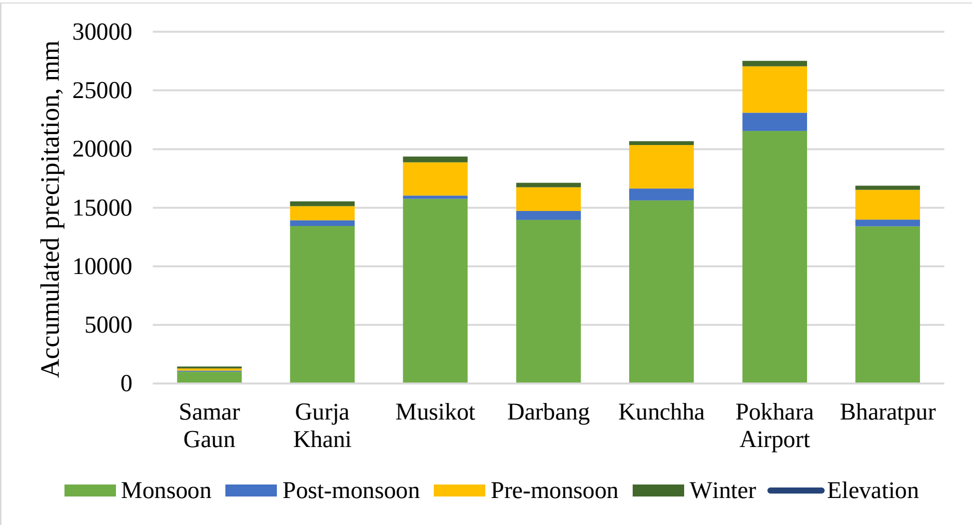


Figure 3-7: Seasonal distribution of precipitation on the basis of elevation

From the Figure 3-7 figure it can be seen that the monsoon is the wettest season receiving nearly 85% of the total annual precipitation. The driest season is winter with the occurrence of only 1% of annual precipitation.

3.6 Temperature

Temperature distribution in Nepal is mainly influenced by altitude. It varies considerably North to South depending on the local wind and the aspect of slope, also seen in Figure 3-8 The mean annual temperature in Terai and lower Siwalik range is above 20°C. In the upper Siwalik zone and Middle Mountains temperature ranges from 15°C to 20°C. The mean annual temperature remains below 10°C in High Mountains and in High Himalayas, it remains below 5°C. (Basnet, 1989-1992).

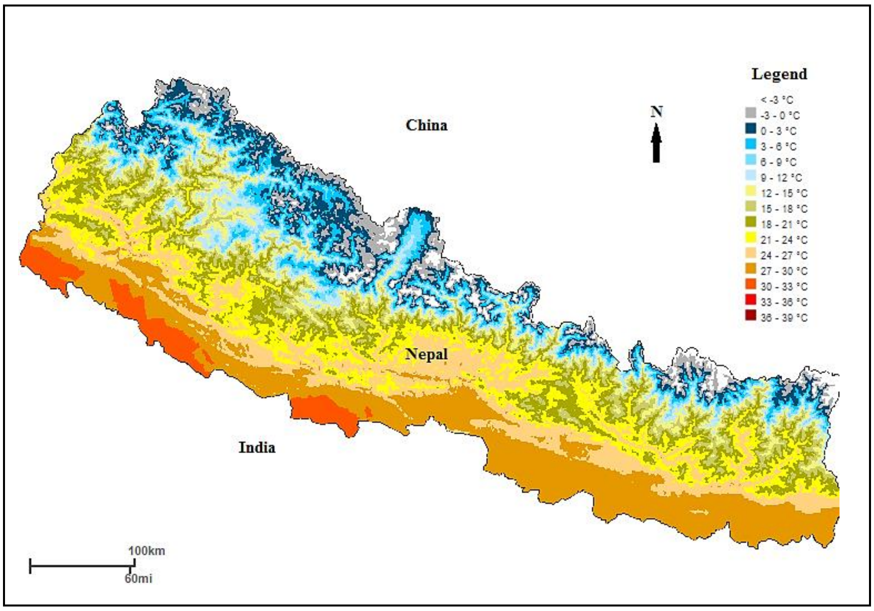


Figure 3-8: Average annual temperature map of Nepal(ICIMOD, 2005)

The summer season (April to September) and is very hot with the mean temperature of 25°C in Terai. In the other zones, temperature is mild even in summer. Winter season (November to February) is characterized by very cold condition in hills and mountains and mild temperature in Terai. The lowest minimum temperature is experienced in January throughout the study area.

3.7 Relative Humidity

Relative humidity is the measure of ratio of partial pressure of water vapor to the equilibrium vapor pressure at the same temperature. As also seen from Figure 3-9, it is a function of pressure and temperature of the particular place. In Nepal, due to the high precipitation and temperature during the monsoon season, relative humidity is relatively higher. It is very low in the dry months (winter and post monsoon). During summer, relative humidity can be very high in Terai belt and very low in the arid leeward side of the mountain at the same time.

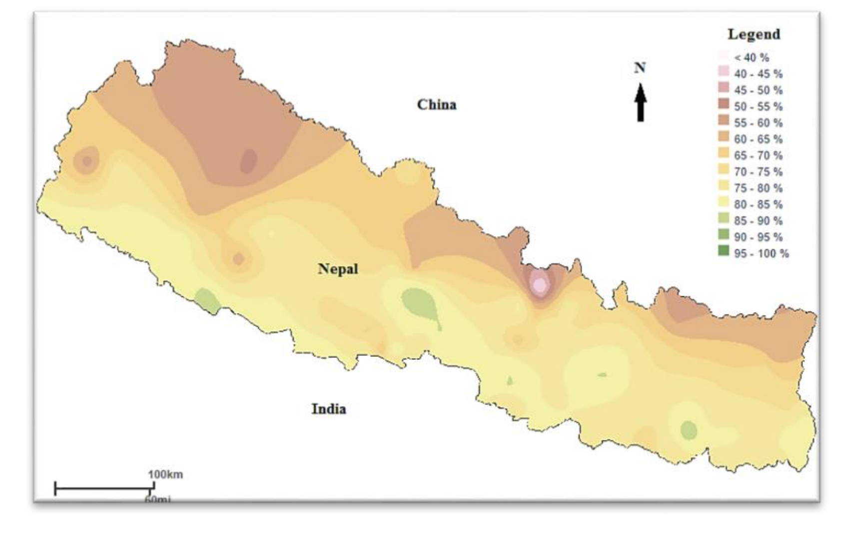


Figure 3-9: Relative humidity over the year in Nepal (ICIMOD, 2005)

3.8 Radiation

In Nepal, solar radiation is observed to be higher at the higher altitude in a clear day. This value is higher for the southern slope as Nepal falls in the Northern hemisphere. Solar radiation also depends on the climatic condition as the radiation value is higher for most part of Nepal during the summer season and lowest during the winter.

3.9 Wind Speed

Wind plays a vital role in precipitation and local wind effect might affect the temperature regime of a particular area. Due to the high altitude variation (nearly 7,000 m) in 200 kms North to South, different wind regimes are observed throughout the country. From the existing 29 functional wind stations, located at different parts of the country, it is seen that wind speed in low altitude (less than 3000 masl) are lower in magnitude than those in high altitude valleys (higher than 3000 masl) and mountain ridges. (B.N. Upreti, 2009). The Terai plains have in general low wind speed than the mountain regions.

Rainfall in Nepal is highly dominated by the orographic rainfall system due to the high Himalayas located on the North of Nepal. The northwestward wind that flows from the head of Bay Of Bengal (BOB) transports moisture to Nepal. The meridional wind component is dominant over the zonal wind component near the southern edge of Nepal. The influence of this summer circulation patterns is unequally distributed over the Himalayas, with greater rainfall in the central-eastern part and less in the northwestern part.

3.10 Land Use

In this study, the land use focus was on the forest cover, lake cover, glacier cover and reservoir cover. In Nepal, 5.75 million hectares of area (39.1% of the total land) is covered by forests. 4.5 thousand hectares is covered by lakes which is 0.03% of the total land area. Similarly, glaciers comprise of 1.2 million hectares of area of which 93% lie in the high mountains (Uddin et al., 2015). There is only one hydropower reservoir in Nepal with the area of 220 hectares. Figure 3-10 shows the integrated land-use map of Nepal.

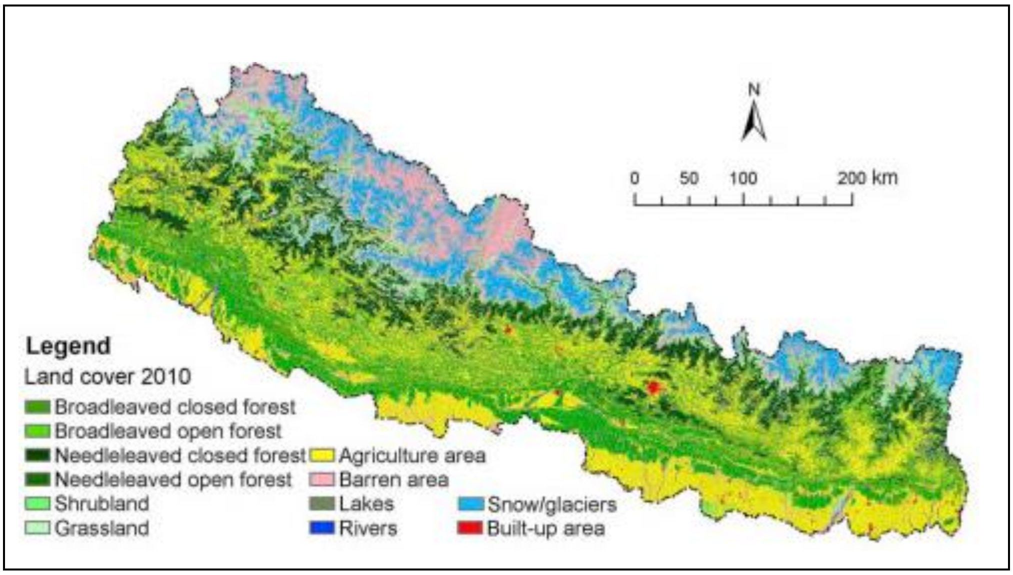


Figure 3-10: Land use map of Nepal (Uddin et al., 2015)

4. Data collection and processing

This chapter discusses the step-wise procedure of input data processing for hydro-meteorological records of the selected stations.

4.1 Introduction

Meteorological and hydrological data needed for the study was collected from the stations assigned by Department of Hydrology and Meteorology, Nepal. Altogether there are 282 meteorological stations and 51 hydrological stations all over Nepal. The data from the period of 2005 to 2012 is taken for the study purpose, the first five years of data was used for calibrating the model and the remaining 3 years of data for the model validation.

4.2 Meteorological data

4.2.1 Precipitation

Of the 282 meteorological stations, data from 13 precipitation stations was taken. The stations were selected on the basis of data availability for the given period, i.e 2002-2014. The list of stations is given in appendix 1 and the data availability is listed in the Figure 4-1 below.

4.2.1.1 Missing data interpolation

In order to input the precipitation data in to the model it was necessary to check the continuity first. The data was plotted against the years and missing years were noted. Some figures denoted by 'T' in the series was confirmed as trace precipitation and was replaced by 2.5 mm (Mekis, 2005).

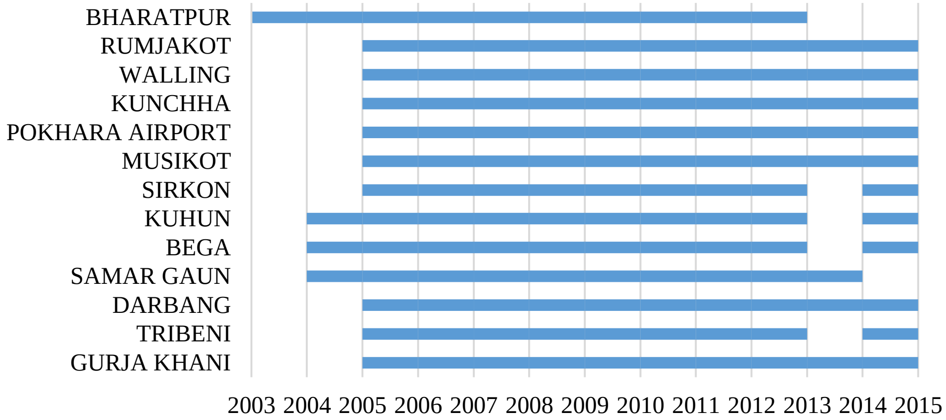


Figure 4-1: Precipitation data availability

These missing data were then interpolated using the following Normal Ratio Method from the nearby stations. (Dingman)

$$p_o = \frac{1}{G}\sum_{1}^{G}\frac{P_o}{P_g} * p_g$$

Where,

p_o = Estimated missing data for a particular day at the gauging station

P_o = Annual average precipitation of the gauging station with missing value

p_g = Observed precipitation for corresponding day at g = 1, 2, 3 …G

P_g = Annual average precipitation of the nearby station

This formula is valid for the stations where annual average precipitation differs by more than 10%. For simplicity, it was assumed that the annual precipitations for all the stations will differ by more than 10%. The stations with a yearlong missing data were left out in the process. The correlation between the stations, besides the annual average was checked before using them to fill up the missing data. Excel sheets were used by constructing simple macro-tools and the filling up process was carried out successfully. Some data which were suspicious were cross checked with the satellite precipitation (www.worldweatheronline.com) and the quality was confirmed hence after.

4.2.1.2 Data Quality Check

Changing of the gauging stations, exposure to the vegetation growth, instrumentation change may lead to the change in the precipitation data. Therefore, it is a must to check the quality of data before using them for further analysis. Trend analysis helps to check possible errors in the data. The trend analysis was carried out by plotting the sum of monthly precipitation of 7 stations with completely filled-up data. The plot is shown in Figure 4-2 below.

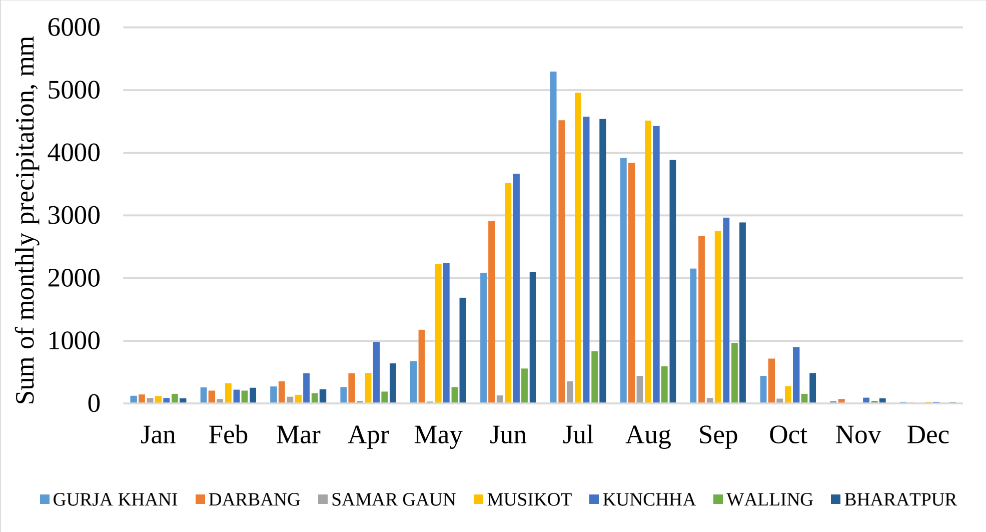


Figure 4-2: Sum of monthly precipitation for the selected stations

From the graph it can be seen that for all the stations, maximum precipitation is observed during the monsoon season (June to September) and the least precipitation is observed in December. The data from the above selected station is further checked for consistency using the seasonal precipitation distribution, also another trend analysis method as in Table 4-1.

Table 4-1: Seasonal distribution of precipitation

Station	**Winter (Dec-Feb)**	**Pre-Monsoon (Mar-May)**	**Monsoon (Jun-Sept)**	**Post-Monsoon (Oct-Nov)**
Gurja Khani	3%	8%	87%	3%
Darbang	2%	12%	81%	5%
Samar Gaun	12%	13%	70%	6%
Musikot	2%	15%	81%	1%
Kunchha	2%	18%	76%	5%
Walling	9%	15%	71%	5%
Bharatpur	2%	15%	79%	3%

While checking for the seasonal distribution also, more than 70% of the precipitation was found to occur during monsoon. This shows that the data from the stations are not skewed in terms of uniformity.

In order to confirm whether or not the gauging station was subjected to the change during the period of study (2005-2012), accumulation curve was plotted for the resulting stations which can be seen in Figure 4-3 below.

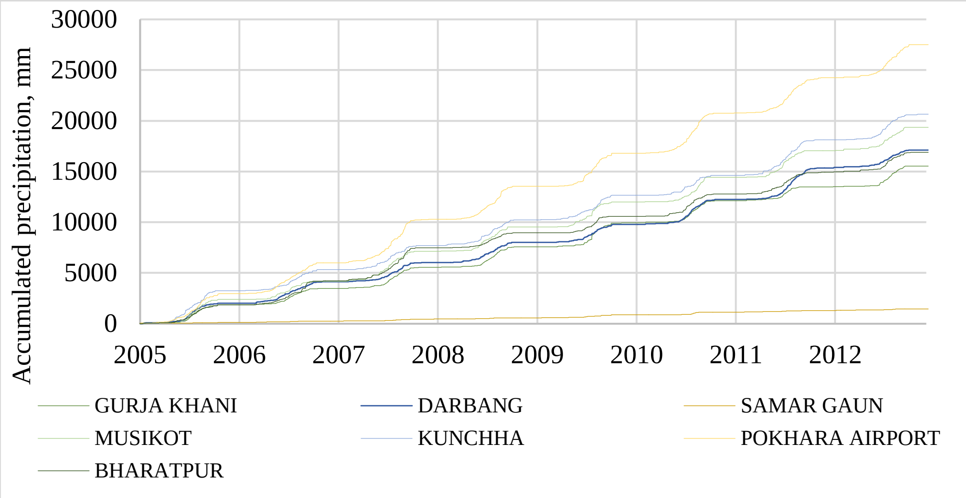


Figure 4-3: Accumulated precipitation of the stations

Previously, the accumulated runoff graph for all the stations except Walling, was found consistent. The data from Walling was rechecked and it was found that the precipitation in 2012 was higher than the precipitation of any other years between 2005 and 2011. Hence, Walling was not considered for further analysis. Instead of Walling, Pokhara Airport was checked for the aforementioned process and was found suitable for further analysis.

4.2.2 Temperature

Temperature data are required to compute type of precipitation, snow melt and evapotranspiration whenever applicable. Altogether 5 temperature stations were found suitable for the study purpose which are listed in Appendix 1. Daily maximum and minimum values were collected for these stations from Department of Hydrology and Meteorological station in Nepal. The average temperature was taken as the mean of maximum and minimum daily temperature.

4.2.2.1 Missing data interpolation

Available data were then plotted against time scale and the gaps in the data was found out. It can be seen from Figure 4-4 below that the data from 2005 to 2012 is complete for all the stations.

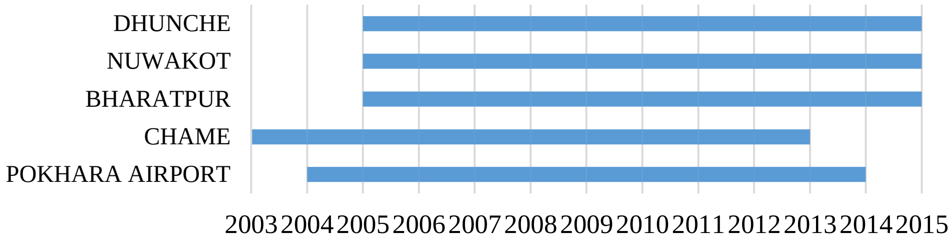


Figure 4-4: Temperature data availability

However, there were some daily data missing in this time series also which was filled using the following formula.

$$T_o = T_g + (\text{avg.}T_o - \text{avg.}T_g)$$

Where,

To = Interpolated daily mean temperature from ungauged station

Tg = Daily mean temperature at known gauging stations g = 1, 2, 3 …g

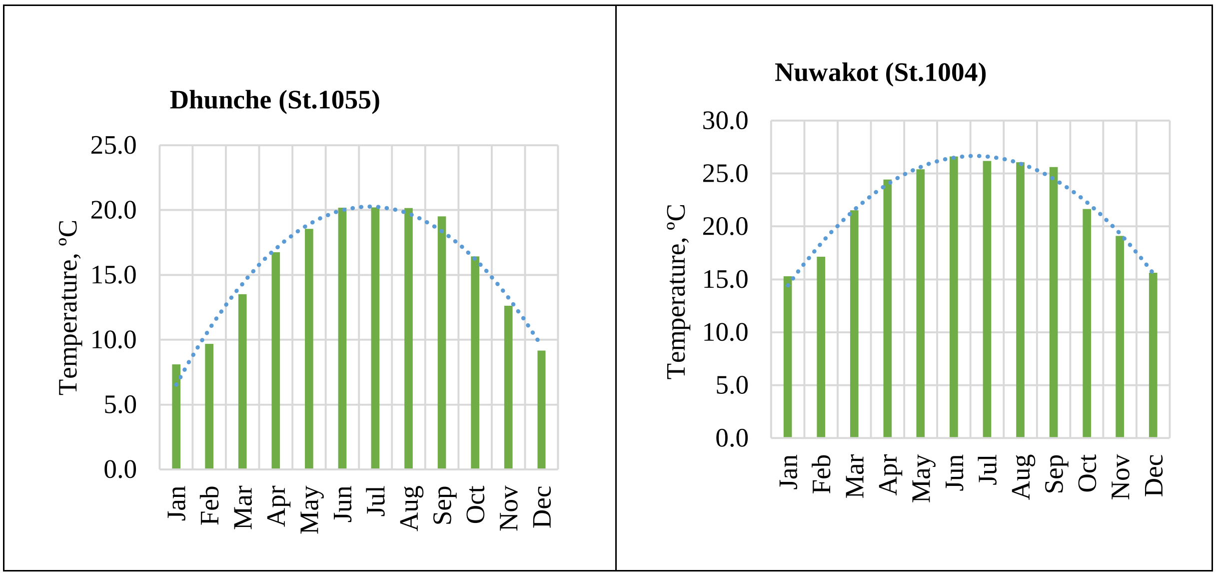
avg.T_o = Average daily mean temperature from ungauged station

avg.T_g = Average daily mean temperature from known gauging stations.

The filling process was carried out using simple formula in excel.

4.2.2.2 Data quality check

Trend analysis was carried out to check the consistency and possible error in the data. For this, the average monthly temperature for each station was plotted. The results are shown in Figure 4-6 below.



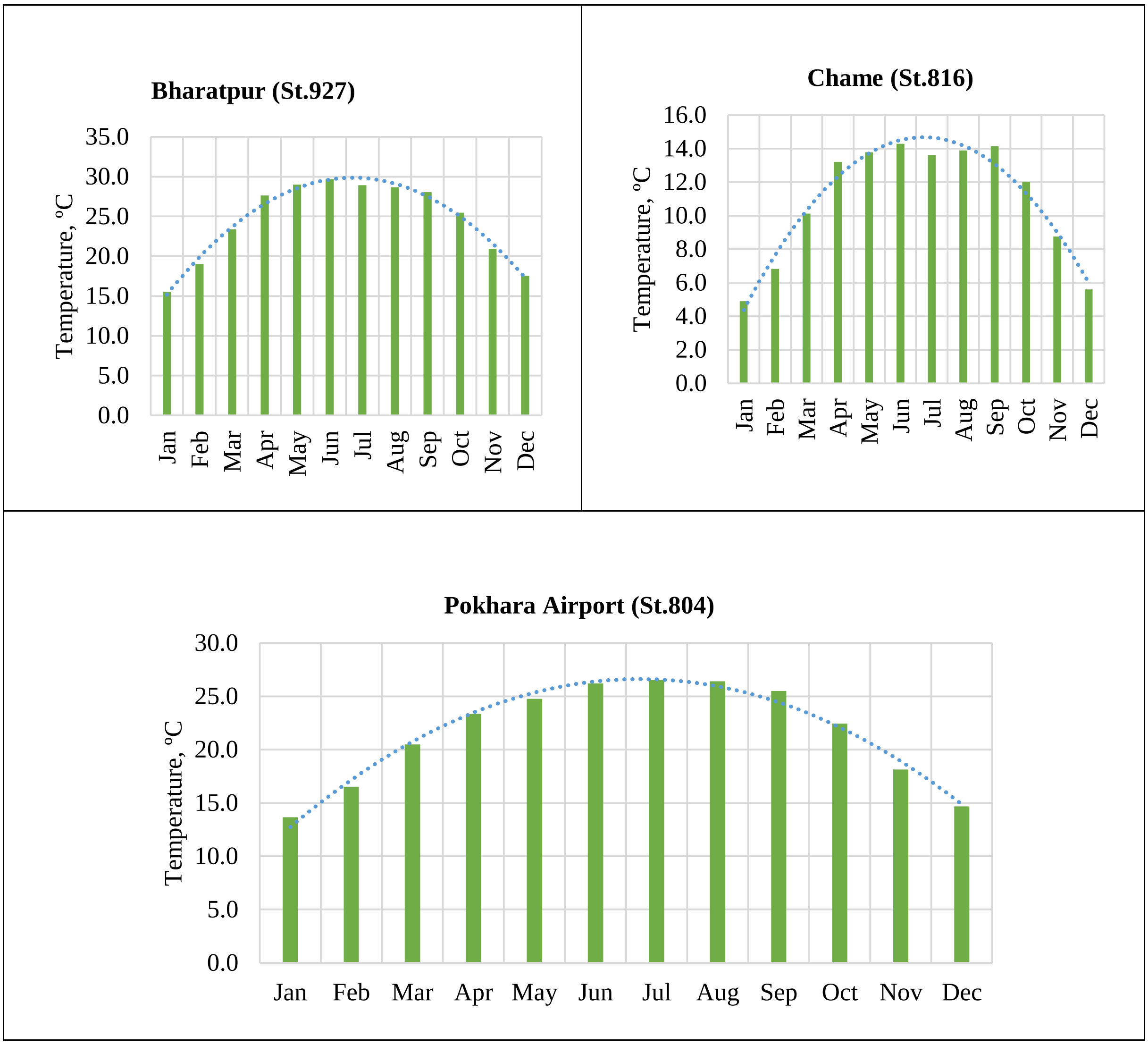
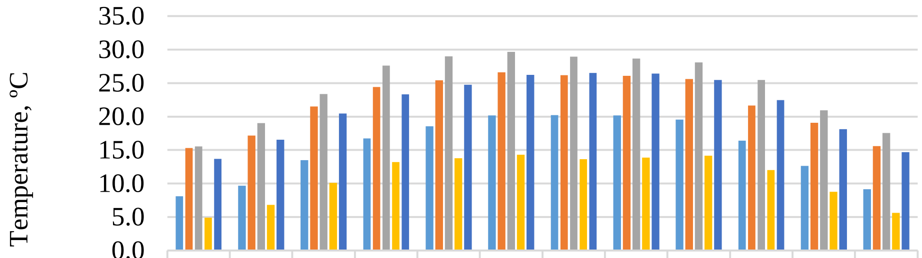


Figure 4-5: Trend analysis observation using mean monthly temperature



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
DHUNCHE	8.1	9.7	13.5	16.8	18.6	20.2	20.2	20.2	19.5	16.4	12.6	9.2
NUWAKOT	15.3	17.1	21.5	24.4	25.4	26.6	26.2	26.1	25.6	21.7	19.1	15.6
BHARATPUR	15.5	19.0	23.4	27.6	29.0	29.7	28.9	28.6	28.1	25.5	20.9	17.5
CHAME	4.9	6.8	10.1	13.2	13.8	14.3	13.6	13.9	14.1	12.0	8.8	5.6
POKHARA AIRPORT	13.7	16.5	20.5	23.3	24.8	26.2	26.5	26.4	25.5	22.4	18.1	14.7

Figure 4-6: Consistency check of the stations using mean monthly averages

From the above Figure 4-6, it can be seen that the monthly distribution of the temperature is fairly consistent throughout the study area between 2005 and 2012. The maximum temperature is observed during the monsoon season and the minimum temperature in the winter season. There is no negative average temperature in any of the stations. Hence, the data from the above 5 stations is good enough to be used for the further analysis.

4.2.3 Solar Radiation, Wind Speed and Relative Humidity

Since ERA-Interim dataset was used as the source of these data, the checking process was not carried out. The filling however was achieved by using SHyFT. In this study, the sensitivity of precipitation and temperature data was more than the other three meteorological data. However, the cross checking of the calibration results from ERA-Interim data was carried out using the data from www.worldweatheronline.com and the quality of data was assured hence after.

4.3 Hydrological data

4.3.1 Runoff

Daily discharge from 14 of the gauging stations distributed all over Narayani Basin was obtained from Department of Hydrology and Meterology. The table in appendix 1 gives the list of the stations under consideration.

The data series available from 2000 until 2014 was plotted against the time scale in Figure 4-7 to check for any missing years.

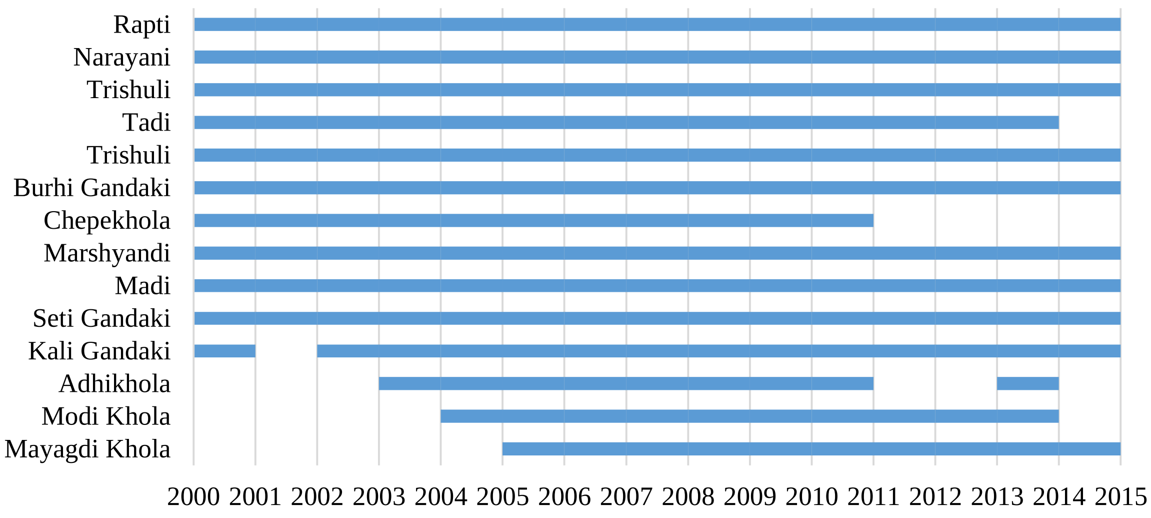


Figure 4-7: Data available for the runoff stations

14 hydrological stations in Narayani basin were used in the study. These stations are well distributed along the tributaries. The data from these 14 stations were selected for the overlapping years to find out the missing runoff data.

4.3.2 Missing data analysis

Runoff data can be very crucial in the hydrological modelling, therefore it is important to fill in the missing data properly. The data gaps may be attributed by number of factors including the effects natural disaster like landslides and floods, equipment failure, mistreatment of the observed data by field personnel. In order to fill up the missing data different methods such as regression analysis, interpolation approaches can be used. For this purpose, the following equation was used.

$$Q_o = \frac{1}{g}\sum_{1}^{g}\frac{AvgQ_o}{AvgQ_g} \cdot Q_g$$

Where,

Q_o= interpolated daily discharge of the unknown station

Q_g = Daily discharge of the known gauging station

$AvgQ_o$ = Average daily discharge of the unknown station

$AvgQ_g$ = Average daily discharge of known station

4.3.3 Data Quality Check

The quality of data is checked by simply plotting the discharge series against the time scale. If the resulting graph showed some flat regions or lots of spikes, then it was understood that there was something wrong with the data series.

The quality checking process was carried out for all the 7 major tributaries of Narayani basin in Table 4-2.

Table 4-2: Station Id of the gauging stations at the tributaries

Station Id	**Name**	**Location**
420	Kali Gandaki	Kotagaun
430.5	Seti Gandaki	Damauli
438	Madi	Shisaghat
439.7	Marshyandi	Bimalnagar
445	Burhi Gandaki	Arughat
449.91	Trishuli	Kalikhola

Station Id	Name	Location
460	Rapti	Rajaiya

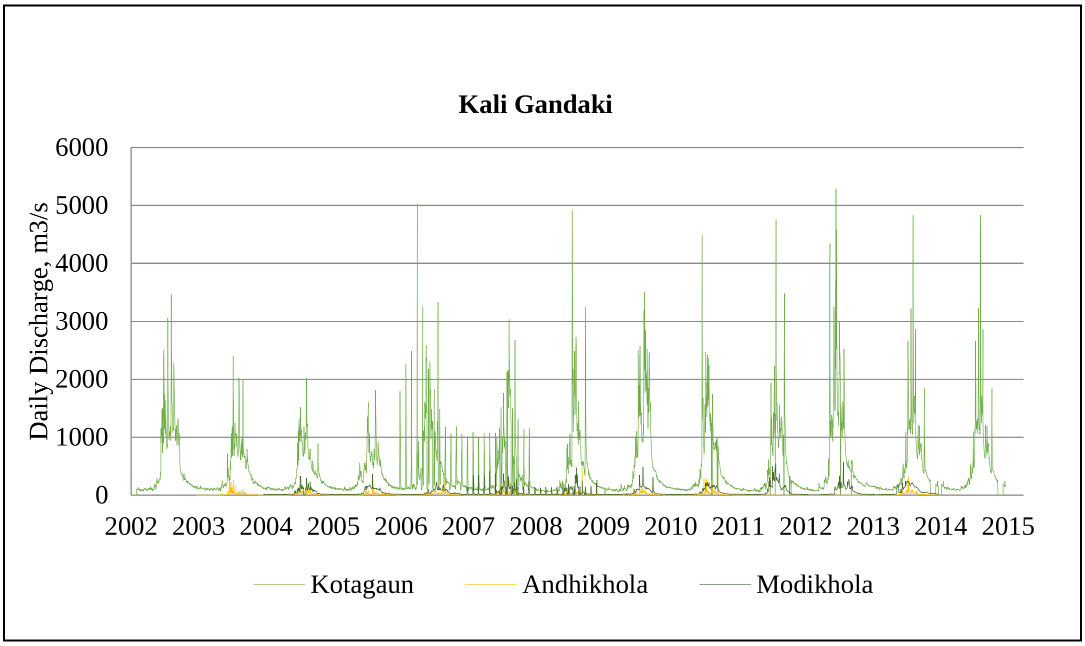


Figure 4-8: Observed daily runoff at Kali Gandaki

In the data series of Figure 4-8, station data from 2006 and 2007 is seen with lots of spikes for Kotagaun and Modikhola. This was checked with the data series of Narayani, which is located further downstream of Kotagaun, it was observed that the data was perfectly normal to the other years data. Also, the station Andhikhola, which is located at the upstream of Kotagaun was also carefully examined and it was found to have a consistency in the runoff between 2006 and 2007 (see appendix 5). Therefore, the data series between the above mentioned years was interpolated from the nearby stations by using the station average of the nearby stations with high correlation.

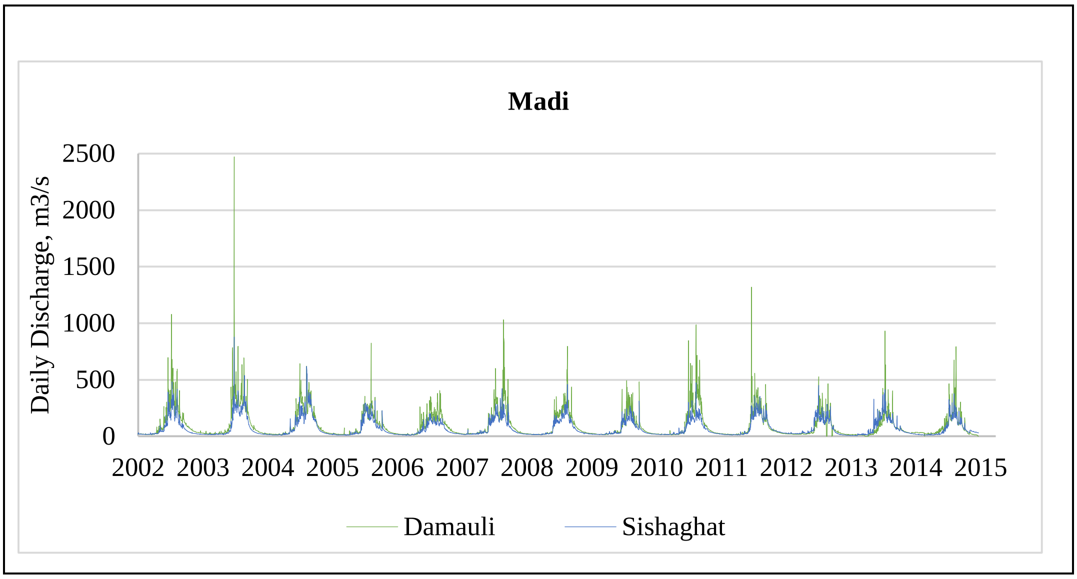


Figure 4-9: Observed daily runoff at Madi

From Figure 4-9, the data for both the tributaries seemed consistent throughout the period of 12 years. A constant high flow in some days in monsoon is observed in Damauli (Seti Gandaki), this is because of the high precipitation in the catchment.

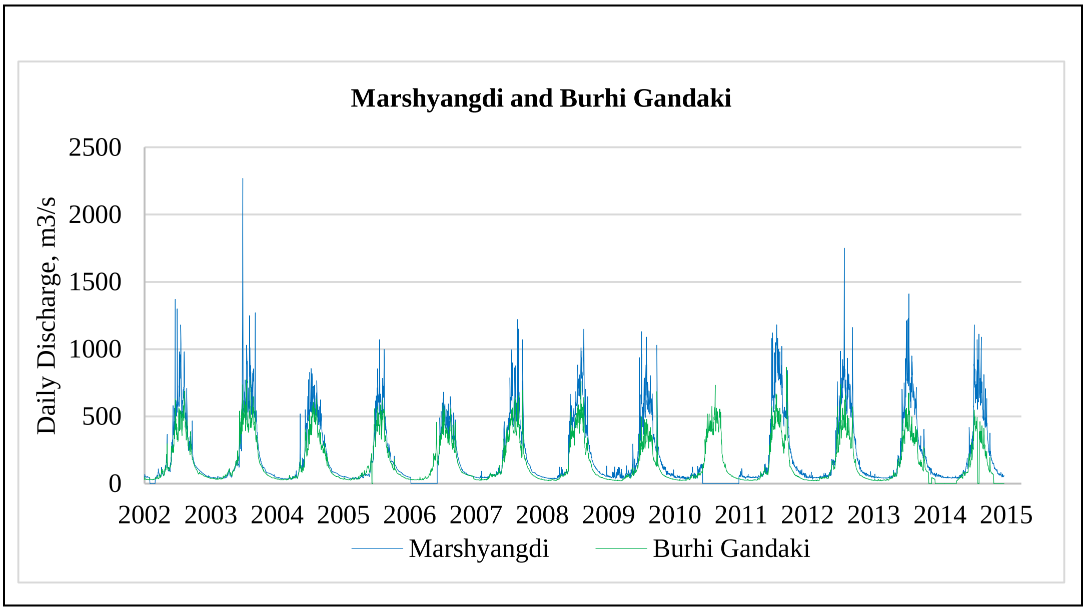


Figure 4-10: Observed daily runoff at Marshyangdi and Burhi Gandaki

Similar to Seti Gandaki, Marshyangdi also has a higher runoff in the year 2003 as seen from Figure 4=10, again attributed by the higher precipitation in the catchment. Other than that, the data looks regular for the given time series.

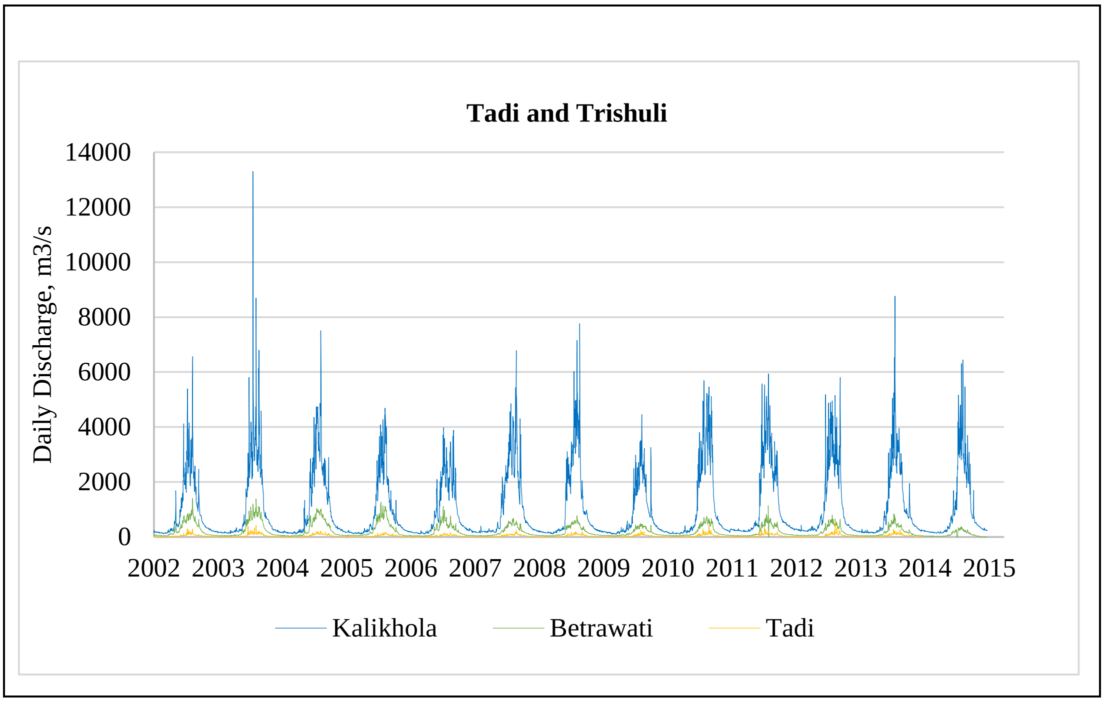


Figure 4-11: Observed daily runoff at Tadi and Trishuli

For this tributary, the data was also checked using the catchment response. It is observed from Figure 4-11 that the response of the Tadi is earlier than Betrawati, which is what was expected because of the catchment size and aspect. This difference is consistently varying throughout the observed time period.

For the final tributary, East Rapti, the data was checked with Narayani between 2005 and 2014 due to the lacking stations for cross checking the data and was found okay (see appendix 5).

4.4 Summary of hydro-meteorological data

As per the quality check of the data, the data series between 2005 and 2012 is found to be good for the hydrological modelling.

4.5 ERA-Interim Dataset

ERA stands for 'European Reanalysis' which produce various datasets and information about the global climate reanalysis carried out by the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA-Interim dataset contains atmospheric and surface parameters with global spatial coverage from 1st January 1989. This includes the gridded datasets of large variety ranging from 3-hourly surface parameters, 6-hourly upper-air parameters, vertical integrals of atmospheric fluxes, monthly averages for many parameters and other derived fields.

4.5.1 Observation and assimilation of data

The number of data processed in ERA-Interim counts to approximately 10^7 per day. The majority of these data originate from the satellite observations, which includes clear-sky radiance measurements from polar-orbiting and geostationary sounders and imagers, atmospheric motion vectors derived from geostationary satellites, scatterometer wind data, and ozone retrievals from various satellite-borne sensors.

The in-situ measurement that is a conventional way of measuring, can be an obligatory constraint in the reanalysis.

In-situ measurement, is used as an obligatory constraint in the reanalysis process. This measurement is achieved for upper-air temperatures, wind component (u/v), and specific humidity from radiosondes, pilot balloons, aircraft, and wind profilers (Dee et al., 2011). Also, in-situ measurement of surface pressure, 2 m temperature, 2 m relative humidity and near-surface winds is done using ships, drifting buoys and land stations are integrated in steady numbers for reanalysis. (Dee et al., 2011).

The assimilation process consists of the two steps: Forecasting and reanalysis. (Dee et al., 2011). The forecasting model estimates the changing state of the global atmosphere and its underlying surface.

> This involves computing a variational analysis of the basic upper-air atmospheric fields (temperature, wind, humidity, ozone, surface pressure), followed by separate analyses of near-surface parameters (2 m temperature and 2 m humidity), soil moisture and soil temperature, snow, and ocean waves which are used to initialize a short term forecast to be used in the next cycle.
>
> (Dee et al., 2011).

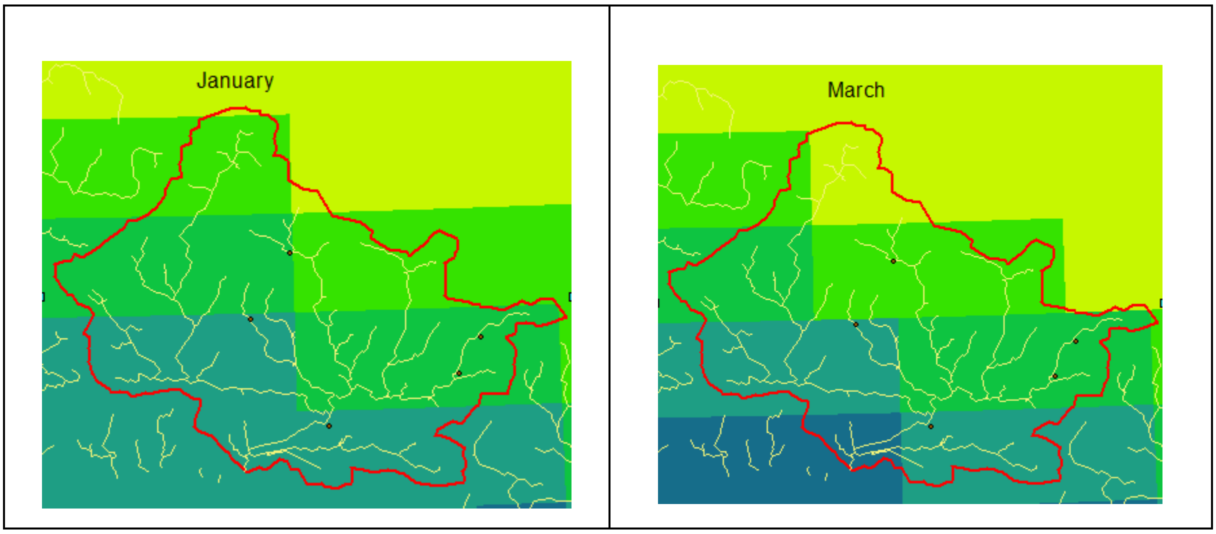
The atmospheric forecast model used for ERA-Interim has a 30 minute time step and a spectral T255 horizontal resolution, which corresponds to roughly uniform 79 km spacing for surface and other grid point fields. (Bumke et al., 2016).

Atmospheric reanalysis is then carried out for the forecasted period using the conventional observing system as a constraint in spite of their much lower data volume. During the assimilation of data, the volume of the data is very high. A quality control approach is applied for retreating this data volume and to exclude the expected negative impact on the reanalysis process.

4.5.2 Use of ERA-Interim data in study

Most of the meteorological stations are not located in the geologically challenging locations in Nepal. Even if they are located, there is no maintenance of these stations which create a question on the data quality. The spatial distribution of the hydro-meteorological stations is very poor, especially for the wind-speed and relative humidity whose stations are very few in number. This is a big concern for the research projects and development projects such as hydropower which require a good spatially distributed data for the analysis of snow melt, precipitation and other hydrological processes. Also, some of the rivers in Nepal originate from Tibet which again makes it very difficult to get the proper analysis result because of the difficulty of accessing the database from that side. The leeward and the windward side of the mountains in Nepal also have a very different hydro-meteorological conditions which are impossible to obtain from the existing meteorological stations in Nepal. Also, the quality of the station data from Department of Hydrology and Meteorology have always been a major concern for the already available data. Therefore, with the gridded data resource like ERA-Interim, that has the global database with already processed data, the extraction of the necessary data will be very easy. The success in use of this database in the current study would provide a new platform for generating the discharge from an ungauged catchment in the Himalayan basin like Narayani.

Spatial coverage of the ERA-Interim has 0.7 degrees approximate resolution incorporating 256 latitudes x 512 longitudes. The data is available in the grid ranging from 0.125x0.125 to 3x3 size (ESRL, 2014). This grid is the representation of the value of a particular element in a particular day (specified time resolution). Figure 4-12 is the illustration of the grid size 0.5x0.5 and parameter resolution for 2 meter temperature in °C in ERA-Interim.



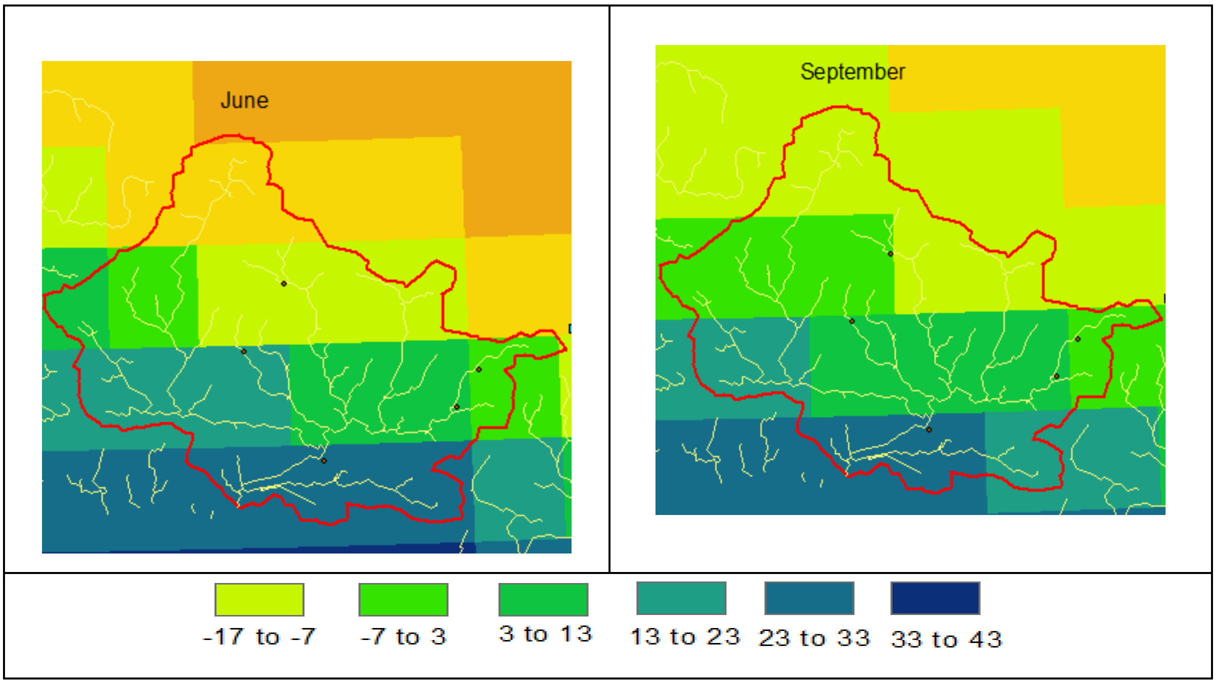


Figure 4-12: ERA-Interim Resolution for 2 meter temperature

The area attribute also has a range of options. The default range varies from 89.5S to 89.5N and 0E to 359.3E.

For this study purpose, geopotential (elevation), precipitation, temperature, global radiation, relative humidity and wind speed data are taken from ERA-Interim database. The time steps for these data is taken as 24 hours initiating at 00 UTC, but can vary depending on the complexity and precision requirement. These data are available at http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/. The date was chosen from 1st January 2005 to 31st December 2012 as the other hydro meteorological data. These data are downloaded as a NETCDF file for a global spatial coverage with the resolution of 0.5°x 0.5°. Since these data are pre-processed for the quality control, no quality control measures were applied in the process. The retrieved NetCDF files was directly read by the SHyFT repository.

5. Geographical data preparation

One of the most important steps in this study was preparation of the geographical data. This was achieved by using ArcGIS tool which will be thoroughly discussed in this chapter.

5.1 Introduction to ArcGIS

ArcGIS is a software program developed by Environmental Systems Research Institute (ESRI). It is a very strong tool for creating and analyzing the geospatial data by managing the geographic information in the database. ArcGIS is an integrated software that consist of three components: Arc Catalog, ArcMap and Arc Toolbox.

5.2 Point Networks

The geographical input to ArcGIS is the location of hydro-meteorological stations in x, y and z coordinates. The station's geo-spatial location collected from Department of Hydrology and Meteorology were changed to UTM coordinates and are added to the GIS system. Nepal lies in the UTM zone 44N and 45N. UTM zone 44N is suitable for the coordinates between 78°E and 84°E. The Narayani Basin lies mostly within UTM zone 44N, therefore the points are projected in WGS 84/ UTM zone 44N coordinate system.

5.3 Watershed Delineation

Watershed of a point is region upstream of it draining into that point as a river, river system or a body of water. In order to manage landscapes as watersheds, it is necessary to delineate the boundaries of it so that watershed can be used as a common spatial terminology. It is very important feature in hydrological modelling. Historically, watershed delineation was carried out manually using the topographical map. For this study purpose, Archydro, an extension of ArcGIS will be used along with Digital Elevation Model (DEM).

Listed below are the requirements needed for watershed delineation using ArcGIS

5.3.1 Digital Elevation Model

DEM is the main dataset required for watershed delineation. For any study area, this dataset can be downloaded from Hydro 1K. The resolution depends on the purpose of study. For this study purpose, the continent based resolution of 1kmx1km was used from Hydro 1K for Asia region which was later clipped for the Narayani Basin Catchment area using the shape file. More information about data extraction can be found on https://lta.cr.usgs.gov/HYDRO1K.

5.3.2 Stream Networks

Stream networks are shape file based supportive data for GIS. For Asia region, the shape file for stream networks was downloaded from http://hydrosheds.cr.usgs.gov/index.php.

5.3.3 Outlet points

The geo-spatial location of gauging stations collected from Department of Hydrology and Meteorology were changed to UTM coordinates and are added to the GIS system.

5.4 Procedure

At first, it is necessary to project the data set to the same spatial reference system. For this study, the dataset were projected to WGS 84/ UTM zone 44N. This can be done using,

- Arc Tolbox >> Data Management Tools >> Projections and Transformation

Then the continental based datasets both DEM and Stream Networks was clipped for the required area. First the clipping was carried out for Nepal as a whole. The raster data was clipped using,

- Arc Toolbox >> Data Management Tools >> Raster >> Raster Processing >> Clip

And the vector data or features were simply clipped using

- Geoprocessing >> Clip

The end result of this was the DEM and the stream network projected for Nepal.
The sequence involved in carrying out the catchment delineation using the hydrology tool from Arc tool box is listed below.

- DEM Manipulation
- Fill Sinks
- Flow Direction
- Flow Accumulation
- Create Watershed Pour Points
- Snap Pour Point
- Watershed

Each of the above mentioned functions is briefly described below.

Fill Sinks

This function fills sinks or gaps in a grid and removes small imperfections in DEM.

- *Terrain Preprocessing >> DEM Reconditioning >> Fill Sinks*

Flow Direction

This function computes the flow direction of a given grid indicating the direction of steepest descent from that cell.

- *ArcToolbox >> Spatial Analyst Tool >> Hydrology >> Flow Direction*

Flow Accumulation

This function computed the flow accumulation grid that contains the accumulated numbers of cells upstream of a cell, for each cell in the input grid.

- *ArcToolbox >> Spatial Analyst Tool >> Hydrology >> Flow Accumulation*

Create Watershed Pour Points

This represents the outlet point of the catchment. This can be done by creating an empty shapefile in the working environment and defining the outlet point with proper coordinates in thus created shapefile.

- *ArcCatalog >> New >> Shapefile*

Snap Pour Point

This function snaps pour to the cell of highest flow accumulation within a specified distance.

- *ArcToolbox >> Spatial Analyst Tool >> Hydrology >> Snap Pour Point*

Watershed

This function determines the contributing area above a set of cells in a raster.

- *ArcToolbox >> Spatial Analyst Tool >> Hydrology >> Snap Pour Point*

Watershed will depend on the number of pour points, for 5 pour points in our study, there are five watersheds. They are, Kali Gandaki, Seti Gandaki, Marshyangdi, Trishuli and East-Rapti.

5.5 Land Use

For the land use the raster was obtained from USGS. The vector data had 22 classes. For SHyFT input, the land use has to be classified to Reservoir fraction, Lake Fraction, Glacier Fraction and Forest Fraction. For this purpose **Reclassify** function was used. Since there was no reservoir in the study area, the reservoir fraction was ruled out of the dataset and was simply assigned a value of 0. Details of Land use reclassification from USGS to SHyFT is presented in Table 5-1 below.

Table 5-1: Land use reclassification

USGS Class	**SHyFt**
16, 17, 18	Lake Fraction
19	Glacier Fraction
11, 12, 13, 14, 15, 20, 21, 22	Forest Fraction

After reclassifying the catchment the raster was changed to polygon and then following procedure was carried out.

1. Grid for the catchment was created using Data management tool>>fishnet. The cell size was set to 2000x2000 and geometry type was selected as polygon.
2. The grid created in 1 was clipped using reclassified polygon created earlier for the catchment. For each of the classification, this process was carried out separately. e.g, forest fraction following lake fraction following glacier fraction
3. From the attribute table of the joins and relates following processes were completed.
    a. On what do you want to join, join data from another layer based on spatial location was selected.
    b. On choose layer to join this layer, the polygon converted from clipping the DEM was selected and then average, minimum and maximum options were checked.
4. This process set the grids with the reference elevation.
5. In the attribute table a field was added by the parameter desired and area, x-coordinate and y-coordinate was calculated.
6. Next, bounding box was checked from the attribute table generated from step 2, the area outside the bounding box was deleted.
7. In the attribute table a field was added by the parameter desired and area was calculated.
8. Using Analysis>>spatial join, step 5 and 7 were joined the desired join space was set as step 5.
9. The area fraction was calculated using the attribute table and the results were exported as a text file.

5.6 Input format in SHyFT model system

SHyFt has its typical data format to take in the GIS data. After the 9th step is completed, the file was compiled with forest fraction, lake Fraction and glacier fraction and converted to text file with the name cell_data.txt. The conversion process following this step is mentioned in chapter 6. The input format in SHyFT model is listed in Appendix 3.

6. SHyFT

This chapter will focus on the modelling tool SHyFT used during the study.

6.1 Introduction

Statkraft Hydrological Forecasting Toolbox (SHyFT) is an integrated hydrological modelling tool developed by Statkraft. It is an open source software which is optimized for highly efficient modelling of hydrological processes by following the model of distributed and lumped models. Its recent developments has introduced more physically based process-level methods.

6.1.1 Physically based distributed model

Models may be classified as distributed when they utilize spatial distributions of input variables and spatial distributions of physical properties in a catchment in combination with algorithms that calculate spatially distributed hydrologic behavior

Such models attempt to increase the simulation accuracy by explicitly accounting the areal distribution of spatially non-uniform quantities and processes. This increased simulation accuracy comes at the expense of significant increase in computational effort and data preparation requirements (Rinde, 1996). These models deal with the distributed parameters, inputs and initial conditions. The brief summary of this is given in Table 6-1 below.

Table 6-1: Distributed model explanation

Principle distributed parameters		
Parameters	**Direct influence**	**Indirect influence**
Elevation	Temperature, Precipitation, wind exposure	Snow accumulation/melting, potential evapotranspiration
Soil storage capacity	Soil water retention, runoff generation, actual evapotranspiration	Flood response
Land use	Interception capacity, infiltration capacity	Snow distribution/melting, actual evapotranspiration
Aspect	Incoming radiation	Evapotranspiration, snow melting
Principle distributed inputs		
Precipitation, Temperature, Wind Field (Speed, Direction), Solar radiation (Net Shortwave radiation, Net longwave radiation), Air humidity.		
Principle distributed initial conditions		
Initial soil wetness, Initial snow status, Seasonal vegetation status.		

6.1.2 Principles of physically based distributed models

Due to the inadequacy of the lumped model to represent the spatially distributed parameters, it was necessary to introduce the complex physically based distributed models. The underlying principle of this approach is stated below (Rinde, 2015).

- All processes are explicitly described in each grid-cell, starting from an initial state and limited by boundary conditions defined by the hydrological states in their neighboring cells.
- Spatial variability and topographical effects (elevation, slope and aspect) cause all cells to have different climatic inputs.
- Physical properties vary between cells (topography, vegetation, surface permeability, soil conductivity, etc.).
- Sub-cell variability is ignored.
- Water fluxes are described by partial differential equations embedded in iterative finite difference or finite element computational schemes.

These model principle are further illustrated in Figure 6-1.

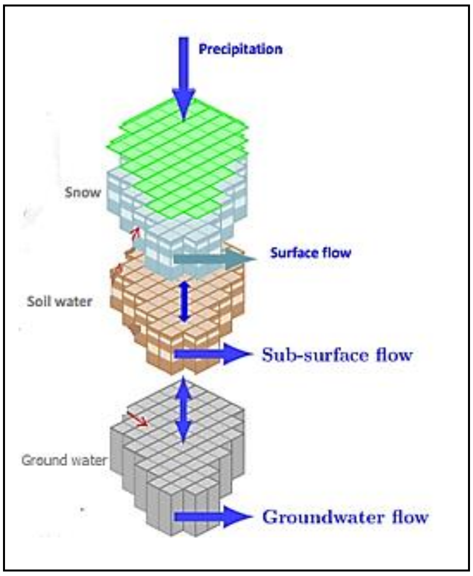


Figure 6-1: Spatial representation of the system

By using these principles, it was easier to produce the distributed hydrological responses, which further helps to improve the ability to model in the ungauged catchments.

6.2 Regional modelling

As mentioned before, one of the major challenges in hydrology is to determine the flow through an ungauged catchment. This can be made possible by a regional hydrological model, which allows the prediction of dynamics of hydrology, water balance and hydrological parameters at the ungauged basin. Regional modelling tool implies the repetitive use of the model everywhere within the region using the global set of parameters. The calibration process also attempts to make a relationship between model parameters and modelling units so that the generated calibrated parameters can be applied to the ungauged area.

One of the regional modelling tool, ENKI uses the similar principle. The model is executed for multiple catchments and iterated for possible range of several free parameters. The calibrated parameters are judged by Nash-Sutcliff efficiency and transferred through basins to predict the runoff. (Shrestha et al. ,2012)

ENKI is already a very strong tool for the regional modelling but due to higher flexibility in adding new simulation routines, it was useful for research purpose but not so much for the operational hydrology which requires running a large number of models for the prognosis. Hence, there was need for a tool that runs faster and calibrates faster, the result was the development of SHyFT.

The main reason for the development of SHyFT was to increase the execution speed of the hydrological process. For this, some flexibility features in ENKI was compromised.

6.3 SHyFT in hydrological modelling

SHyFT is an open source tool, which makes it accessible by everybody. The tool works on a python environment, the code for which is based on early initiative for distributed hydrological simulation, ENKI which was developed at Sintef.

SHyFT can work for both distributed and lumped models. It is also a strong modelling tool to work with the regional modelling. Like ENKI, SHyFT also has a flexibility in adding new simulation routines through its 'orchestration' system which is described later. It works with the YAML setup which is a bit more complex to extend with the new hydrological routines, but it also has an advantage of controlling over the output through its Python interface. Since, SHyFT is a new tool, there are some requirements to compile and run this modelling tools which is described in appendix 3. SHyFT also has a unique model system which is described in below.

6.4 Model Components

SHyFT is a complex system of routines, but these routines are arranged in a manner that is understandable to the first time users. Once you open the SHyFT folder, the folders are arranged in same order as Figure 6-2.

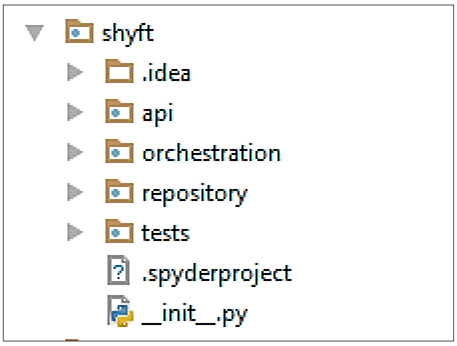


Figure 6-2: SHyFT folder arrangement

A brief description of each sub-folder is mentioned in below.

6.4.1 api

This directory contains the python wrappers for the shyft core. A core contains basic data structure, cell-models, models and algorithms. The api structure is very simple as seen from Figure 6-3.

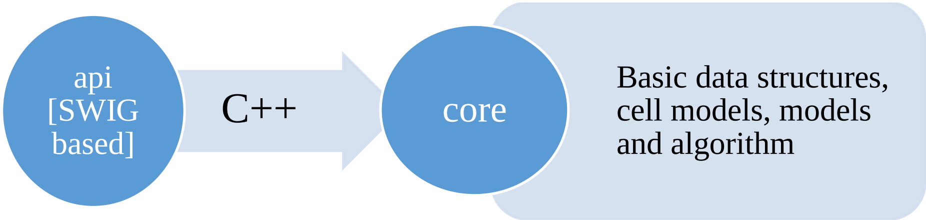


Figure 6-3: api structure

6.4.2 orchestration

The process involving the ingestion of data coming from the observations (meteorological stations, catchment discharges etc.) to perform a simulation is called orchestration in SHyFT. Prior to ingestion, there might be a round of calibration to fill up the internal data structure of SHyFT.

The orchestration code uses YAML configuration files in order to define a calibration or simulation run and provides basic infrastructure to read them, still allowing the users to add their own code to tailor the calibration/simulation.

Although the core of the SHyFT simulation code is written in C++, all the orchestration code is written in Python and uses the Python-API. In addition, the orchestration code adds another abstraction layer that allows the end user to write their own repositories classes with a minimal (or none) knowledge of SHyFT core internals.

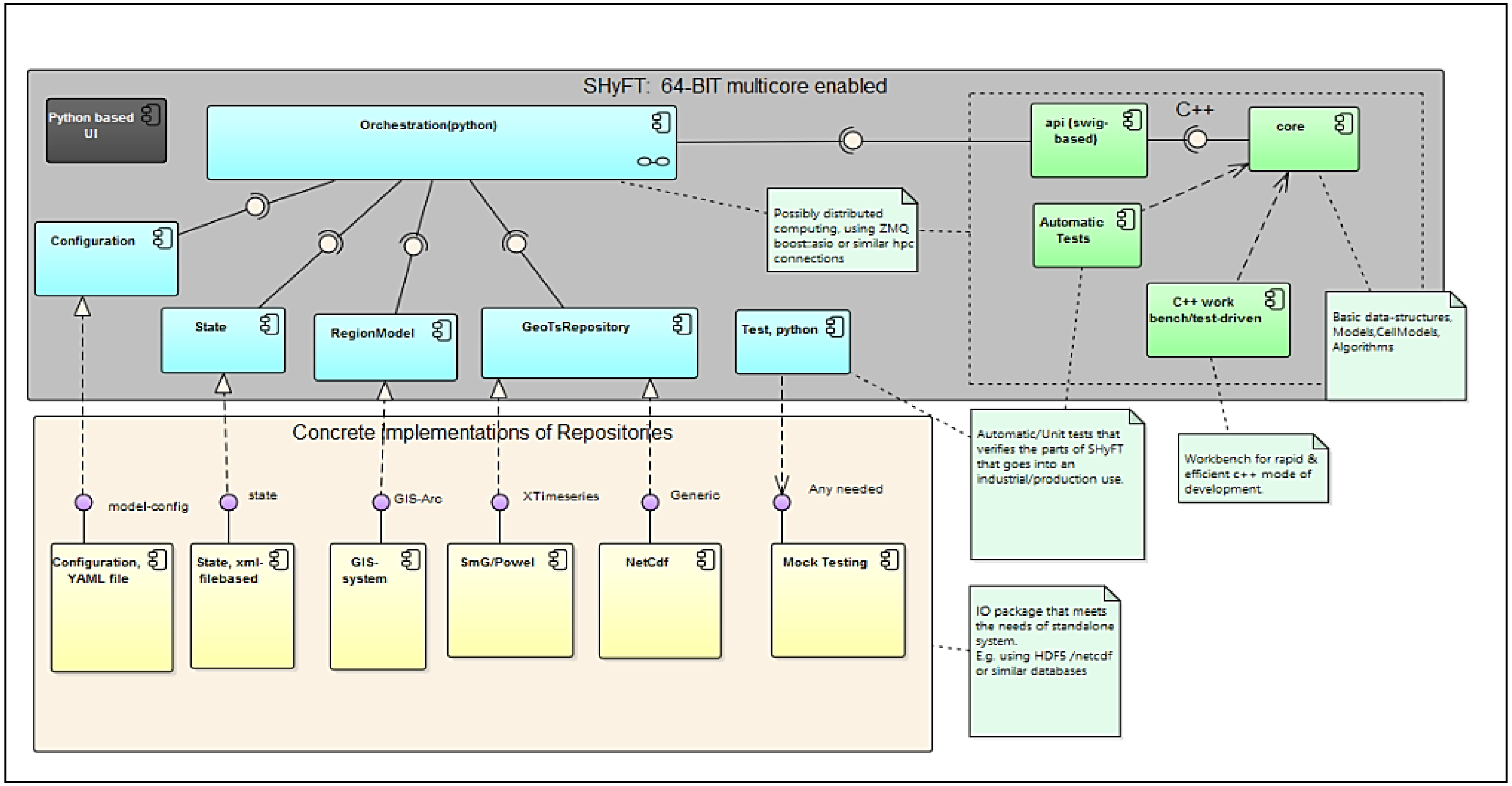


Figure 6-4: SHyFT model diagram

6.4.3 repository

In general, a repository is the collection of data in some arbitrary format, but in SHyFT it refers to the Python code that can read this data and feed it to the SHyFT core.

The ‘shyft.repository’ package offers interfaces so that the users can provide concrete implementations of their own repositories. In this regard, users can provide their own repositories customized to their own configuration files (typically slight variations of the YAML config files that come with SHyFT) and a diversity of data formats (maybe a whole collection of files, databases, etc.). eg. ERA-Interim repository.

Configuration files with orchestration and repositories further described below.

1. The region configuration

It maps to the RegionModelRepository in the Figure 6-4 and is responsible for providing the static geographic properties of the region (or pointers/references to find those).

2. The model configuration

It maps to the cell-model parameters (green objects in the C++ section) in Figure 6-4. This section also maps to the cell model parameters along with the model type. e.g. pt_gs_k

3. The datasets

It corresponds to the GeoTsRepository in Figure 6-4 and is responsible to provide geo-located time series when the orchestrator ask for time-series within a specified bounding box and time-period.

4. The interpolation configuration

Describes how the geo-located time-series, as fetched from the GeoTsRepository should be projected onto the area of each cell.

6.4.4 tests

This folder verifies and contains all parts of SHyFT that goes into the industrial production. So, this folder has the information of all the routines involving equations and methods which makes an integral part SHyFT operation. e.g. actual_evapotranspiration_test, hbv_snow_test etc. The netcdf files inside tests contains all the YAML files which is the key to SHyFT functioning. These are the configuration files. They are six in number connecting the input data files (physiographic, climate and calibration in NetCDF format) to the simulation and calibration process files.

The NetCDF folder inside SHyFT tests contains following major configuration files which are in YAML file format.

1. Region.yaml

This YAML file connects to the physiographic data which is stored at the cell_data in the netcdf format inside shyft-data folder.

2. Dataset.yaml

This YAML file is responsible to return the climate data stored as precipitation, temperature, wind_speed, relative_humidity and global_radiation as a netcdf file inside shyft-data folder.

3. Model.yaml

This YAML file contains model parameter set for the simulation and calibration purpose.

4. Simulation.yaml

This file calls and returns the region.yaml, dataset.yaml and model.yaml files with user defined function to specify the start time, time step and number of steps to be used in the simulation and calibration process.

5. Calibration.yaml

This file uses the simulation.yaml as the simulator using one of the three optimizers (min bobyqa, sceua and dream) to calibrate the discharge using the input file (discharge.nc) stored inside shyft-data folder.

6. Interpolation.yaml

This file returns the interpolation algorithm in the simulation.yaml file during the simulation.

6.5 SHyFT Model structure

Since SHyFT is a distributed hydrological model, it works down from the regional level to the cell level by distributing the input parameters (hydro-meteorological data) in to the individual geo-referenced cells. Figure 6-5 below gives the overview of how this process completes in SHyFT.

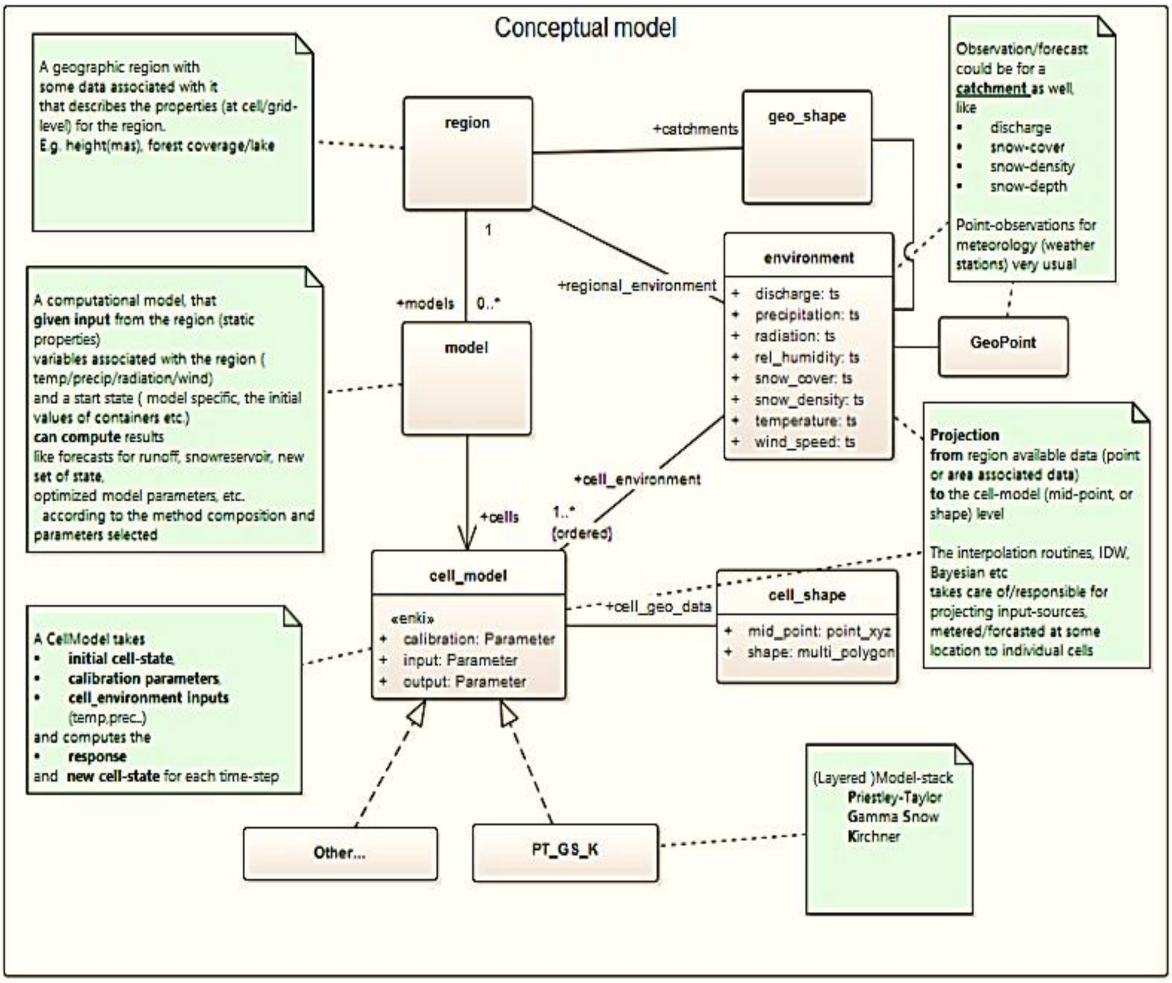


Figure 6-5: SHyFT model structure

The model file is one of the most important part of SHyFT. This contains two sections as seen in Figure 6-6.

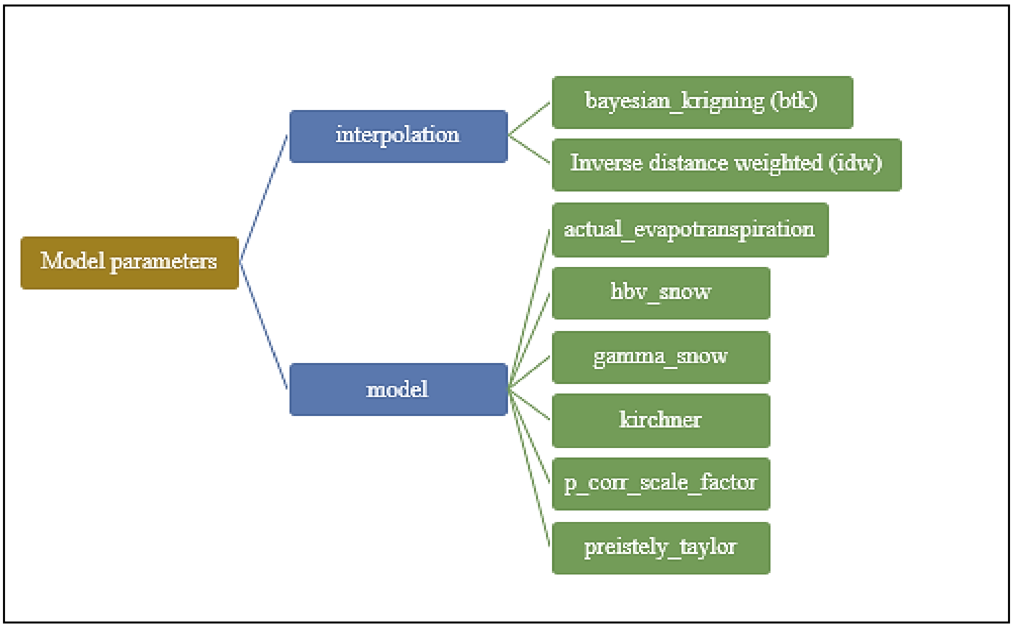


Figure 6-6: SHyFT model functions

The interpolation parameter uses following two approaches

Bayesian Kriging interpolation

This approach is used basically for the temperature data. In this process the interpolation is carried out by scattering a time series of data from the given source location at the target location using a Bayesian Temperature Kriging method where the vertical distances used during the semivariogram's covariance calculations are scaled with a height factor using the similar semi-variogram graph of Figure 6-7. .

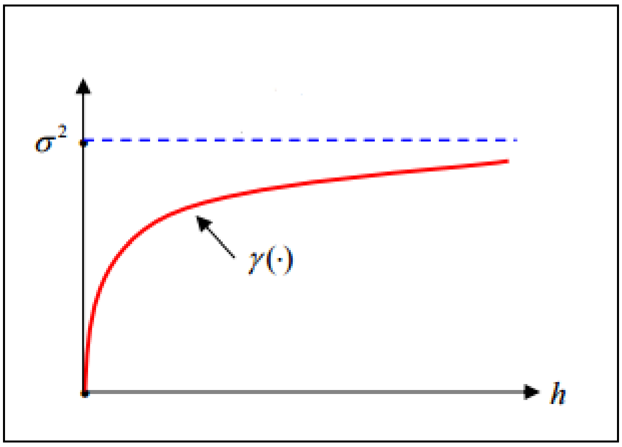


Figure 6-7: Semi-variogram for Bayesian Krigning interpolation (Smith and LeSage, 2004)

In addition to interpolating values at the destinations, this interpolation also estimates the sea level temperature and temperature gradients at each time step.

Inverse distance weighted interpolation

This approach is used for the interpolation of all the input data except temperature. In this interpolation, the time series data at the given source location is transformed to a new set of time series data at the grid location. In doing so, a weighted average of maximum number closest neighbors with available data is used for each target location. The algorithm does not use neighbor locations that exceeds the maximum distance. Also, the interpolation is limited by the maximum numbers of neighboring members. The algorithm also uses gradient correction for the precipitation data.

The IDW method used in SHyFT is similar to the idwtemp used in ENKI by Sjur Kolberg/Sintef.

The model stack of the model parameter uses the following functions.

Actual_evapotranspiration

Actual evapotranspiration is calculated on the basis of the potential evapotranspiration calculated from Preistley taylor. Besides that, this function also uses water level, snow fraction, scale factor and fraction of time in the algorithm. This algorithm is derived from ENKI programmed by Kolbjørn Engeland and Sjur Kolberg.

$$\text{Actual evapotranspiration} = \text{pot. evapotranspiration } (1 - e^{-\frac{water\ level}{scale}})(1 - \frac{snow\_fraction}{dh})$$

Where,

dh = (dt)/hour (or any calendar time)

scale = scale factor*dh/3.0

scale factor = 1.5

HBV snow

The snow melt and accumulation is calculated using the HBV model concept as seen in Figure 6-8 which based on the input precipitation and air temperature. The parameter used for the HBV routine is listed in appendix.

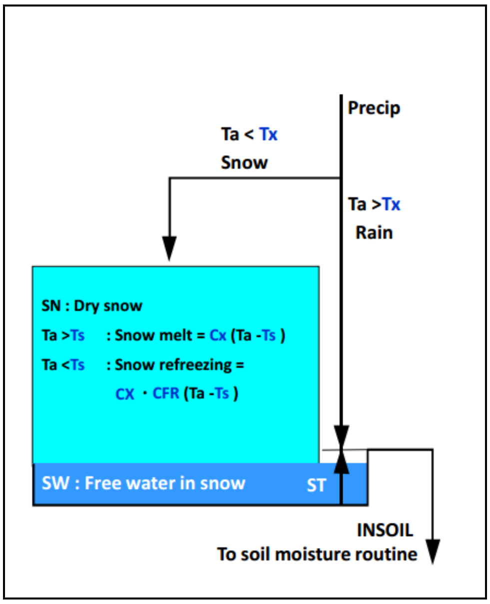


Figure 6-8: HBV snow melt model (Rinde, 2015)

Gamma snow

Gamma snow is used in SHyFT as an alternative to the HBV snow model. Snow-rain threshold temperature and snow melt sensitivity to wind speed are free parameters in this model. The outflow is the melt release from the saturated snow based on the Gamma distribution snow depletion curve (SDC) in Figure 6-9.

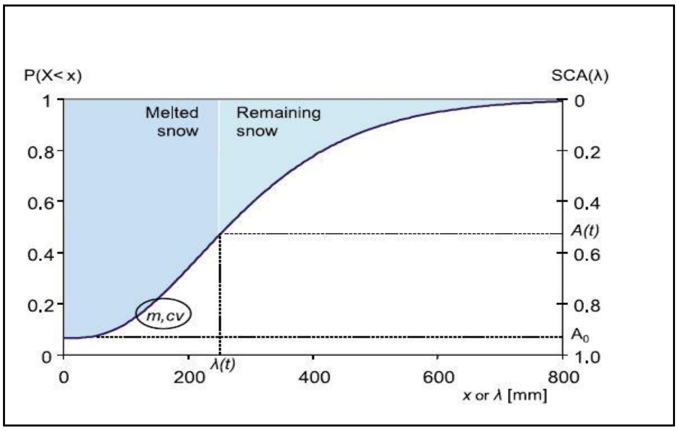


Figure 6-9: Snow Depletion Curve (SDC) (Kolberg and Gottschalk, 2010)

This curve in Figure 6-9 is gives a three parameter model. Amongst the three, two parameters quantify the snow pack by a Gamma distribution and the third parameter quantifies the maximum snow covered area in a cell during the snow melt. These two parameters are given by the following equation,

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

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

In the equation, p is the probability density function, F is the cumulative probability distribution function. F is equal to the incomplete Gamma function (γ) with shape and scale arguments. A (t) is the snow covered area of the at time t (Kolberg and Gottschalk, 2010). Other variables define the snow pack in each cell as,

m = average snow water equivalent [mm]

cv = coefficient of variation of snow water equivalent

A_0 = snow covered area at the beginning of melt season

λ(t) = accumulated melt depth since the melt season

There are three classes set to achieve the snow melt in the model which is given in the appendix.

Kirchner

Kirchner uses three parameters to calculate the discharge on the basis of observed precipitation and evapotranspiration data. This function achieves the soil moisture routine and runoff generation in the model.

$$\ln(g(Q)) = \ln\left(\frac{-\,dQ/dt}{Q}\bigg|_{P \ll Q, E \ll Q}\right) \approx c_1 + c_2 \ln(Q) + c_3 (\ln(Q))^2$$

Where,

c1: first parameter in the Kirchner model

c2: second parameter in the Kirchner model

c3: third parameter in the Kirchner model

g(q): Sensitivity function, this function expresses the sensitivity of discharge to changes in the storage.

The minimum q value is set at 0.0001 mm of water level for any state. If the supplied q state is less than this value, the algorithm forces the new q state to the minimum value to ensure stability.

There are some assumptions for the model algorithm for Kirchner in SHyFT. They are,

- The outflow, Q depends entirely on the amount of water stored (S) in the catchment.

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

- The flow in the catchment is controlled by the release of water from the storage rather than bypassing flow from direct precipitation.
- The saturated and unsaturated storages are hydraulically connected and the net ground water flow across watershed boundary is zero (Kirchner, 2009)

P_corr_scale_factor

This function corrects the accumulation of precipitation for a particular region under calibration. The initial scale factor is set at 1 but changes during calibration. The algorithm used in SHyFT is adapted from ENKI method programmed by Kolbjørn Engeland and Sjur Kolberg.

Precipitation correction = precipitation x scale factor

Presitley Taylor

Preistley Taylor method is used to calculate the potential evapotranspiration in SHyFT. The algorithm uses the Preistley Taylor equation as mentioned below.

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

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. (Priestley and Taylor, 1972)

6.6 Use of SHyFT system and model in this study

In this study, all the systems and models of SHyFT were carefully used to generate the runoff from the given input system. The interpolation was carried out using IDW interpolation for precipitation, wind speed, radiation and relative humidity and BTK for interpolation of temperature in interpolation.yaml. In this study, actual evapotranspiration, Preistley Taylor, Gamma Snow and Kirchner algorithm were used from the model for generating the outflow. The method stack used for this study is given by the name pt_gs_k. The flow of the processes in this method is given in Figure 6-10 below.

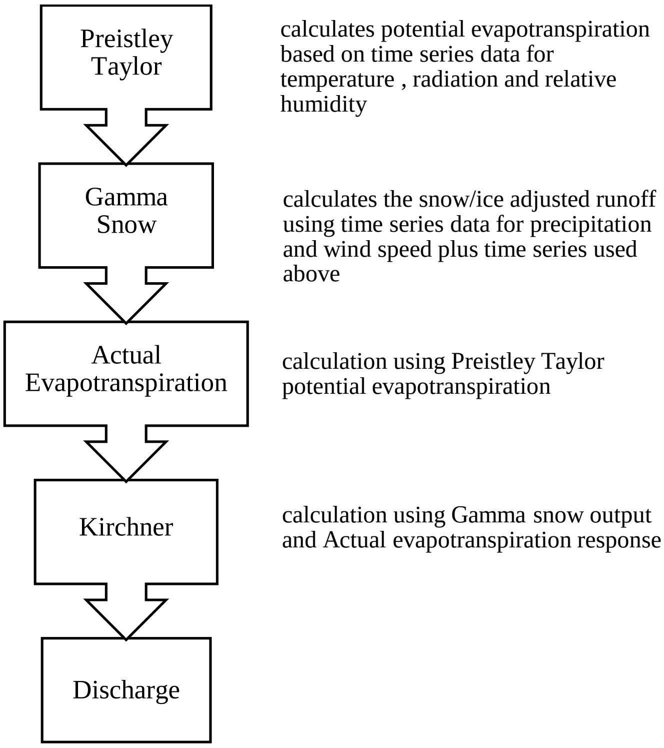


Figure 6-10: pt_gs_k method stack

The range of these parameters was set at calibration.yaml file. SCE-UA was primarily used as the optimizer and the fitness in the model was checked using NSE which was again mentioned in calibration.yaml.

The setup of the model was done for two different inputs. In the first case station data were used in the model and 'shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepositoryrepository' was used to read the *.nc files from the model.yaml file. In the second case however, ERA-Interim dataset was used and 'shyft.repository.netcdf.erainterim_data_repository.ERAInterimDataRepository' was used to read the *.nc files for ERA in the same model.yaml file.

Since regional modelling was carried out in the study, different outlet stations with their catchment id and uid (station name as in discharge.nc) were mentioned in the calibration.yaml file and the model was simulated for all those stations. The catchment indices were given the same catchment id as the outlet stations for n number of stations used in region.yaml file before the simulation. The bounding box of the catchment limitation was set in model.yaml file. For this the lowest latitude and longitude value in UTM were written for lower_left_x and lower_left_y and the number of steps to obtain the maximum co-ordinate of the bounding box was set at nx and ny, step_x and step_y were kept the same as the grid size of the cell data.

Tests were run using run_shyft.py and calibration was carried out using calib_shyft.py. Before carrying out these tests, the number of years and the time step was set at simulation.yaml file.

The output was extracted using sim_obs.py which was a self-written python routine. Achieving the output in excel further required packages xlsxwriter and pandas in addition to the already installed python packages. The codes and routines used under tests are in appendix 2.

Figure 6-11 gives the detail breakdown of the features used during the study in SHyFT.

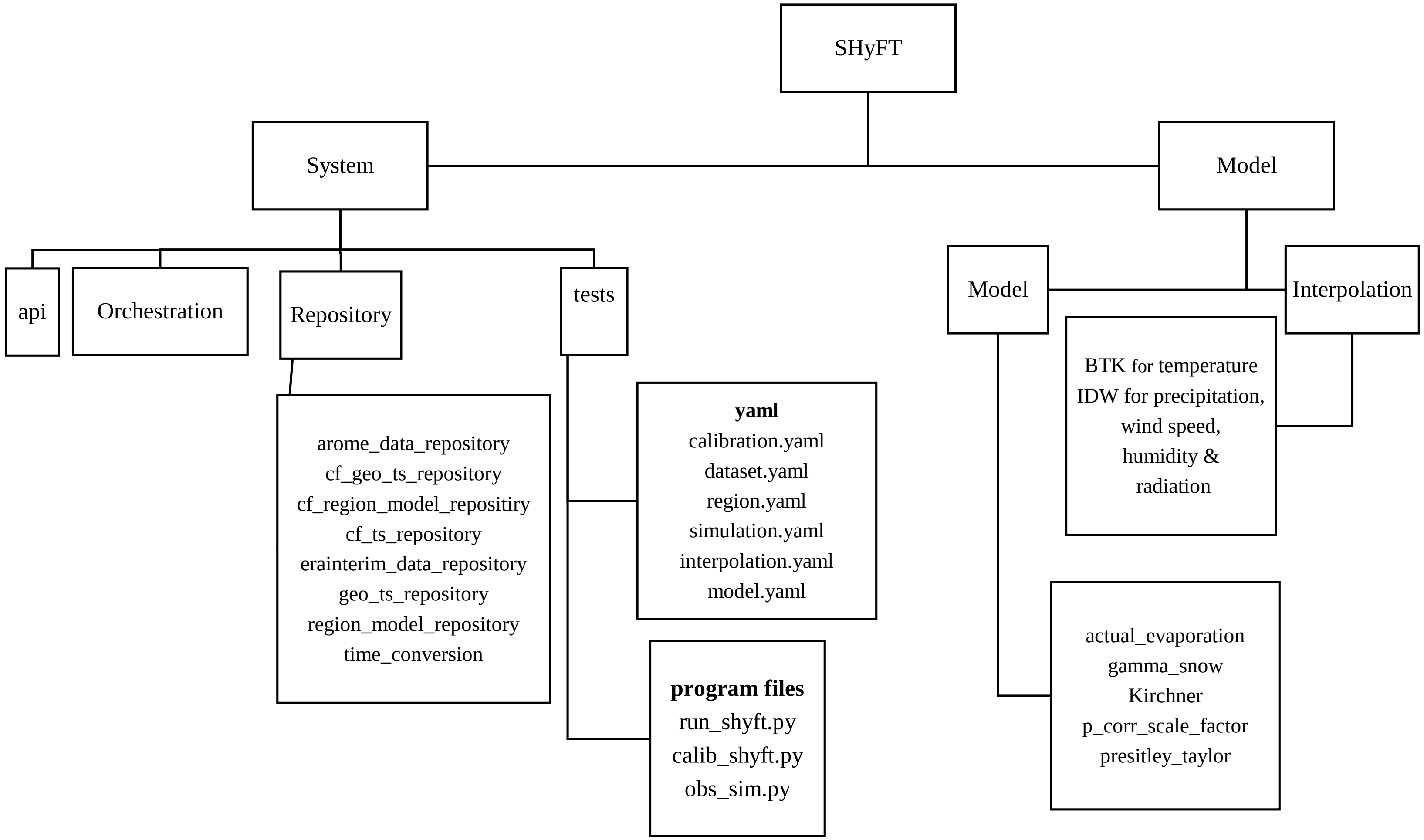


Figure 6-11: SHyFT model as used in the study

6.7 Data preparation for SHyFT

Like any other hydrological modelling tool, there is a unique style for data input in SHyFT system. SHyFT takes in the data in netcdf format. So the processed data in excel has to be saved as a text file in a folder with precipitation, temperature, relative_humidity, wind_speed, glonal_radation, discharge and cell_data. Thus converted text files are further converted to NetCDF file using a python routine as in appendix 2. This process is simplified in Figure 6-12. The final product from the conversion can be used as the input file in SHyFT.

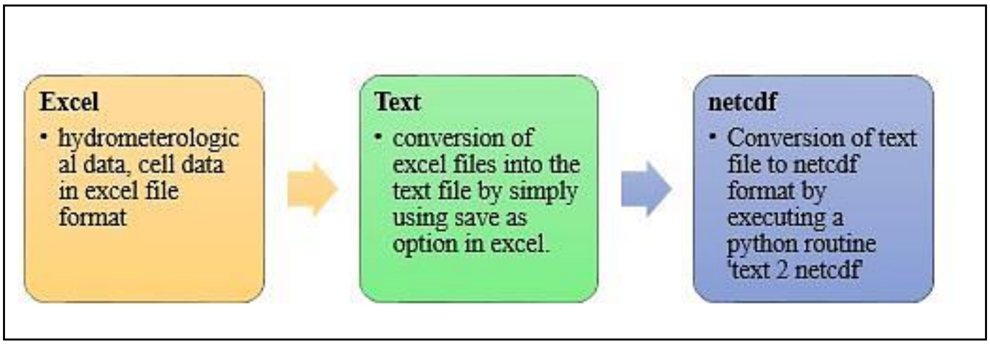


Figure 6-12: Data conversion process in SHyFT

This study used the YAML files to read the input data in the form of NETCDF files. These files were prepared by fulfilling the requirements as mentioned above. There were two sets of NetCDF files used during the study, which is given in Table 6-2.

Table 6-2: NetCDF file used during study

Dataset	Set I Hydro-meteorological station	Set II ERA-Interim dataset
Precipitation	precipitation.nc	tp.nc
Temperature	temperature.nc	t2m.nc
Relative humidity	relative_humidity.nc	d2m.nc, sp.nc
Wind speed	wind_speed.nc	u10.nc, v10.nc
Global Radiation	radiation.nc	ssrd.nc
Discharge	discharge.nc	
Physiographic data	cell_data.nc	

6.7.1 File format for ShyFT

Appendix 3 gives the overview of the file format in Excel. It should be remembered that while converting the files into text files, there should be no space after the last text in a line. The conversion routine changes the text data in a one dimensional array. Therefore, the presence of space will indicate the end of text file and the routine fails to read anything beside that point.

The EPSG in the format refers to the coordinate system for that catchment/study area. This value is 36232 for 32N projection. The missing values shall be left as -999, SHyFT has a buildup system that fills up these missing data in each cell by different interpolation methods discussed earlier. The precipitation units is taken as mm/hr, therefore the data has to be adjusted likewise. e.g. If the precipitation data is available in mm/day, then each data has to be divided by 24 to get them in mm/hr. The coordinate system follows the UTM coordinate format and the catchment id refers to the unique number assigned for each catchment by the user.

6.8 Handling ERA-Interim data in SHyFT

SHyFT has a repository built for the ERA-Interim datasets. This repository reads all the global netcdf files and prepares them for the model region by clipping the latitude and longitude boundaries from the global dataset.

However, the input file has to be in a cluster. e.g if it is a wind speed data then the ERA-Interim netcdf file should have the information regarding the time axis, u10, v10, latitude, longitude and elevation in a single file. Sometimes this requires merging of two netcdf files, this process can be carried out using python. The ERA-Interim datasets required to process any file is well mentioned in the repository as,

Precipitation = Total precipitation [tp], geopotential value [z], latitude, longitude, time

Temperature = 2 meter temperature [t2m], z, latitude, longitude, time

Wind speed = x-wind [u10], y-wind [v10], z, latitude, longitude, time

Radiation = Surface solar radiation downwards [ssrd], z, latitude, longitude, time

Relative Humidity = dew point temperature [d2m], surface pressure [sp], z, latitude, longitude, time

These YAML files have to be connected with this repository to run the simulation and calibration. Any change in the file location has to be carried out in dataset.yaml and simulation.yaml.

7. Calibration

The calibration processes used in the study will be discussed in this chapter.

7.1 Introduction

Model calibration is the process of adjusting the model parameters to fit within the margins of uncertainties in order to get the actual system response. In hydrology is usually carried out by comparing the modelled and observed hydrographs for a number of locations and a small number of input parameters that are varied in a trial and error mode to achieve a desired response (Gupta, 1994). Usually, the end result of the calibration process is the suitable set of the parameter values which are difficult to measure directly from the field data.
Principally calibration can be carried out by three different ways.

1. Trial and Error method or manual parameter optimization
2. Automatic method or numerical parameter optimization
3. Combination of the above 2 methods.

7.2 Evaluation of the calibration

Evaluation of the calibration results is a vital process in hydrological modelling. There are two methods primarily used to evaluate the model performance; qualitative and quantitative.
The qualitative method includes the visual inspection of the simulated and observed hydrographs. For good calibration the simulated and the observed hydrographs should produce the correct flow variation pattern and recession curvatures. It should also be able to produce the similar correct overall volume for simulated and observed discharge. This method is highly subjective because of the lack of the numerical function, therefore a quantitative approach should also be examined.
A quantitative method is an objective function to determine the model performance in simulation. An objective function is a numerical function that is used to check the goodness of fit between the calibrated or modelled results.
In this study, the modelled result was checked using the following quantitative methods.

- Explain variance or Nash Sutcliffe efficiency, R^2

Nash Sutcliffe efficiency is defined by R^{2-}which is given as the equation below.

$$R^2 = 1 - + \frac{\sum(Qs - Qo)^2}{\sum(Qo - Qo_m)^2}$$

Where,

Qo = Observed Runoff

Qs = Simulated Runoff

Qo_m= Mean of observed runoff

R^2 value varies between -∞ to +1, where 1 indicates a perfect fit, 0 indicates that the model prediction are as good as the mean of averaged data and negative values indicates that the observed mean is better than the model value.

- Accumulated Difference

The accumulated difference is the difference between the sum of the simulated and observed discharge for each value setup. It is also defined by an objective function which is given by the equation below.

$$\text{AccDiff} = \sum_{t=0}^{n}(Qs - Qo)$$

Where,

Qs = Simulated discharge, m^3/s

Qo = Observed discharge, m^3/s

For a good calibration result, the accumulated difference should be close to zero, which is very difficult to obtain. Therefore this criteria can be very useful in comparing between the two case results rather than determining the goodness of fit.

- Water balance

The water balance is the balance between the input and the output of the hydrological system. It is given as,

$$Q = P - Et \pm \Delta S$$

Where,

Q = Observed discharge [mm]

P = Precipitation [mm]

Et = Actual Evaporation [mm]

ΔS = change in the soil moisture storage, or snow storage [mm]

In this approach, the balance in the equation will indicate a proper model calibration.

7.3 The calibration process using SHyFT

In SHyFT there are two methods of achieving the calibration process; manual and automatic calibration. This process of calibration is completed in two major YAML files calibration.yaml and model.yaml. The calibration.yaml file has the option of setting up the optimization procedure through one following three optimization algorithms.

- SCE-UA optimizer

  This optimizer has the advantages of globally converging the parameters even during the presence of multiple region of attraction. Apart from that it is also very robust and has the ability to handle the high parameter dimensionality (Duan et al., 1994).
- MIN-BOBYQA

  MIN-BOBYQA is a local search algorithm that converges at the response surface using a quadratic approximation for unconstrained upper and lower bounds where the function value improves. Therefore, it can handle a lot of calibration parameter but it comes at the cost of never knowing whether a global optimizer is found or not.
- DREAM

  It is an adoptive metropolis sampler and can be best used for the likelihood based PMs.(Shrestha et al., 2012)

For this study purpose SCE-UA optimizer was used dominatingly, but in some cases MIN-BOBYQA was also used to check the calibration.

Apart from the optimizers, the calibration.yaml file also provides the range of calibration parameters.

The model.yaml file contains the model parameters that are changed during the calibration process. This is where the parameter change in each step gets stored. The final calibration parameter is applied in model.yaml and can be obtained by using a simple python routine. In the calibration process, this was added in the calib_shyft.py file inside shyft>>shyft>>tests.

7.4 Calibration cases

For this study, three calibration cases were evaluated. For all the cases the data from 2005-2009 was used for calibration and 2010-2012 for validation. The model calibration was carried for Gamma Snow algorithm for snow routine along with Kirchner's for soil moisture routine and Preistley Taylor for evapotranspiration routine as mentioned in the previous chapter. Gamma Snow was chosen over the HBV snow routine because of the earlier test results the HBV snow model in the Himalayan catchments generated some accumulated value which was not satisfactory for snow outflow. (Timalsina, 2009)

Case I:

The model was calibrated for the 4 major tributaries; Kali Gandaki, Trishuli, Seti Gandaki and Marshyangdi as seen in Figure 7-1. The choice of these tributaries were based on their unique regional variation with the rainfall pattern, catchment characteristics, catchment sizes and the distribution of their area from west most extreme of Narayani basin to the east most extreme. In this case, the meteorological data (precipitation and temperature) were used from the meteorological stations in Nepal. Wind speed, radiation and relative humidity was taken for Pokhara airport from the weather stations and applied for the all the four catchments.

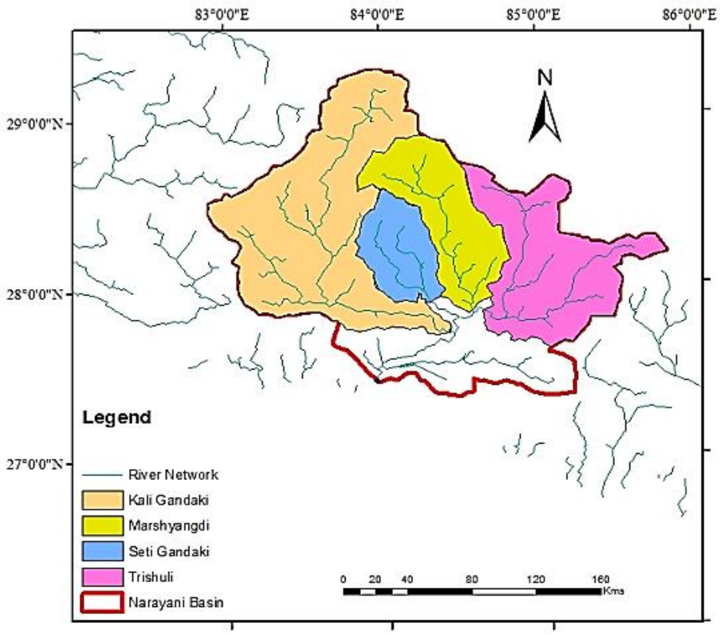


Figure 7-1: Tributaries for case I

Case II:

The model was calibrated for the same setup except the meteorological data was replaced with ERA-Interim dataset.

Case III:

In this case, the calibration was carried out for three tributaries Trishuli, Seti Gandaki and Marshyangdi in Figure 7-2. Kali Gandaki was removed from the calibration as the similar character catchment Marshyangdi was already included to account the climatic variability. The input data for this case was prepared using the station data for precipitation and ERA-Interim data for other climatic data. In the precipitation data, Pokhara Airport was replaced by Rumjakot to minimize the precipitation accumulation due to the local precipitation effect. After this, result of calibration was tested for two characteristically distinct catchments Kaligandaki and East Seti.

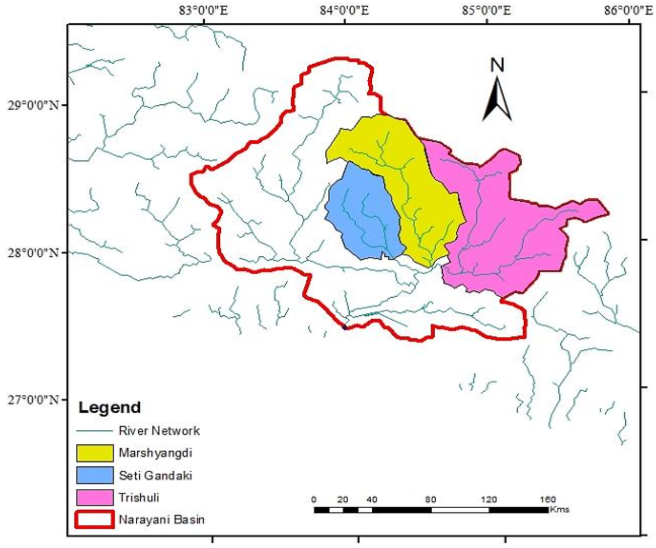


Figure 7-2: Tributaries for case III

Finally, the comparison was carried out using the traditional scaling approach for one of the Himalayan catchments.

7.5 Parameter setup

In the Himalayan catchment like Narayani Basin, the parameter setup is different than the default setup of Norway. Therefore, the parameter was corrected for temperate and precipitation gradient as these two dataset play a vital role in the calibration process.

7.5.1 Temperature gradient

The initial setup of the temperature gradient was -0.6 °C/ 100 m in the model. But for the Himalayan catchment this value is -0.5 °C / 100 m (Timalsina, 2009). Also, if the elevation difference between the selected stations are carefully looked upon then there is a constant falling gradient in the temperature. For this study, this value is taken between the maximum and minimum elevation station by considering the fact that this will represent whole catchment area. The average temperature gradient was taken as -0.5 °C / 100 m based on Table 7-1 below.

Table 7-1: Monthly average temperature gradient ° C/100 m

Temp. Stations	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Chame at 2680 masl	4.9	6.8	10.1	13.2	13.8	14.3	13.6	13.9	14.1	12.0	8.8	5.6
Dhunche at 1982 masl	8.1	9.7	13.5	16.8	18.6	20.2	20.2	20.2	19.5	16.4	12.6	9.2
Nuwakot at 1003 masl	15.3	17.1	21.5	24.4	25.4	26.6	26.2	26.1	25.6	21.7	19.1	15.6

Temp. Stations	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
Pokhara Airport at 827 masl	13.7	16.5	20.5	23.3	24.8	26.2	26.5	26.4	25.5	22.4	18.1	14.7
Bharatpur at 205 masl	15.5	19.0	23.4	27.6	29.0	29.7	28.9	28.6	28.1	25.5	20.9	17.5
Temp. Gradient °C/100 m	-0.43	-0.49	-0.54	-0.58	-0.62	-0.62	-0.62	-0.59	-0.56	-0.54	-0.49	-0.48

7.5.2 Precipitation gradient

The initial precipitation gradient was set at 2% per 100 m. This might be valid for some catchments but for the study area, this value is completely different. The precipitation trend in the Himalayan catchment has been illustrated in many literatures to be in the rising trend up to a certain elevation following a dropping trend after certain elevation range as illustrated by Figure 7-3.

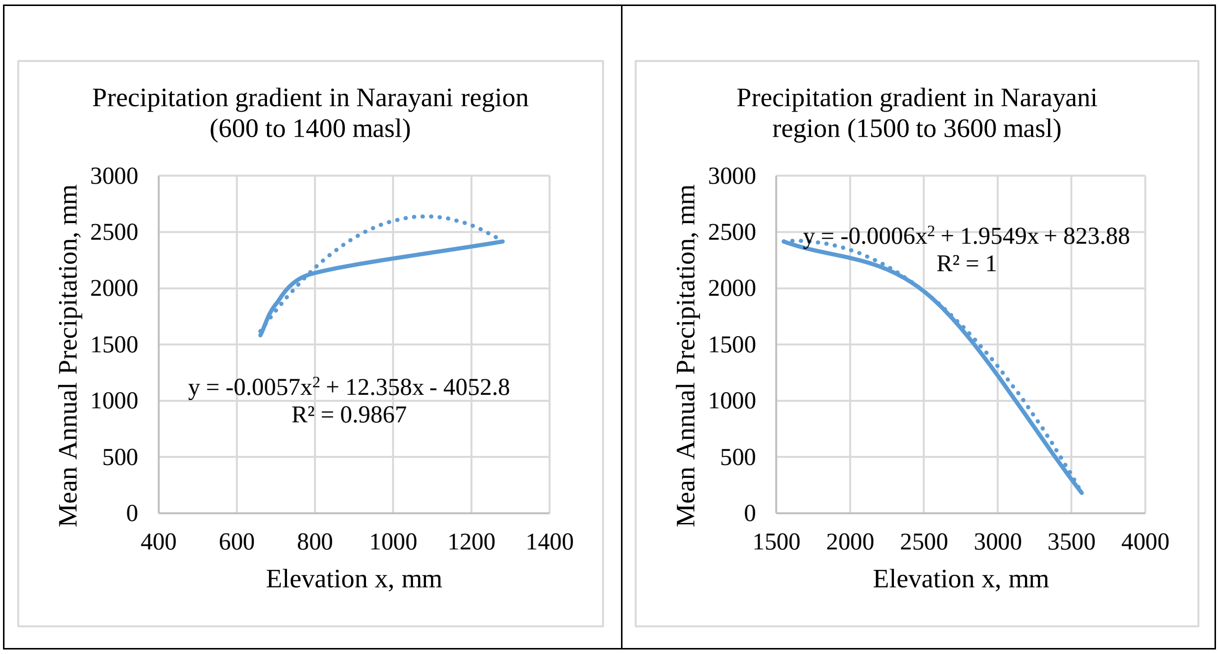


Figure 7-3: Precipitation gradient in the Narayani basin

The R^2 value for both the cases is strong enough to support the fact that the precipitation gradient is rising up to a certain extent in the Himalayan catchment then dropping almost linearly after a certain elevation. The linear gradient is positive in the first case of elevation ranging between 400- 1400 masl and negative in the second case of elevation ranging between 1500-3600 masl. Therefore, to simplify the task, hypsographic curve was considered while taking the value of precipitation gradient.

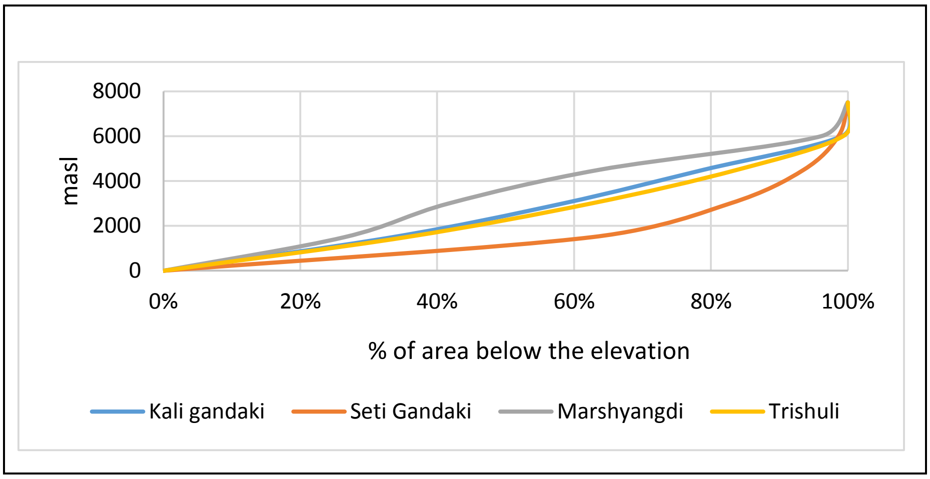


Figure 7-4: Hypsographic curve of the tributaries

The area-elevation distribution of the tributaries are not homogeneous as seen in figure 7-4. For this reason the precipitation gradient was selected based on the distribution of the area as per the elevation. If the area of the tributary under consideration falls mostly below 2500 masl, the rising trend in the gradient was adopted and vice-versa. For the ERA-Interim dataset also, the gradient correction was carried out as before to make the comparison coherent. The areal precipitation in ERA-Interim was observed to be lesser than the meteorological dataset for which the precipitation scale was adjusted to 2 from the default value of 1. (Appendix 4)

7.6 Calibration

The calibration process was carried out by numerical parameter optimization method using SCE-UA optimizer and the resulting calibrated parameter is given in Table 7-2 below.

Table 7-2: Calibration of parameter in SHyFT for Narayani Basin

Description	Symbol	Range	Distribution	Calibrated values		
				Case I	Case II	Case II
Kirchner's model						
Kirchner's Coefficient	c1	-8 to 0	Uniform	-7.854306	-6.95356	-6.131125
	c2	-1 to 1.2	Uniform	-0.094174	-0.67011	0.1551430
	c3	-0.05 to -0.15	Uniform	-0.134291	-0.14612	-0.096684
Actual evapotranspiration						
Actual evapotranspiration	ae_scale_factor	1.5	Constant	1.5	1.5	1.5
Gamma Snow						
Snow/rain threshold temperature [C]	tx	-3 to 2	Uniform	1.816022	1.057002	-1.894830
Slope in turbulent wind function [m/s]	wind_scale	1 to 6	Uniform	2.457624	2.335534	2.540835

Description	Symbol	Range	Distribution	Calibrated values		
				Case I	**Case II**	**Case II**
Maximum liquid water content	max_water	0.1	Constant	0.1	0.1	0.1
Intercept in turbulent wind function	wind_const	1	Constant	1	1	1
Albedo decay rate during melt [days]	fast_albedo_decay rate	5 to 15	Uniform	7.735726	7.152259	10.472206
Albedo decay rate in cold conditions [days]	slow_albedo _decay_rate	20 to 40	Uniform	35.479272	24.88664	23.767640
Surface layer magnitude	surface_magnitude	30	Constant	30	30	30
Maximum albedo value	max_albedo	0.9	Constant	0.9	0.9	0.9
Minimum albedo value	min_albedo	0.6	Constant	0.6	0.6	0.6
Snowfall required to reset albedo [mm]	snowfall_reset _depth	5	Constant	5	5	5
Spatial coefficient variation of fresh snowfall	snow_cv	0.4	Constant	0.4	0.4	0.4
Spatial coefficient variation of fresh snowfall due to forest	snow_cv_forest _factor	0	Constant	0	0	0
Spatial coefficient variation of fresh snowfall due to altitude factor	snow_cv_altitude _factor	0	Constant	0	0	0
Glacier ice fixed albedo	glacier_albedo	0.4	Constant	0.4	0.4	0.4
Precipitation correction						
Precipitation correction	p_corr.scale_factor	1	Constant	1	1	1
Preistley Taylor						
land albedo	albedo	0.2	Constant	0.2	0.2	0.2
alpha	alpha	1.26	Constant	1.26	1.26	1.26

The range of parameter for Kirchner’s was selected by checking the fluctuations in the output generated. The current range was able to produce a good calibration and was able to avoid both the accumulation and inadequacy of water for calibration. Due to the lack of proper data, the range of parameter for Gamma Snow was kept as default. For Preistley Taylor, the alpha value is constant and the albedo was derived as land albedo and kept constant for the study.

The model was then run for different cases.

7.6.1 Parameter Adjustment

The model was calibrated regionally for the aforementioned four catchments for the first two cases. The calibration parameters was obtained as in Table 7-2. This value was checked for the individual tributary. The precipitation scale and the precipitation gradient were adjusted based on the catchment characteristics. Also, the precipitation gradients were adjusted for the individual cases depending on the hypsographic curve and the simulation results after the adjustment. The summary of the adjustment made is presented in Table 7-3.

Table 7-3: Parameter adjustment after calibration

<table>
<tr><th rowspan="2">Catchment</th><th colspan="3">p_corr.scale_factor</th><th colspan="3">Precipitation scale factor (gradient correction)</th></tr>
<tr><th>Case I</th><th>Case II</th><th>Case III</th><th>Case I</th><th>Case II</th><th>Case III</th></tr>
<tr><td>Kali Gandaki</td><td>0.95</td><td rowspan="4">2.0</td><td>NA</td><td colspan="2">0.98</td><td>NA</td></tr>
<tr><td>Seti Gandaki</td><td>0.80</td><td>0.70</td><td colspan="3">0.95</td></tr>
<tr><td>Marshyangdi</td><td>0.98</td><td>1.0</td><td colspan="3">1.04</td></tr>
<tr><td>Trishuli</td><td>1.0</td><td>1.0</td><td colspan="3">1.05</td></tr>
</table>

In case II the precipitation correction scale factor was set to 2.0 to adjust the lower areal precipitation obtained from the ERA-Interim dataset. (See appendix 4)

In case III, the precipitation from Pokhara airport was replaced with precipitation from Rumjakot, to get a fair distribution. The precipitation was corrected for Seti Gandaki with a factor of 0.70 accounting for its small catchment area and to minimize the precipitation accumulation as the input precipitation stations were located to highly influence this tributary. Also, the inverse distance value was limited to 90 kms for the same region for Seti Gandaki.

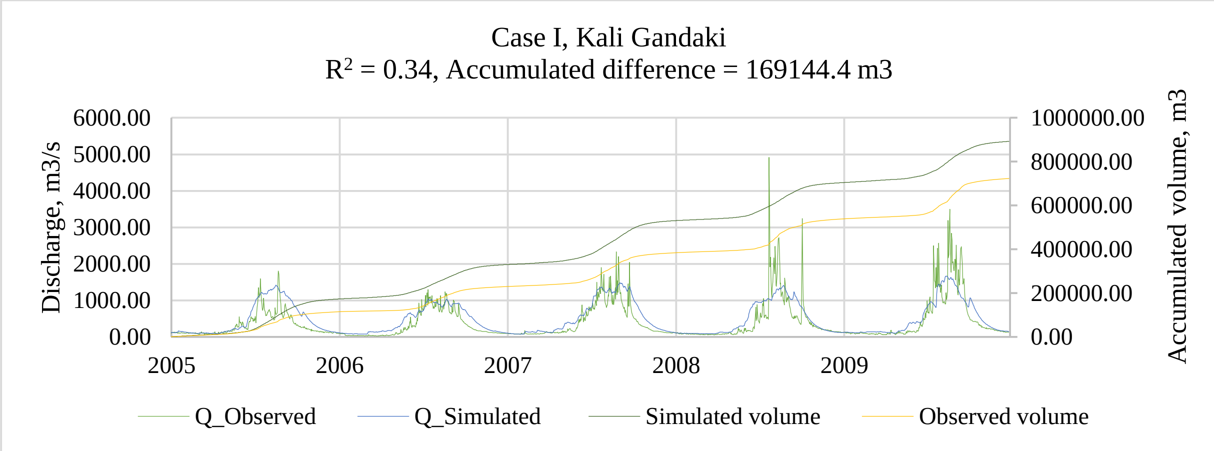
7.6.2 Calibration Evaluation

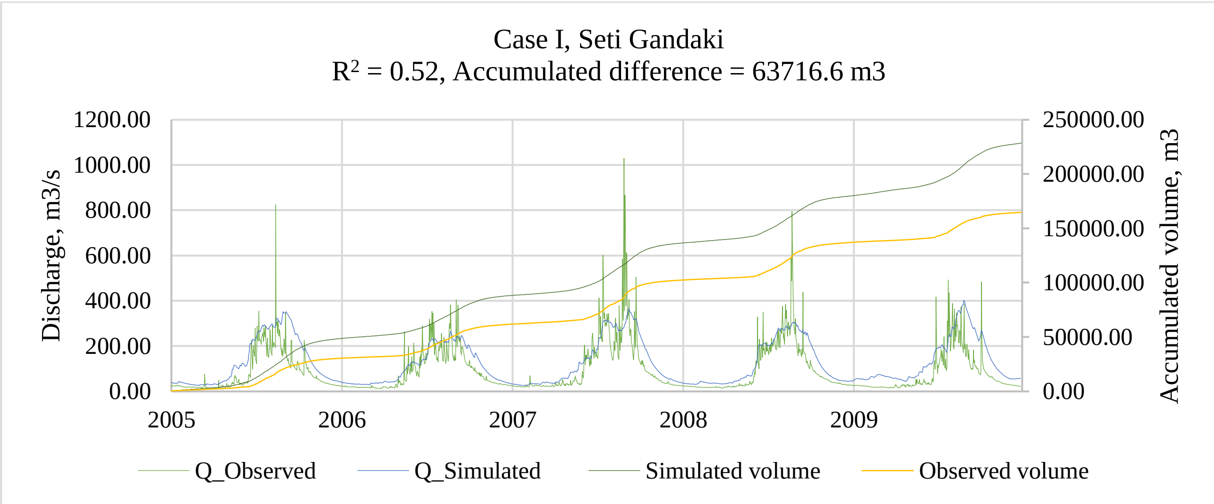
7.6.2.1 Case I

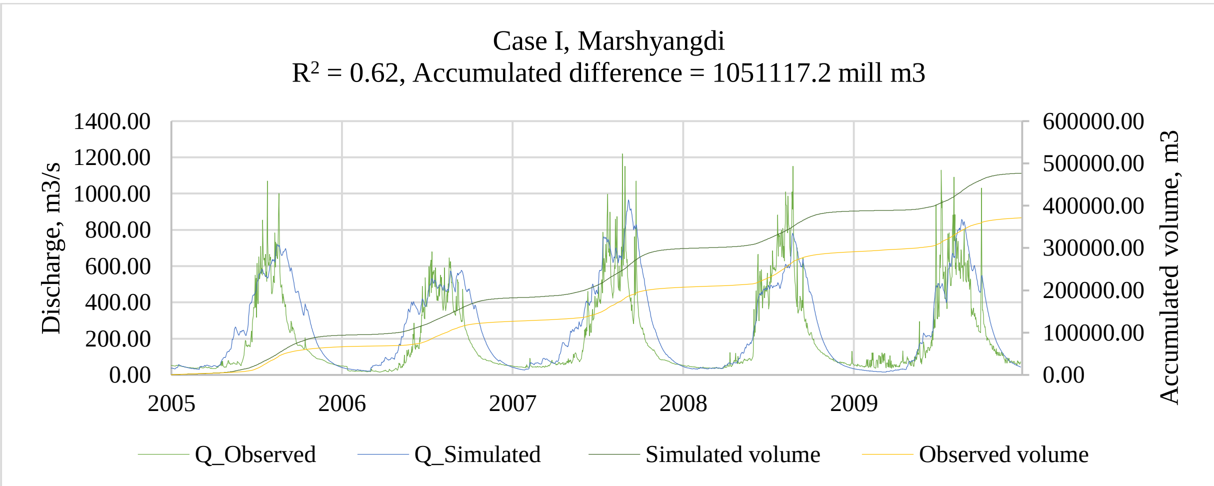
In case I, the Precipitation from Pokhara Airport had a greater influence all over the catchment as it is the area with the largest average annual precipitation all over the catchment and also, the interpolation distance was set at 600 kms. Due to this, the model has calibrated well in Trishuli and Marshyangdi which is the least affected by influence area of this high precipitation region and not so well in Seti Gandaki, owing to its smaller size and large rainfall input(Figure 7-5)

The flow periods are not well simulated for the years 2005-2007 for Marshyangdi and Trishuli.(Figure 7-5) There has been longer episode of low flow than the observed periods in all the catchments. Nevertheless, the recession curve looks better in low flow periods in most of the tributaries.

For small catchment like Seti Gandaki and Marshyangdi, the model is able to simulate the high flow periods, but it completely fails to do so in the larger catchments as seen in Figure 7-5.







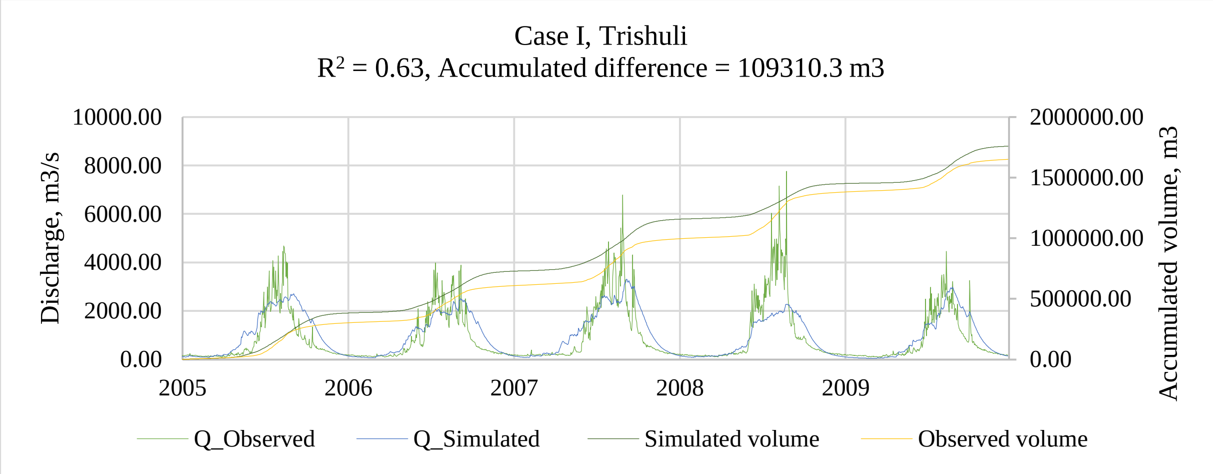


Figure 7-5: Evaluation of the model calibration for case I

The R^2 value and the accumulated difference does not have particular trend for any catchment as seen in the above chart. For same year, the model is well simulated for Marshyangdi and Trisuli but not so well for Seti Gandaki and Kali Gandaki due to the lower precipitation in Pokhara Airport in 2005 than other years. This was verified using objective function in Table 7-4.

Table 7-4: Calibration verification using objective functions for case I

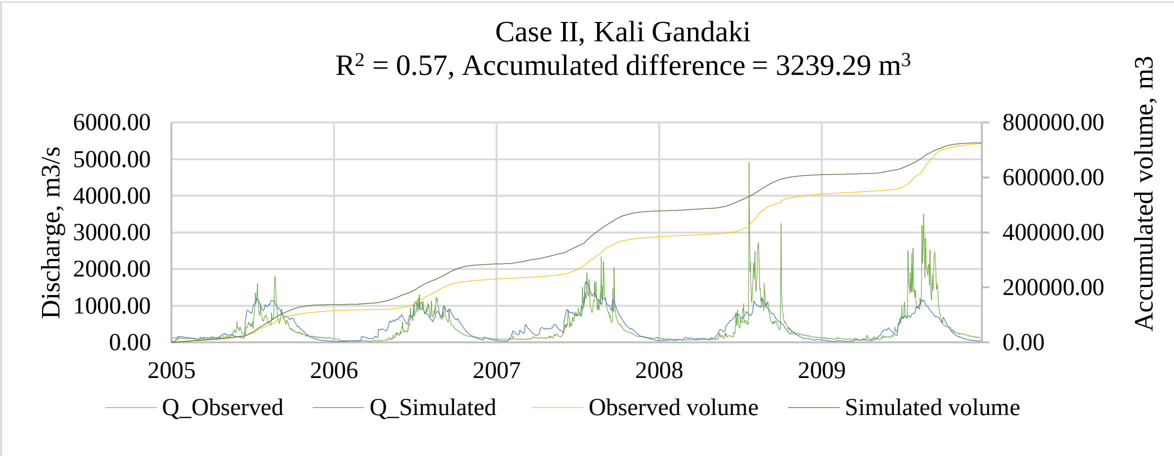
Tributaries	Objective function	2005	2006	2007	2008	2009
Kali Gandaki	R^2	0.19	0.77	0.72	0.51	0.72
	Accumulated Difference	58397.57	41922.10	46310.55	18992.77	3186.51
Seti Gandaki	R^2	0.25	0.71	0.62	0.71	0.12
	Accumulated Difference	18215.87	8280.07	7635.91	8772.77	20083.70
Marshyangdi	R^2	0.56	0.56	0.48	0.77	0.70
	Accumulated Difference	27163.22	28201.36	35863.76	4853.71	9342.29
Trishuli	R^2	0.62	0.71	0.67	0.55	0.71
	Accumulated Difference	80215.86	39074.51	42643.50	-91166.63	38839.91

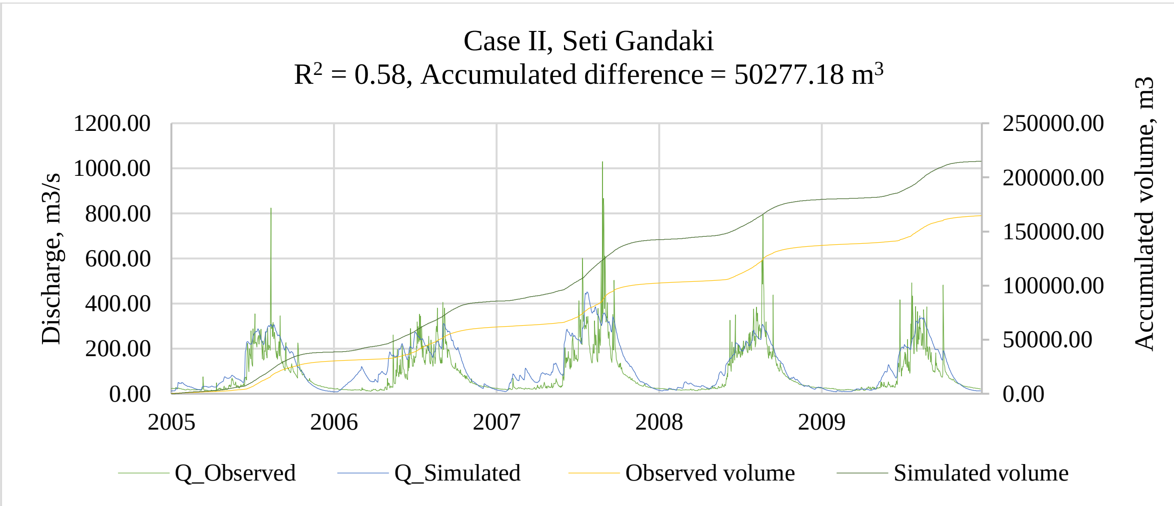
Hence, the results obtained from the meteorological stations did not simulate well for all the catchments as expected, therefore the next step was to check the model with data from ERA-Interim to reduce the local effect of precipitation in a particular catchment. Also, this simulation would be a good check to see whether there was any discrepancy due to the input data.

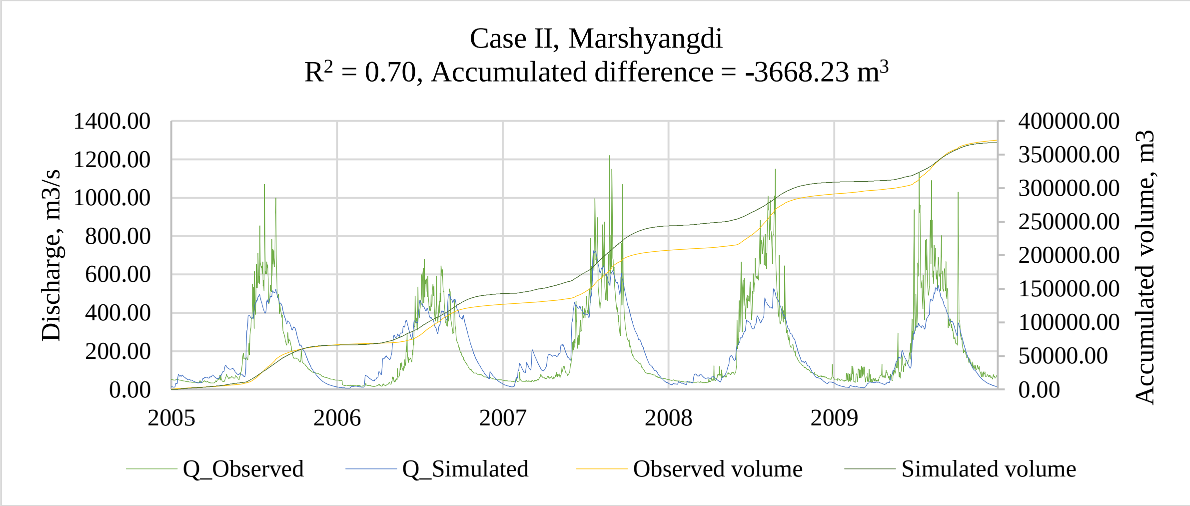
7.6.2.2 Case II

In this case, a clear distinction can be observed between the R^2 value of Kali Gandaki and Seti Gandaki in 2005 which was not affected by the local precipitation effect of Pokhara which certainly improved the local precipitation state than the previous case. (Figure 7-6)

However, this model has failed to generate a higher discharge in the year 2007 for Seti Gandaki despite the fact that was the wetter year. This could have been due to the assimilation process in ERA-Interim which thins out the data value during the quality control. Also the flow periods are simulated for much less than the observed flow except for Trishuli. Even this model has failed to simulate high discharges in Trishuli, but has an improved calibration in Kali Gandaki.(Figure 7-6) The areal precipitation generated from this calibration was much lower than that of the first case. (See appendix 4)







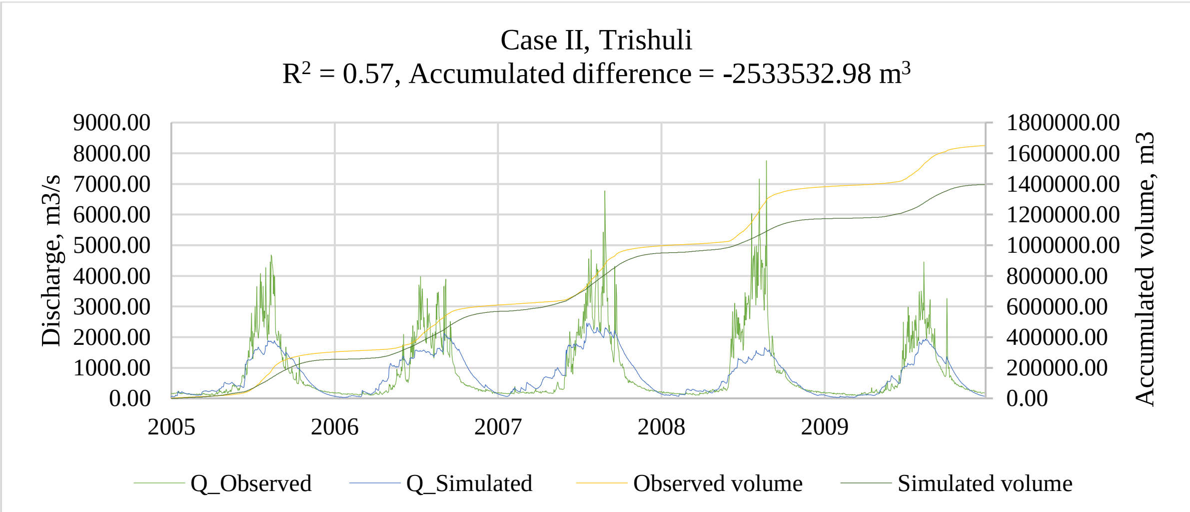


Figure 7-6: Evaluation of the model calibration for case II

Table 7-5: Calibration verification using objective functions for case II

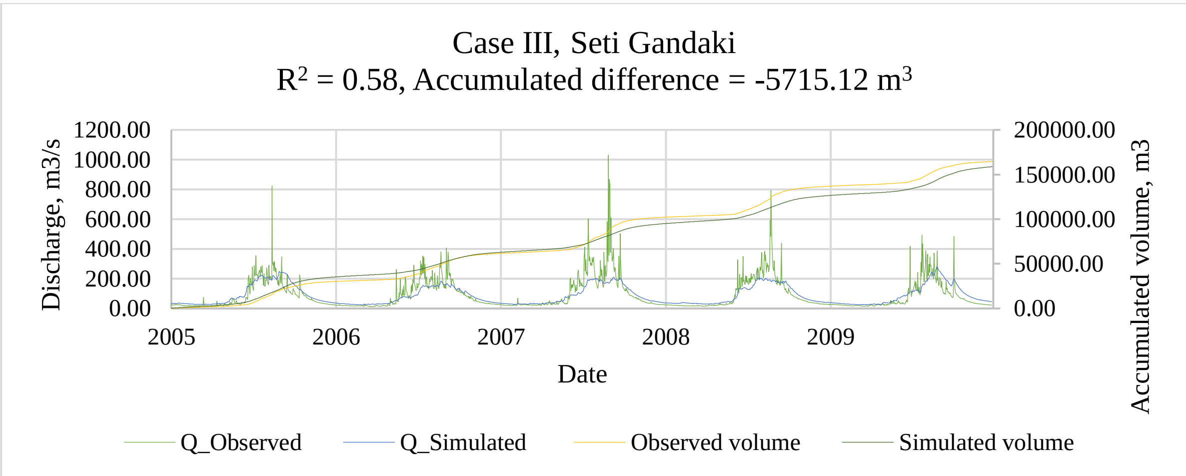
Tributaries	**Objective function**	**2005**	**2006**	**2007**	**2008**	**2009**
Kali Gandaki	R^2	0.62	0.67	0.75	0.48	0.53
	Accumulated Difference	22753.56	32663.67	39028.68	-22986.4	-65615
Seti Gandaki	R^2	0.65	0.39	0.49	0.80	0.59
	Accumulated Difference	8423.161	15543.11	16226.91	2346.06	8003.794
Marshyangdi	R^2	0.80	0.62	0.75	0.68	0.68
	Accumulated Difference	-177.313	16238.19	20415.04	-18749.6	-20349.9
Trishuli	R^2	0.65	0.61	0.68	0.37	0.73
	Accumulated Difference	-47040.6	5725.353	-5098.25	-160817	-44346.5

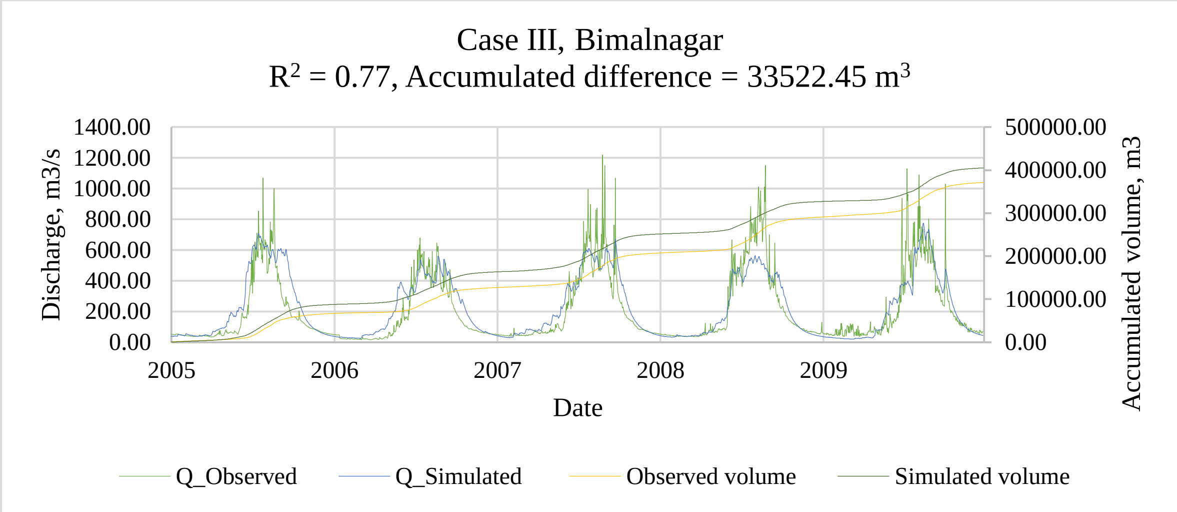
This was verified using objective function in Table 7-4.

For this case also, the model does not perform well in all the years. It is unable to incorporate the big changes that were happening in the catchment. Nonetheless, this model was able to demonstrate that the reduction in the local precipitation effect improves the calibration.
Hence, the next case was performed with the improved precipitation data from the meteorological stations and other climatic data from the ERA-Interim dataset.

7.6.2.3 Case III

In this case, the high precipitation station at Pokhara Airport which was causing the accumulation in Case I was replaced by the precipitation station at Rumjakot with an average precipitation variation. Doing this reduced the spatial difference and improved the interpolation in the cell level. The calibration was carried for three tributaries Trishuli, Marshyangdi and Seti Gandaki. The calibration parameters as obtained in Table 7-2 was further adjusted as for the above two cases with the values from Table 7-3.
The accumulated difference in the discharge was minimized to a greater extent in this case. This curve is well calibrated in the recession zone. The fit between the observed and simulated discharge also look better than the two former cases. However, even this model was unable to process the high runoff periods. Also, the average R^2 value of the calibration improved from -1.34 in case I and 0.39 in case II to 0.66 in this case. Even for the individual catchment, this value improved to 0.58, 0.77 and 0.63 as seen in Figure 7-7. Also the accumulated difference of the observed and simulated runoff decreased by some degree as seen from Table 7-6. The fluctuation in the response time of the catchment reduced from the last two calibrations (Figure 7-5, Figure 7-6 and Figure 7-7) By far, this case looks better than the former two cases. Therefore, case III was used to validate the model and compare the response on larger and smaller ungauged catchments.





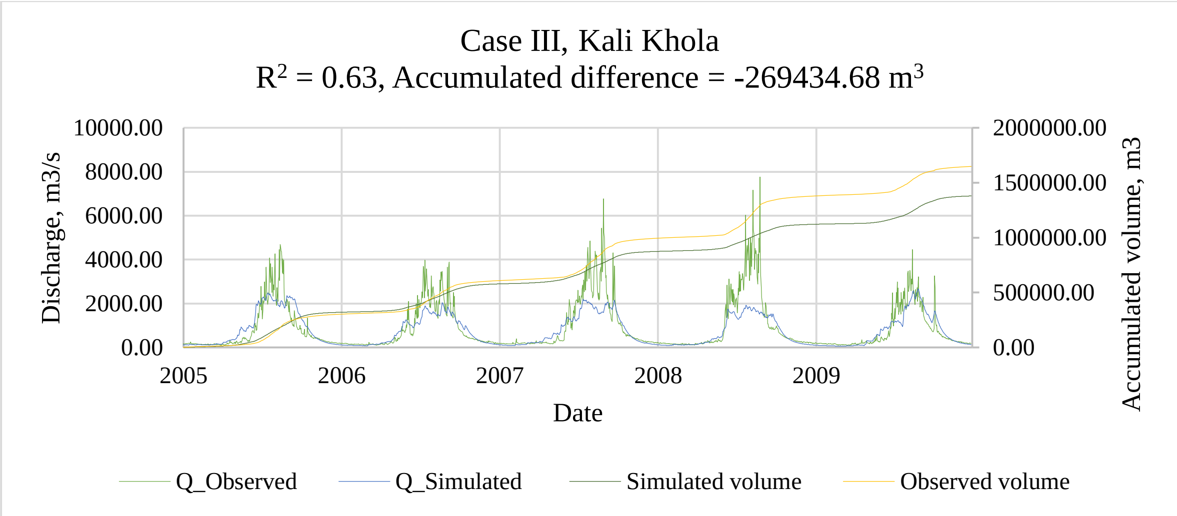


Table 7-6: Calibration verification using objective functions for case III

Tributaries	Objective function	2005	2006	2007	2008	2009
Seti Gandaki	R^2	0.67	0.66	0.48	0.64	0.60
	Accumulated Difference	4928.21	-3747.59	-8471.92	-3195.16	4124.20
Marshyangdi	R^2	0.67	0.77	0.80	0.81	0.76
	Accumulated Difference	-23971.8	-18623.7	-27583.7	-32220.2	-17335.1
Trishuli	R^2	0.70	0.73	0.61	0.49	0.79
	Accumulated Difference	17762.14	-47720.1	-90032	-137888	-10323.8

7.6.3 Validation

Since case III was further used for the study, the calibration was validated for the period of 2010-2012 for each tributary, the result of which is shown below in Figure 7-7 and Table 7-7 .

Table 7-7: Result of model validation for case III using R^2

Date	2005	2006	2007	2008	2009	2010	2011	2012	**Average**
Seti Gandaki	0.668	0.663	0.482	0.636	0.604	0.558	0.581	0.715	**0.613**
Marshyangdi	0.756	0.817	0.792	0.753	0.756	0.808	0.859	0.795	**0.792**
Trishuli	0.685	0.737	0.614	0.497	0.803	0.678	0.784	0.688	**0.686**

From this result, the model seems well calibrated. The R^2 value has not changed drastically between calibration period and validation period. In some cases, this value has improved for the validation period as seen in Table 7-7. The average R^2 value is lesser than that can be accepted in a normal rainfall runoff modelling, but it has to be kept in mind that the regional modelling comes at an expense of losing some catchment characteristics.

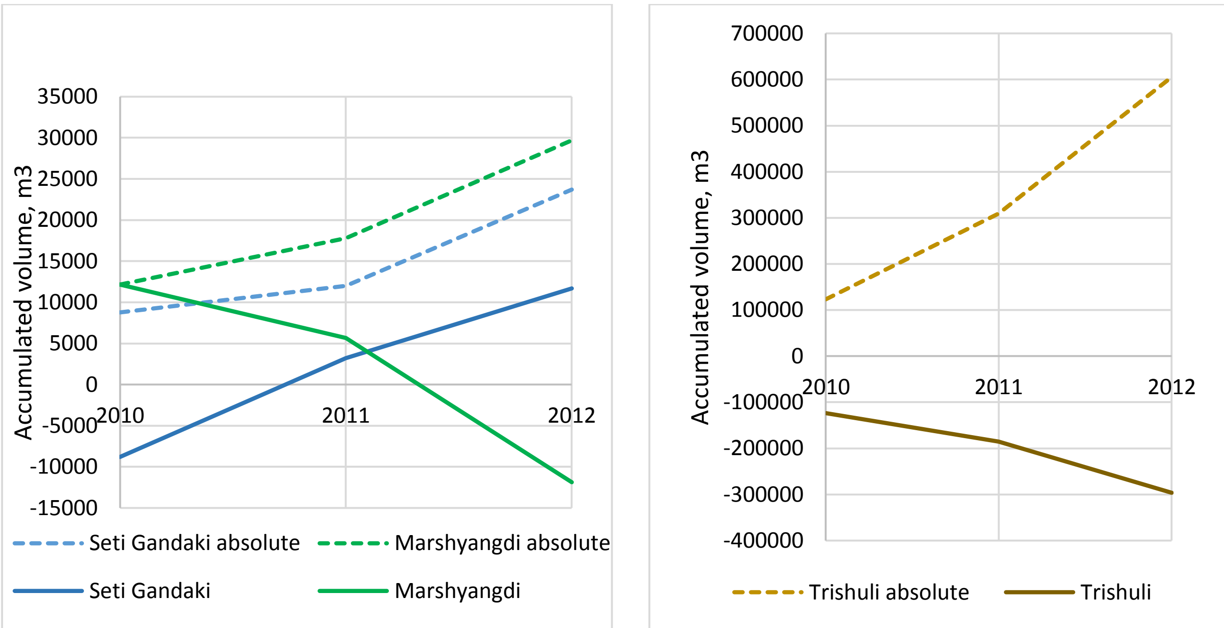


Figure 7-7: Accumulated volume for the validation period

The actual deviation in the catchment is still noteworthy as seen from the absolute accumulation graph of Figure 7-7, yet the model is able to simulate in the flow in the catchment independent of the input dataset.

Therefore, this validation result is accepted as good and the model is further used to generate the results.

7.7 Discussion on the model setup

The first two models are special in their own way despite their results. Generation of good calibration from case II with ERA-Interim data would have created big data resource for

Himalayan catchments where the meteorological stations are located wide apart and not spatially well distributed in elevation. Due to some assimilation process used in ERA-Interim, the intensity of the data might have been disturbed, which failed to give a better result. Even in the first case with the meteorological dataset from the gauging stations in Nepal, the results obtained were not very coherent for all the tributaries.

For all the setups, the precipitation was interpolated spatially by inverse distance weightage method using one gradient factor. Since the precipitation regime is very different with elevation, the gradient factor was unable to process the areal precipitation correctly. The average value of areal precipitation in Trishuli was obtained as 4099.67 mm for gradient of 4% per 100 m which is much higher than the real precipitation observed in Trishuli (2400 mm/year) (Devkota et al., 2008). This is not correct for a mountainous region and should have been a negative gradient after a certain height which further would have produced a lower precipitation value in the higher elevation and higher precipitation in the lower elevation on the southern side of the mountain. Unlike HBV model, it was difficult to manually manipulate the gradient input in the catchment with the elevation, therefore the model had to have a lower precipitation scale factor to accommodate the negative scale effect. But since most of the tributaries are snow fed the model had to exaggerate the precipitation factor to simulate the higher discharge in the catchment.

Unlike the Norwegian catchment where the snowmelt is triggered by the rise in temperature, the snow melt mechanism is completely different in Himalayan catchment as the temperature on the mountains never go above the melting point. The lack of snow data like snow depth, reflectivity, gamma snow model parameters were not changed during the process which further failed to accurately support the outflow from snowmelt in the catchment.

Also, the IDW interpolation used in precipitation calculation can give much skewed result towards the areas with higher number of stations. The result of this was generation very high discharge in Seti Gandaki and Kali Gandaki in case I. Therefore, selection of the scaling factor and maximum distance was another crucial process in this calibration.

The IDW method also projects the elevation to a constant gradient in the catchment. This method completely fails to interpolate in the sinks and grooves, which forms an important part of the land topography. Use of this interpolation made it impossible to change the gradient to positive and negative during the simulation, thus disabling the model to acknowledge the gradient change with elevation.

Nonetheless, the model seems to have interpolated the temperature data well in all the three cases, the gradient control was the only factor that was changed and this was able to generate a reasonable areal temperature series for all the tributaries.

8. Results

This chapter focusses on the result generated using all the theories and processes mentioned earlier.

For the further generation of results, case III was used. A comparison was carried out for the large snow-fed tributary and a small monsoon rainfall generated tributary. The catchments used were Kali Gandaki and East Rapti, both of which have a very different characteristics, details of which are given below in Table 8-1 and Figure 8-1.

Table 8-1: Catchment characteristics distinction

Catchment	Kali Gandaki	East Rapti
GIS-area [km^2]	12031.36	626.17
Glacier Area [km^2]	3.58	0
Elevation Range [masl]	395 – 6772	369 – 2172
Catchment Origin	High Himalayas	Middle Mountains
Precipitation gradient correction	0.98	1.05
Precipitation scale correction	1.0	1.0
Location in the catchment	West	East
Annual average runoff [m^3/s]	432.58	30.98

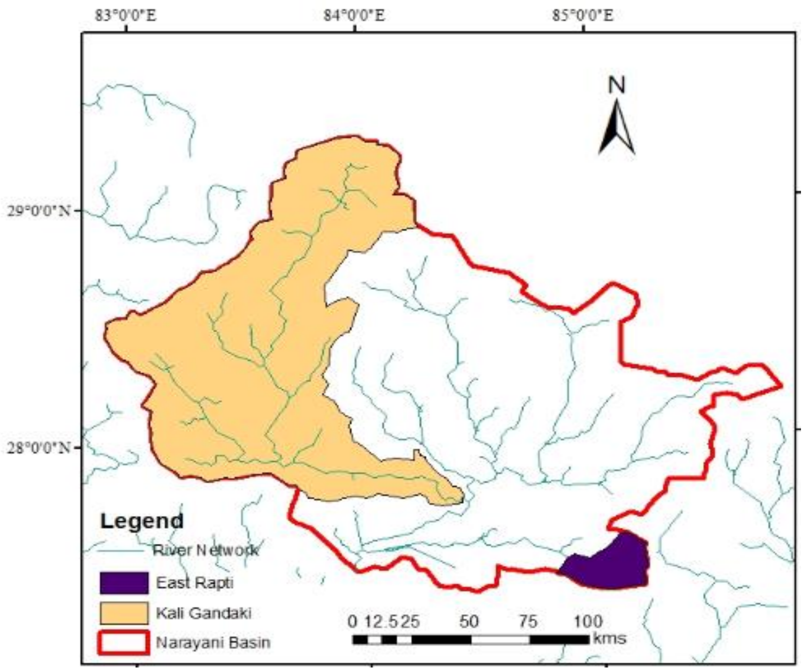
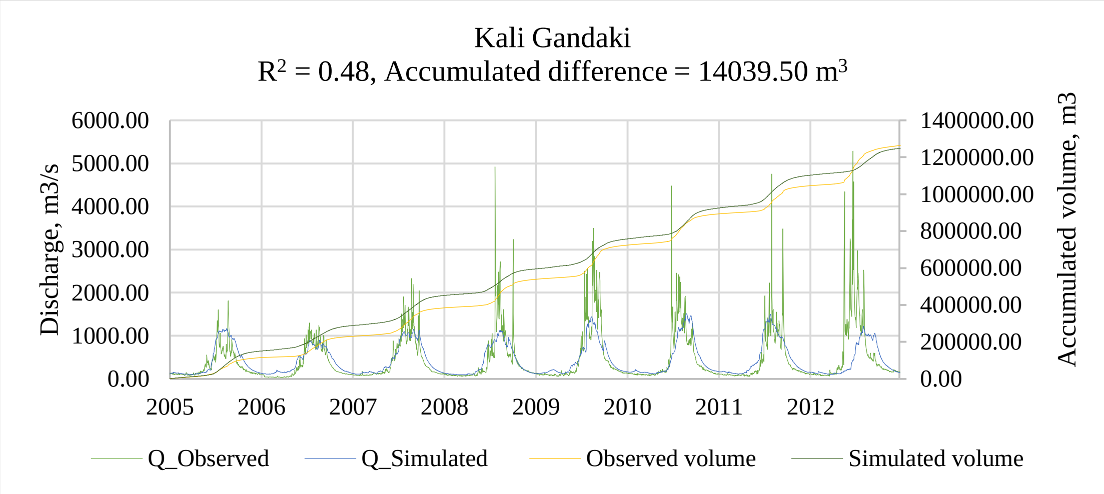


Figure 8-1: Location and size of the catchment under study

8.1 Results and Discussions

8.1.1 Time series of flow

The model generated using case III was further used to predict the runoff of the catchments Kali Gandaki and East Rapti. This was done to check whether or not the calibrated model was able to generate a good results for the ungauged catchments. The only correction carried out in the calibrated model was the gradient correction for precipitation for both the catchments. The model was used between the periods of 2005-2012 the results of which are given below in the chart of Figure 8-2.



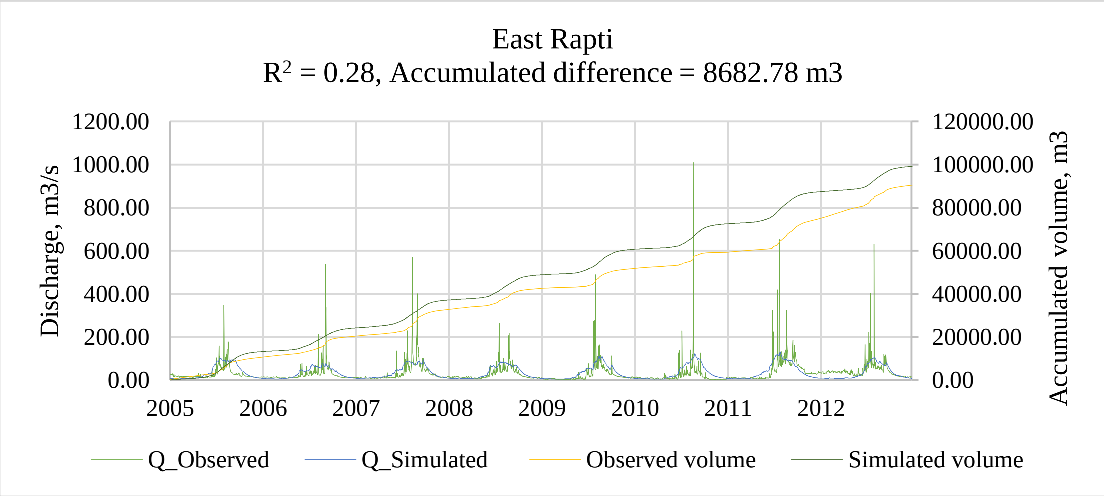


Figure 8-2: Runoff generation using Case III

As seen from the above two results in Figure 8-2, the model seems to be able to predict the flow in the snow-fed rivers than in the monsoon generated rivers. For Kali Gandaki, the R^2 value is reasonable except for the year 2012. While rechecking the input data, it was observed that particular year had very different runoff pattern than the other years under observation as seen in the above chart. When that year was neglected the overall R^2 value improved to 0.65 from 0.48 as seen in Table 8-2. Whereas for East Rapti, the model is unable to generate a good runoff value for the whole period. Furthermore, the model is not able to generate the high flow but is better in the low flow scenario. The difference in the simulated and observed runoff volume seen in Figure 8-2 is due to this deficiency of the model.

Table 8-2: R^2 value for different year under observation

<table>
<tr><th rowspan="2">Year</th><th colspan="4">R²</th></tr>
<tr><th>Kali Gandaki</th><th>Average</th><th>East Rapti</th><th>Average</th></tr>
<tr><td>2005</td><td>0.53</td><td rowspan="5">0.67</td><td>0.14</td><td rowspan="5">0.31</td></tr>
<tr><td>2006</td><td>0.83</td><td>0.24</td></tr>
<tr><td>2007</td><td>0.82</td><td>0.36</td></tr>
<tr><td>2008</td><td>0.52</td><td>0.42</td></tr>
<tr><td>2009</td><td>0.63</td><td>0.38</td></tr>
<tr><td>2010</td><td>0.55</td><td rowspan="3">0.43</td><td>0.10</td><td rowspan="3">0.24</td></tr>
<tr><td>2011</td><td>0.64</td><td>0.40</td></tr>
<tr><td>2012</td><td>0.09</td><td>0.21</td></tr>
<tr><td>Overall</td><td colspan="2">0.48</td><td colspan="2">0.28</td></tr>
</table>

8.1.2 Water Balance approach

Further analysis was carried out using the water balance approach the result of which is given in Table 8-3 below.

Table 8-3: Water balance approach for validation of model

Year	Kali Gandaki				Rapti			
	Precipitation [mm]	**EA [mm]**	**Q [mm]**	**Water Balance [mm]**	**Precipitation [mm]**	**EA [mm]**	**Q [mm]**	**Water Balance [mm]**
2005	1405.45	23.94	1074.18	307.32	1865.89	43.48	1818.72	3.68
2006	1252.66	21.97	991.86	238.84	1556.63	36.42	1512.71	7.51
2007	1537.13	25.51	1169.60	342.01	1832.66	41.39	1788.25	3.02
2008	1373.63	22.86	1034.07	316.71	1641.44	38.52	1619.40	-16.47
2009	1458.25	24.46	1162.78	271.01	1677.72	39.71	1633.15	4.86
2010	1512.29	24.96	1202.93	284.39	1668.93	38.13	1633.76	-2.97
2011	1593.40	27.78	1276.81	288.81	2111.99	50.62	2055.27	6.10
2012	1353.17	23.58	1050.86	278.73	1657.20	39.76	1625.63	-8.18

From the water balance Table 8-3, it can be observed that actual evapotranspiration is higher for Rapti than Kali Gandaki which proves that temperature data has been used properly by the model. There is still some imbalance in the net input and output. Therefore, measures can be taken to minimize this effect. The imbalance is observed more in the bigger catchment with snow-melt outflow than the smaller catchment.

Further, the model was checked for the ability to generate the flow duration curve and extreme events.

8.1.3 Flow duration curve

The flow duration curve was plotted between the years 2005-2012 for both the catchments with observed discharge and simulated discharge as in Figure 8-3. In both the cases the simulated flow is unable to produce the high flows. Since, Kali Gandaki is a snow fed river, there is a high base flow in the river throughout the year. The monsoon period gives the extreme flows which the model has failed to regenerate. But for the 40-50% of the year, the model can reproduce a fairly good flow for Kali Gandaki.(Figure 8-3) This can be useful in planning hydropower and irrigation plants but not for predicting the extreme events like flood.

For Rapti also, the high flow periods are not simulated well by the model. Nonetheless, the model is able to predict the low flow which constitutes about 60% of the flow of the year.

The model parameter calibrated here, is therefore not valid for the extreme flow events.

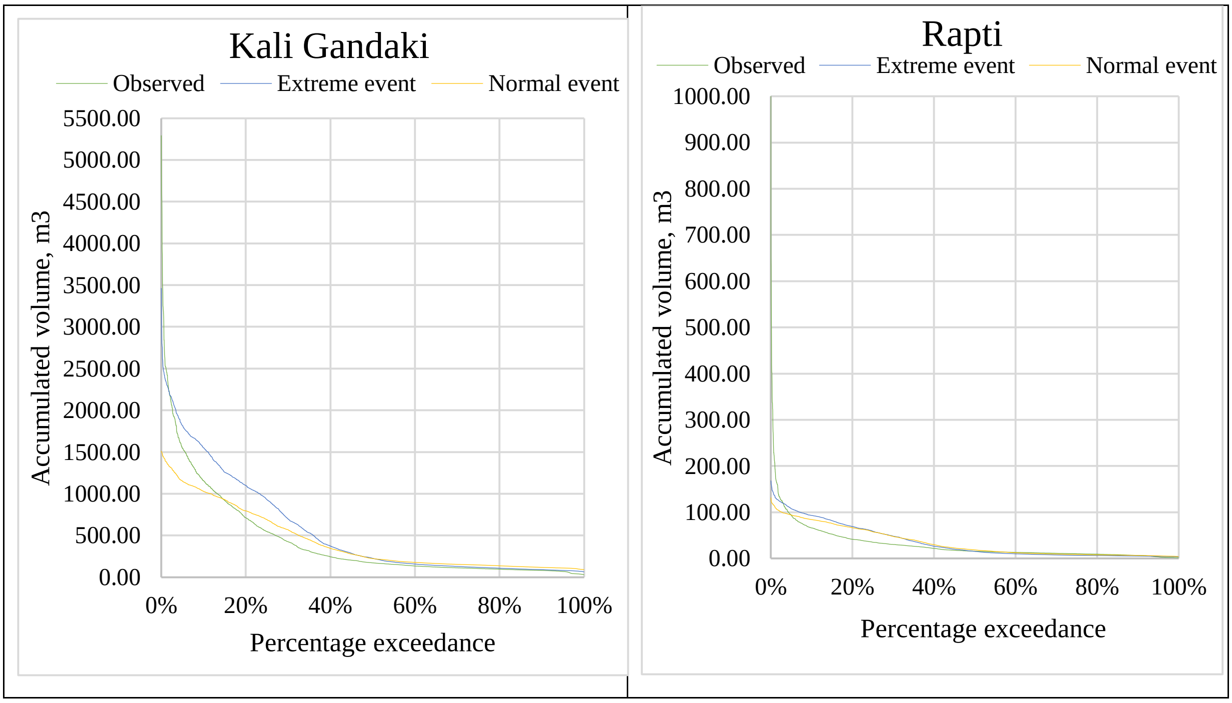


Figure 8-3: Flow duration curve

The model was again calibrated to work for the extreme flow periods by changing some calibration parameters like c1 from -6.1331125 to -4.9060913, c2 from 0.1551430 to 0.6055143 and c3 from -0.096684 to -0.0696684, this improved the extreme event to certain extent but the R^2 value dropped to 0.27 from 0.48 for Kali Ganadaki and 0.22 from 0.28 for Rapti. This improved the prediction for Kali Gandaki to some extent but not for Rapti as seen by the extreme event curve in Figure 8-3. Also, the model performance in the low flow region is still very good after the parameter adjustment. But the cost of changing the calibration parameter comes at the low NSE therefore, the extreme event prediction performance of the model was dropped and further tests were carried for the flow generation in the ungauged catchment using the traditional scaling approach.

8.1.4 Tradeoff between the traditional scaling approach and regional modelling to predict the flow in the ungauged catchment

In order to compare between the traditional scaling method and regional modelling, area proportion method was used for Kali Gandaki. The most suitable catchment for scaling the runoff was selected on the basis of the following factors.

1. Dimensions of the catchment (i.e. area, perimeter)
2. Specific runoff
3. Elevation distribution
4. Climate regime
5. Form of catchment (shape)
6. Surface type (land use, vegetation)

Marshargdi was found to be the most suitable catchment meeting all the above criteria better for Kali Gandaki. The comparison between the traditional approach and regional modelling is given in Figure 8-4.

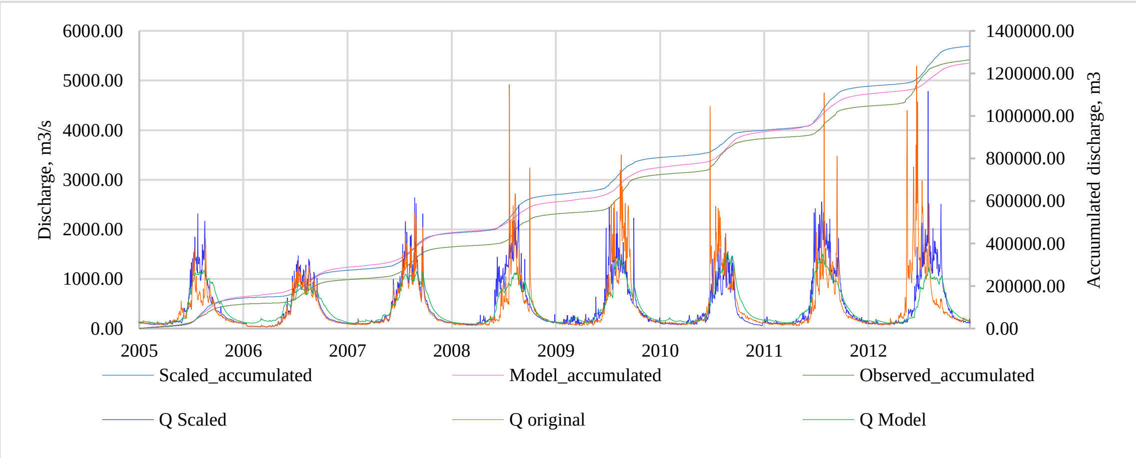


Figure 8-4: Trade-off between traditional approach and regional modelling using SHyFT

The R^2 value obtain by using both methods is listed in the table below.

Table 8-4: Objective function to check the trade-off

Year	2005	2006	2007	2008	2009	2010	2011	2012	**Overall**
From Shyft	0.53	0.83	0.82	0.52	0.63	0.55	0.64	0.09	**0.48**
From Scaling	0.24	0.97	0.97	0.52	0.56	0.55	0.41	0.17	**0.47**

As seen from the Figure 8-4 and both objective function in Table 8-4 and Figure 8-5, neither the model nor the traditional approach is perfectly fitting with the observed discharge. The main tradeoff is on the high flow region, which is well obtained from the scaling approach. However, the recession curve fits better with the SHyFT model except for 2012. Hence, the traditional scaling approach can be used only for generating the high flow in the catchment. To check this flow duration curve was plotted for Kali Gandaki for the traditional scaling approach, modelled flow and the original flow in the catchment in Figure 8-5.

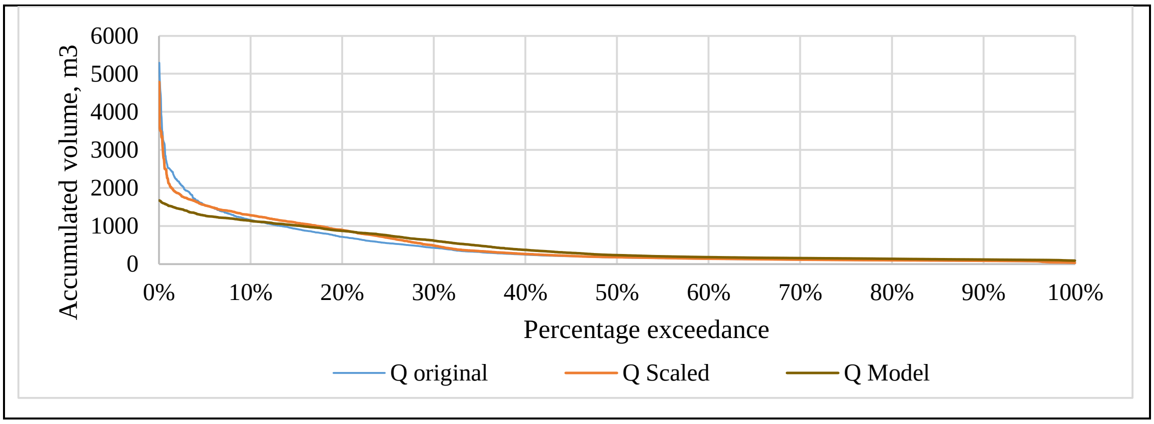


Figure 8-5: Flow duration curve using traditional scaling approach

As seen in Figure 8-5, traditional scaling approach can be best applied when the flow is on the higher side. For the smaller flow cases, it will generate more uncertainty due to its mono-dynamic character. In this method, the flow in the ungauged catchment will solely depend on the feature of the gauging station used for scaling. This method might be applied when there is no modelling tool but there are numbers of gauged catchments on the periphery of ungauged catchment.

8.1.5 Snow covered area and snow melt study

As mentioned before the snow melt mechanism in the Himalayan catchment is very different than that of the Norwegian catchments. However, for both the catchments, the snow cover starts to reduce after the onset of spring and reappear again after autumn. This was tested for the above two catchments. Since Rapti is not a snow-fed river and the elevation is very low for the snow accumulation, it was expected that there would be no snow cover area in this catchment throughout the year. The results obtained from the model also suggested that there was no snow cover in this catchment.

This process was repeated for Kali Gandaki and the results obtained from that is represented in Figure 8-6.

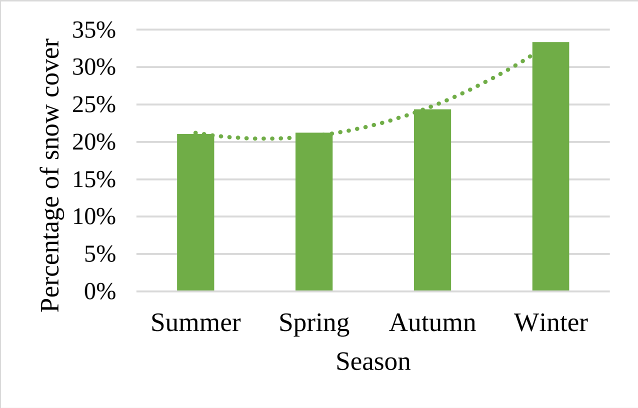


Figure 8-6: Seasonal variation of snow covered area for Kali Gandaki

As seen from the chart 8-6, the snow cover area starts to reduce in spring when the snow melt begins and continues to reduce in summer. The autumn is the beginning of the snow accumulation in the Himalayas and also the graph depicts the same scenario. The winter season has the most snow area coverage because of the low temperature and radiation which slows down the melt process. One interesting observation from Figure 8-6 is that the snow in the Himalayas never drops to zero. There is always some snow reserve throughout the year. Also, in winter due to the low precipitation there is not much added snow, but due to the low snow-melt, the snow cover area increases. For the Himalayan catchments, the nature of snow cover area changes with the season but it never goes to zero as depicted by the chart above.

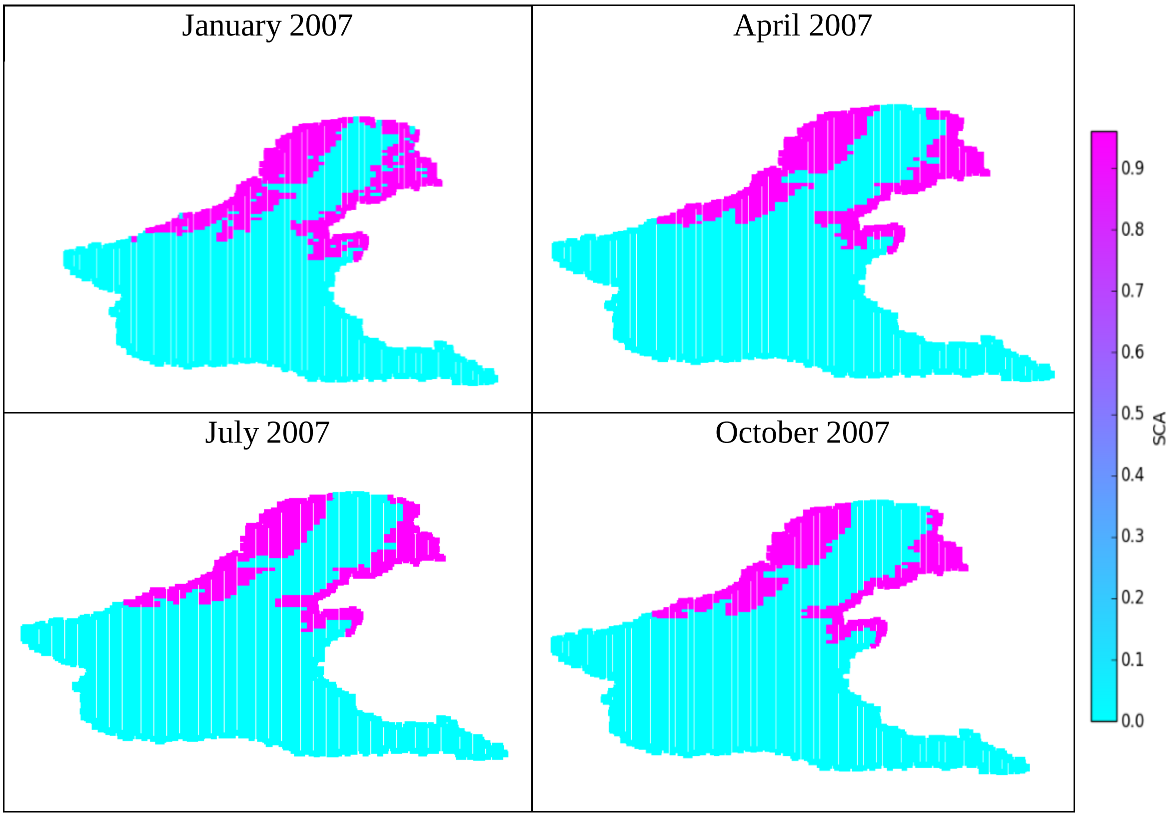


Figure 8-7: Seasonal variation map of the snow cover area (SCA) in Kali Gandaki

Similarly, from Figure 8-7, it is further clarified that the snow accumulation process in the catchment begins after autumn, the melting starts at the beginning of spring. The winter season has the highest snow coverage and the summer season has the lowest. This fact was further validated by the snow-melt outflow in Kali Gandaki in Figure 8-8.

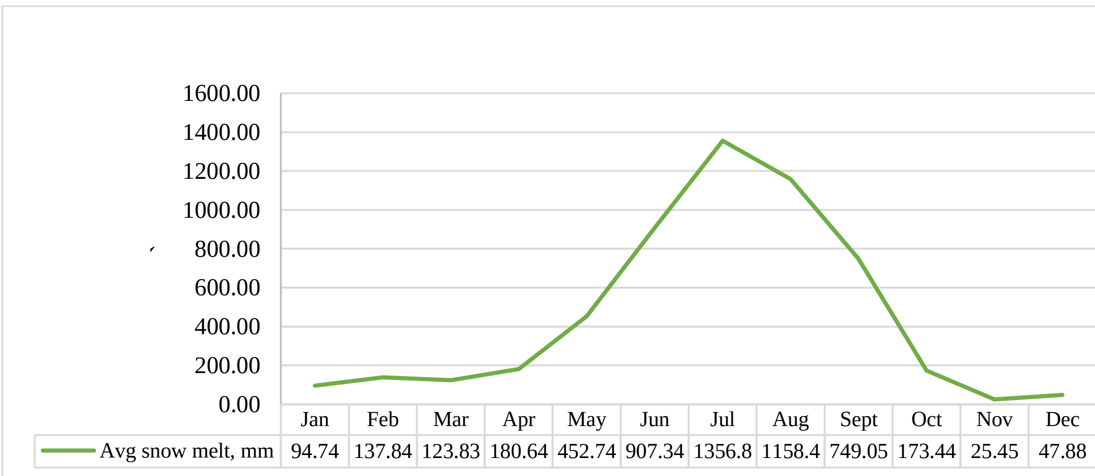


Figure 8-8: Average snow melt outflow per month in Kali Gandaki

Outflow due to the snow melt is highest in the summer season as seen in Figure 8-8,. This slowly drops down in the autumn and becomes very negligible in winter. Again, after spring, the process picks up and the cycle continues.

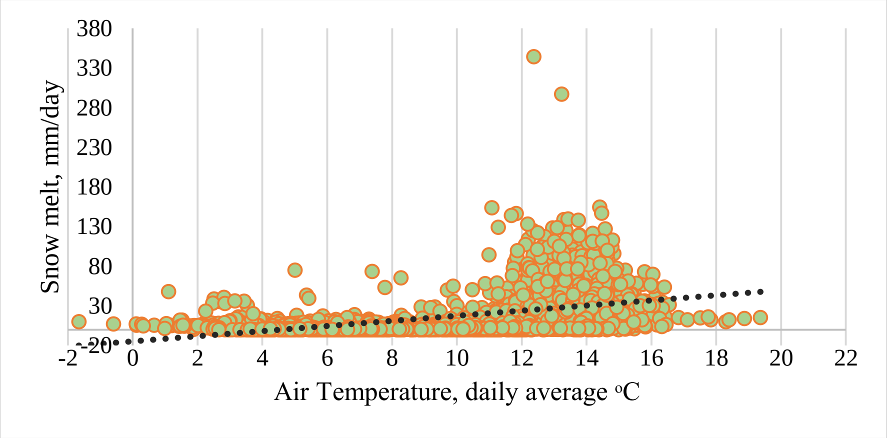


Figure 8-9: Correlation between average daily temperature and daily snow melt.

The correlation seen in Figure 8-9 is good enough to explain the fact that the areal temperature is a good index of energy available for snow melt (Killingtveit et al., 1995). Except for the two points in the graph, the temperature index looks good. The snow melt process as seen from the figure has increased with the rise in temperature. It can also be conferred that the maximum snow melt in the catchment occurs when the temperature is between 12 to 16 degrees, which is the maximum temperature that is likely to be observed in the area generated by the winter snow. Above that elevation, will probably have a permanent snow line with a very different snow regime.

However, additional climatic parameters like wind speed, short wave radiation, relative humidity and snow albedo has to be added in the study to achieve a complete energy balance of snow pack.

9. Discussion and further recommendation

From the results obtained, the model's performance range seems wide. The quality however is still questionable for some observations. There are some explanations to every observation which will be elaborated in this chapter. Finally, some recommendations on the use and improvement of the model will be covered in this section.

9.1 Discussion

From the three cases applied in the study, the weightage of the precipitation data in the model is clarified. The more spatial distribution of the precipitation stations will properly assign the weightage on a particular cell during the calibration process thus eliminating the precipitation accumulation problem on a particular region.

The model performance on the large catchment and smaller catchments clearly reflects the constraints in this model. Had there been more meteorological data towards the Rapti catchment, the simulation result would have produced a better fit. Hence, it is of utmost importance that the input data in the model be spatially well distributed especially for temperature and precipitation stations.

The base flow in the Himalayan River is mostly generated by the snow-melt. The peak flow is generated by the precipitation mostly in the monsoon season. The model was therefore unable to generate higher flow in both the cases, which can be observed from Figure 8-2. This is further supported by the R^2 value of 0.28 in Rapti. Since Rapti is a monsoon-generated river, the model was unable to simulate the precipitation value into the catchment properly which resulted in the poor fit, contrary to the Kali Gandaki River which seems to have a proper base flow generated from the snow melt-outflow.

When the model was set free to calibrate with varying p_corr_scale_factor in model.yaml, then the calibration result gave a p_corr_scale_factor (0.84 for Trishuli, 0.85 for Kali Gandaki, 2.99 for Seti Gandaki, 0.91 for Marshyangdi while calibrating individually) which is very different than the forcefully set factor of 1. This again shows that for the monsoon-generated rivers like Seti Gandaki, the model tries to increase the precipitation so that the simulated runoff will be close to the observed runoff especially during monsoon when the flow increases nearly by 10 times of the base flow. But the regional modelling required this factor be constant in order that the model have a common parameter set.

As explained in section 7.7, the IDW interpolation in the precipitation fails to appreciate the gradient value. For that matter, the calibrated parameters are further changed for gradient correction for individual catchment. This is another important feature of this model. It is flexible enough to adapt to unique parameter change value for each catchment.

For the water-balance, generally the local inflow should be non-negative. High evapotranspiration, infiltration may cause negative value over a long term in the catchment. (Killingtveit et al., 1995). The results from both the catchments show a positive inflow. However, both are missing the infiltration data. Considering some degrees of storage for both the catchments, the balance looks better for Kali Gandaki than for Rapti.

The actual evapotranspiration is higher for Rapti than for Kali Gandaki as the average temperature in Rapti is generally higher which shows that the model is able to interpolate the temperature data very well all over the catchment.

The downside of this model is the prediction of the extreme events like flood as seen from the flow duration curve. The exaggeration of the p_corr_scale_factor also was not able to push the model to generate the high flows. Hence, the setup of the model should be proper with a good input data varying spatially. For this region, the model is currently unfit to be used for such kind of predictions. The use of traditional scaling approach will solve this problem to a greater extent. But it should be kept in mind that the fit in the base flow will be hugely compromised while doing so as seen from the accumulated curve that is constantly widening between the scaled runoff and observed runoff in Figure 8-4.

The model however has a good fit for the snow outflow calculation. For the snow-fed river like Kali Gandaki, the model was able to generate an average flow of 453 mm against the 1120 mm output which makes up to 40% of the flow which is the typical feature of the snow-fed rivers in the Himalayas. The melt period also fits properly with the actual scenario in which the melting in triggered in spring until late autumn.

The snow cover area also does not completely vanish, which again is an important feature of the Himalayas. The area typically shrinks in summer and expands in winter by roughly 15%. Also, the model is able to predict whether or not the catchment is snow-fed as for Rapti, the same calibrated parameters did not show any snow covered area perfectly justifying the absence of glaciers and ice reserves in the input file.

9.2 Conclusion

From the results and discussion, the model can be used to predict inflow from ungauged catchment with some limitations. This model was unable to simulate the extreme flow cases, therefore, traditional scaling approach has to be used for extreme flow cases. However, the low flow periods are properly simulated by the model. Therefore, it can be used for the cases where development can be achieved by tapping water from low flow period.

The snow-melt in the Himalayan catchment was properly demonstrated by the model. But the accuracy of the result will be enhanced by properly setting the calibration parameter by recognizing the characteristics feature of the Himalayan snow. Nonetheless, the snow study can be achieved to some extent using this model.

Due to the use of Python interface, the model is very strong for plotting the graphs, charts and directly storing the values in some external folder.

Also, the flexibility to use between HBV and Gamma snow model for snow melt makes it easier to choose from the options. For HBV snow, the model has to be calibrated again as the scope of current parameter set is unable to cover the HBV snow model.

9.3 Further recommendation

Based on the discussion above, there are some improvements that will help the model to generate a better fit results. Some recommendations to improve the future research on this model are explained below:

- Further research on the ERA-Interim dataset with respect to the precipitation correction will be able to create a better fit for the simulated discharge
- Spatially well distributed meteorological stations has be used to account for the better interpolation of the cell data.
- More knowledge on the snow parameters of the Himalayan catchment will be able to generate an accurate estimation of the outflow from snow-melt and also establish its correlation to snow water equivalent.
- More research on the Kirchner’s parameter used during the study will explain in detail the nature of the hydrograph generated from the simulated runoff.
- High resolution GIS data will provide accurate watershed delineation and correct land-use fraction.
- Inclusion of some more catchments of different sizes in the southern side of the basin in the calibration will further produce a better regional model.

10. References

1. *Temperature Map of Nepal* [Online]. Available: http://www.bestcountryreports.com/Temperature_Map_Nepal.php.
2. B.N. UPRETI, A. S. 2009. WIND ENERGY POTENTIAL ASSESSMENT IN NEPAL. 6.
3. BAJRACHARYA, T. R., ACHARYA, S. & ALE, B. B. 2011. Changing Climatic Parameters and its Possible Impacts in Hydropower Generation in Nepal (A Case Study on Gandaki River Basin). *Journal of the Institute of Engineering*, 8**,** 160-173.
4. BASNET, K. 1989-1992. Temperature Variation in Nepal. *The Himalayan Review,* XX-XXIII**,** 10.
5. BEVEN, K. J. 2011. *Rainfall-Runoff Modelling : The Primer*, Chicester, Wiley.
6. BUMKE, K., KÖNIG-LANGLO, G., KINZEL, J. & SCHRÖDER, M. 2016. HOAPS and ERA-Interim precipitation over sea: Validation against shipboard in-situ measurements. *Atmospheric Measurement Techniques Discussions***,** 1-26.
7. DEE, D. P., UPPALA, S. M., SIMMONS, A. J., BERRISFORD, P., POLI, P., KOBAYASHI, S., ANDRAE, U., BALMASEDA, M. A., BALSAMO, G., BAUER, P., BECHTOLD, P., BELJAARS, A. C. M., VAN DE BERG, L., BIDLOT, J., BORMANN, N., DELSOL, C., DRAGANI, R., FUENTES, M., GEER, A. J., HAIMBERGER, L., HEALY, S. B., HERSBACH, H., HÓLM, E. V., ISAKSEN, L., KÅLLBERG, P., KÖHLER, M., MATRICARDI, M., MCNALLY, A. P., MONGE-SANZ, B. M., MORCRETTE, J. J., PARK, B. K., PEUBEY, C., DE ROSNAY, P., TAVOLATO, C., THÉPAUT, J. N. & VITART, F. 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137**,** 553-597.
8. DEVKOTA, B. D., OMURA, H. & KUBOTA, T. 2008. Vegetation morphology and soil features along unstable road slope: a case study from Mugling Narayanghat Road Section, Central Nepal.
9. DINGMAN, S. L. *Physical Hydrology,* Long Groove, Illinois, University of New Hampshire
10. DUAN, Q., SOROOSHIAN, S. & GUPTA, V. K. 1994. Optimal use of the SCE-UA global optimization method for calibrating watershed models. *Journal of Hydrology*, 158**,** 265-284.
11. ENGELAND, K., BRAUD, I., GOTTSCHALK, L. & LEBLOIS, E. 2006. Multi-objective regional modelling. *Journal of Hydrology*, 327**,** 339-351.
12. ESRL, N. 2014. *Physical Sciences Division* [Online]. Available: http://www.esrl.noaa.gov/psd/data/gridded/data.erainterim.html [Accessed 03/05/2016.
13. GUPTA, H. V., S. SOROOSHIAN, AND P. O. YAPO 1994. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resource Research*, 34**,** 751-763.
14. ICIMOD 2005. Climate and Hydromet. ICIMOD.

15. JEEBAN, P., PIYUSH, D., MADAN LALL, S., SUMAN, A., NIR, Y. K., SONI, M. P., TARENDRA, L., AJAY, K. J., MOHAN, S. & RAMCHANDRA, K. 2015. Spatial and Temporal Variability of Rainfall in the Gandaki River Basin of Nepal Himalaya. *Climate,* 3**,** 210-226.
16. KILLINGTVEIT, Å., SÆLTHUN, N. R., NORGES TEKNISKE HØGSKOLE INSTITUTT FOR, V. & NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET INSTITUTT FOR, V. 1995. *Hydrology,* Trondheim, Norwegian Institute of Technology. Department of Hydraulic Engineering.
17. KIRCHNER, J. W. 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research,* 45.
18. KOLBERG, S. & GOTTSCHALK, L. 2010. Interannual stability of grid cell snow depletion curves as estimated from MODIS images. *Water Resources Research,* 46**,** n/a-n/a.
19. MEKIS, É. 2005. ADJUSTMENTS FOR TRACE MEASUREMENTS IN CANADA. 6.
20. PANDE, R. K. 1987. Altitude Geography of Nepal. Kathmandu: Educational Enterprise.
21. PARIYAR, D. 2008. Country Pasture/Forage Resource Profile Nepal. *FAO***,** 42.
22. PRIESTLEY, C. & TAYLOR, R. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review,* 100**,** 81-92.
23. RINDE, T. 1996. *PINE : a hydrological model with a flexible model structure,* S.l., s.n.
24. RINDE, T. 2015. Class lecture notes on TVM4106 HYDROLOGICAL MODELLING HØST 2015.
25. SHRESTHA, A. B. & ARYAL, R. 2011. Climate change in Nepal and its impact on Himalayan glaciers. *Regional Environmental Change,* 11**,** 65-77.
26. SHRESTHA, J. P., ALFREDSEN, K. & TIMALSINA, N. P. 2012. Regional modelling for estimation of runoff from ungauged catchment, case study of the Saptakoshi basin, Nepal. Institutt for vann- og miljøteknikk.
27. SMITH, T. E. & LESAGE, J. P. 2004. A Bayesian probit model with spatial dependencies. *Advances in econometrics,* 18**,** 127-160.
28. TIMALSINA, N. P. 2009. *Hydrological Modelling of the Upper Trishuli 3A Catchment in Nepal. Comparison of Lumped and Distributed Models.* NTNU.
29. UDDIN, K., SHRESTHA, H. L., MURTHY, M. S. R., BAJRACHARYA, B., SHRESTHA, B., GILANI, H., PRADHAN, S. & DANGOL, B. 2015. Development of 2010 national land cover database for the Nepal. *Journal of Environmental Management,* 148**,** 82-90.
30. ZHANG, Y., VAZE, J., CHIEW, F. H. S., TENG, J. & LI, M. 2014. Predicting hydrological signatures in ungauged catchments using spatial interpolation, index model, and rainfall-runoff modelling. *Journal of Hydrology,* 517**,** 936-948.

Appendix 1: List of Hydro-Meteorological stations used in the study

Precipitation Stations:

Station name	Index	Location	Latitude	Longitude	Elevation [m]
Gurja Khani	616	Myagdi	28° 36' 00"	83° 13' 00"	2530
Tribeni	620	Parbat	28° 02' 00"	83° 39' 00"	700
Darbang	621	Myagdi	28° 23' 00"	83° 24' 00"	1160
Samar Gaun	624	Mustang	28° 58' 00"	83° 47' 00"	3570
Bega	626	Myagdi	28° 28' 00"	83° 36' 00"	1770
Kuhun	627	Myagdi	28° 23' 00"	83° 29' 00"	1550
Sirkon	630	Parbat	28° 08' 00"	83° 37' 00"	790
Musikot	722	Gulmi	28° 10' 00"	83° 16' 00"	1280
Pokhara Airport	804	Kaski	28° 13' 00"	84° 00' 00"	827
Kunchha	807	Lamjung	28° 08' 00"	84° 21' 00"	855
Walling	826	Syangja	27° 59' 00"	83° 46' 00"	750
Rumjakot	827	Tanahun	27° 52' 00"	84° 08' 00"	660
Bharatpur	927	Chitawan	27° 40' 00"	84° 26' 00"	205

Temperature Stations:

Station name	Index	Location	Latitude	Longitude	Elevation [masl]
Pokhara Airport	804	Kaski	28° 13’ 00’’	84° 00’ 00’’	827
Chame	816	Manang	28° 33’ 00’’	84° 14’ 00’’	2680
Bharatpur	927	Chitawan	27° 40’ 00’’	84° 26’ 00’’	205
Nuwakot	1004	Nuwakot	27° 55’ 00’’	85° 10’ 00’’	1003
Dhunche	1055	Rasuwa	28° 06’ 00’’	85° 18’ 00’’	1982

Meteorological Stations:

Station Name	Index	Location	Latitude	Longitude	Catchment area [sq.km]	Elevation [masl]
Mayagdi Khola	404.7	Mangalghat	28° 21’ 10’’	83° 32’ 16’’	1112	914
Modi Khola	406.5	Nayapul	28° 15’ 15’’	83° 43’ 27’’	601	701
Adhikhola	415.1	Borlangpul	28° 15’ 15’’	83° 35’ 58’’	195	749
Kali Gandaki	420	Kotagaun	27° 45’ 00’’	84° 20’ 50’’	11400	198
Seti Gandaki	430.5	Damauli	27° 57’ 12’’	84° 15’ 54’’	1350	290
Madi	438	Shisaghat	28° 06’ 00’’	84° 14’ 00’’	858	457
Marshyandi	439.7	Bimalnagar	27° 57’ 00’’	84° 25’ 48’’	3774	354
Chepekhola	440	Ghambesi	28° 03’ 41’’	84° 29’ 23’’	308	442
Burhi Gandaki	445	Arughat	28° 02’ 37’’	84° 48’ 59’’	4270	485
Trishuli	447	Betrawati	27° 58’ 08’’	85° 11’ 00’’	4110	600
Tadi	448	Belkot	27° 51’ 35’’	85° 08’ 18’’	653	475
Trishuli	449.91	Kalikhola	27° 50’ 08’’	84° 33’ 12’’	16760	220
Narayani	450	Devgat	27° 42’ 30’’	84° 25’ 50’’	31100	180
Rapti	460	Rajaiya	27° 26’ 50’’	84° 58’ 26’’	579	332

Appendix 2: Codes for SHyFT system

Appendix 2.1: YAML files

region.yaml

```
---
repository:
  class:
!!python/name:shyft.repository.netcdf.cf_region_model_repository.CFRegionMo
delRepository
  params:
    data_file: C:\shyft-data\netcdf\Narayani/cell_data.nc

domain:
  EPSG: 32644
  nx: 150
  ny: 100
  step_x: 2500
  step_y: 2500
  lower_left_x: 680000
  lower_left_y: 3020000

catchment_indices:
    - 0
 #   - 1
  #  - 2
  #  - 3
  #  - 4

#parameter_overrides:
#  1228:
#    kirchner:
#        c1: -2.539
...
```

dataset.yaml

```
---
sources:
  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
#  - repository:
!!python/name:shyft.repository.netcdf.erainterim_data_repository.ERAInterim
DataRepository
    types:
      - precipitation
    params:
  #    filename: C:\shyft-data\netcdf\Narayani_era/precipitation_era.nc
      stations_met: C:\shyft-data\netcdf\Narayani_era/precipitation.nc
      selection_criteria: null

  - repository:
!!python/name:shyft.repository.netcdf.cf_geo_ts_repository.CFDataRepository
 # - repository:
!!python/name:shyft.repository.netcdf.erainterim_data_repository.ERAInterim
DataRepository
    types:
      - temperature
    params:
   #   filename: C:\shyft-data\netcdf\Narayani_era/temperature_era.nc
```

```
      stations_met: C:\shyft-data\netcdf\Narayani_era/temperature.nc
      selection_criteria: null

  - repository:
!!python/name:shyft.repository.netcdf.erainterim_data_repository.ERAInterim
DataRepository
    types:
      - wind_speed
    params:
        filename: C:\shyft-data\netcdf\Narayani_era/wind_speed.nc

  - repository:
!!python/name:shyft.repository.netcdf.erainterim_data_repository.ERAInterim
DataRepository
    types:
      - relative_humidity
    params:
        filename: C:\shyft-data\netcdf\Narayani_era/relative_humidity.nc

  - repository:
!!python/name:shyft.repository.netcdf.erainterim_data_repository.ERAInterim
DataRepository
    types:
      - radiation
    params:
        filename: C:\shyft-data\netcdf\Narayani_era/global_radiation.nc
...
```

model.yaml

```
model_t: !!python/name:shyft.api.pt_gs_k.PTGSKModel  # model to construct
model_parameters:
  actual_evapotranspiration:
    ae_scale_factor: 1.5
  gamma_snow:
    calculate_iso_pot_energy: false
    fast_albedo_decay_rate: 7.735726
    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: 35.479272
    snow_cv: 0.4
    tx: 1.816022
    snowfall_reset_depth: 5.0
    surface_magnitude: 30.0
    wind_const: 1.0
    wind_scale: 2.457624
    winter_end_day_of_year: 100
  #hbv_snow:
  #    foo: 10.0
  kirchner:
    c1: -6.554306
    c2: -0.194174
    c3: -0.134291
  precipitation_correction:
    scale_factor: 1.0
  priestley_taylor:
```

```
    albedo: 0.2
    alpha: 1.26
```

interpolation.yaml

```
interpolation_parameters:
  temperature:
    method: btk
    params:
      temperature_gradient: -0.5
      temperature_gradient_sd: 0.25
      nug: 0.5
      range: 200000.0
      sill: 25.0
      zscale: 20.0
    # method: idw
    # params:
      # max_distance: 600000.0
      # max_members: 10
      # distance_measure_factor: 1.0
      # default_temp_gradient: -0.005 # degC/m, so -0.5 degC/100m
      # gradient_by_equation: false
  precipitation:
    method: idw
    params:
      max_distance: 400000.0
      max_members: 10
      distance_measure_factor: 1
      scale_factor: 0.93
  radiation:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
  wind_speed:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
  relative_humidity:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
```

simulation.yaml

```
interpolation_parameters:
  temperature:
    method: btk
    params:
      temperature_gradient: -0.5
      temperature_gradient_sd: 0.25
      nug: 0.5
      range: 200000.0
```

```
      sill: 25.0
      zscale: 20.0
    # method: idw
    # params:
      # max_distance: 600000.0
      # max_members: 10
      # distance_measure_factor: 1.0
      # default_temp_gradient: -0.005 # degC/m, so -0.5 degC/100m
      # gradient_by_equation: false
  precipitation:
    method: idw
    params:
      max_distance: 400000.0
      max_members: 10
      distance_measure_factor: 1
      scale_factor: 0.93
  radiation:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
  wind_speed:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
  relative_humidity:
    method: idw
    params:
      max_distance: 600000.0
      max_members: 10
      distance_measure_factor: 1.0
```

calibration.yaml

```
narayani:
  model_config_file: narayani_simulation_era.yaml
  calibrated_model_file: calibrated_model_era.yaml  # file where the
calibrated params will go
  optimization_method:
  #  name: min_bobyqa # can be 'min_bobyqa', 'dream' or 'sceua'
  #  params:
   #   max_n_evaluations: 5000
   #   tr_start: 0.1
  #    tr_stop: 1.0e-3
    # name: dream
    # params:
      # max_n_evaluations: 1500
    name: sceua
    params:
      max_n_evaluations: 2500
      x_eps: 0.9
      y_eps: 1.0e-1
  target:
  - repository:
!!python/name:shyft.repository.netcdf.cf_ts_repository.CFTsRepository
    params:
      file: C:\shyft-data\netcdf\Narayani/discharge.nc
```

```
      var_type: discharge
    1D_timeseries:
    - catch_id: [0]
      uid: Kotagaun
      start_datetime: 2005-01-01T00:00:00
      run_time_step: 86400 # 3600
      number_of_steps: 2920 # 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: [2]
  #    uid: Bimalnagar
 #     start_datetime: 2005-01-01T00:00:00
 #     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: !!python/name:shyft.api.pt_gs_k.PTGSKOptModel
  calibration_parameters:
    kirchner.c1:
      min: -8.0 # -3.0
      max: 0.0 # 2.0
    kirchner.c2:
      min: -1.0 # 0.8
      max: 1.2 # 1.2
    kirchner.c3:
      min: -0.15
      max: -0.05
    ae.ae_scale_factor:
      min: 1.5
      max: 1.5
    gs.tx:
      min: -3.0
      max: 2.0
    gs.wind_scale:
      min: 1.0
      max: 6.0
    gs.max_water:
      min: 0.1
      max: 0.1
    gs.wind_const:
      min: 1.0
      max: 1.0
    gs.fast_albedo_decay_rate:
      min: 5.0 # 5.0
      max: 15.0 # 15.0
    gs.slow_albedo_decay_rate:
      min: 20.0 # 20.0
```

```
    max: 40.0 # 40.0
  gs.surface_magnitude:
    min: 30.0
    max: 30.0
  gs.max_albedo:
    min: 0.9
    max: 0.9
  gs.min_albedo:
    min: 0.6
    max: 0.6
  gs.snowfall_reset_depth:
    min: 5.0
    max: 5.0
  gs.snow_cv:
    min: 0.4
    max: 0.4
  gs.snow_cv_forest_factor:
    min: 0.0
    max: 0.0
  gs.snow_cv_altitude_factor:
    min: 0.0
    max: 0.0
  gs.glacier_albedo:
    min: 0.4
    max: 0.4
  p_corr.scale_factor:
    min: 1.0
    max: 2.0
  pt.albedo:
    min: 0.2
    max: 0.2
  pt.alpha:
    min: 1.26
    max: 1.26
```

Appendix 2.2: text2netcdf converter

```
# -*- 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 2005-01-01 00:00:00 +00:00','values':t_v},

#{'name':'station_name','dtype':'S2','dims':('station','strlen',),'cf_role'
:'timeseries_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','values':name_v},

{'name':'x','dtype':'f8','dims':('station',),'units':'m','axis':'X','standa
rd_name':'projection_x_coordinate','values':x_v},

{'name':'y','dtype':'f8','dims':('station',),'units':'m','axis':'Y','standa
rd_name':'projection_y_coordinate','values':y_v},

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

{'name':'crs','dtype':'i4','dims':(),'grid_mapping_name':'transverse_mercat
or','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',),'va
lues':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':'projection_x_coordinate','values':x_v},

{'name':'y','dtype':'f8','dims':('cell',),'units':'m','axis':'Y','standard_
name':'projection_y_coordinate','values':y_v},
```

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

{'name':'crs','dtype':'i4','dims':(),'grid_mapping_name':'transverse_mercat
or','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 = 'C:/Users/Anugya/Shyft_Narayani/text' # path to folder
with input text files
    netcdf_folder = 'C:/Users/Anugya/Shyft_Narayani/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':re
ad_hydclim_from_txt,'write_fcn':write_hydclim_to_nc},

'temperature':{'abbrv':'temp','fil_prefix':'temperature','read_fcn':read_hy
dclim_from_txt,'write_fcn':write_hydclim_to_nc},

'wind_speed':{'abbrv':'wind','fil_prefix':'wind_speed','read_fcn':read_hydc
lim_from_txt,'write_fcn':write_hydclim_to_nc},

'relative_humidity':{'abbrv':'rh','fil_prefix':'relative_humidity','read_fc
n':read_hydclim_from_txt,'write_fcn':write_hydclim_to_nc},

'radiation':{'abbrv':'rad','fil_prefix':'global_radiation','read_fcn':read_
hydclim_from_txt,'write_fcn':write_hydclim_to_nc},

'discharge':{'abbrv':'q','fil_prefix':'discharge','read_fcn':read_hydclim_f
rom_txt,'write_fcn':write_hydclim_to_nc},

'cell_data':{'abbrv':'cell_info','fil_prefix':'cell_data','read_fcn':read_c
ellinfo_from_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,))
```

Appendix 2.3: Program files

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\\narayani_simulation.yaml'

cfg = yaml_configs.YAMLSimConfig(config_dir, 'narayani')

# 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()
```

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
ConfigCalibrator7
from shyft.orchestration.simulators import config_simulator

cfg =
yaml_configs.YAMLCalibConfig('C:\\shyft\\shyft\\tests\\netcdf\\narayani_cal
ibration.yaml', 'narayani')
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)
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)))
```

sim_obs.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\\narayani_simulation.yaml'
config_section = "narayani"
```

```
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\\Narayani'))
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'][0:1800,0]
    obs_date = [datetime.utcfromtimestamp(t) for t in
dset.variables['time'][0:1800]]


# 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')
```

Appendix 3: SHyFT

Appendix 3.1: Requirement and installation procedure.

The requirements to compile and run SHyFT in windows are listed below.

- GIT bash and GIT CMD
- A Python3 (3.4 or higher) interpreter
- The SWIG wrapping tool (>= 3.0.5)
- The python packages NumPy, netcdf4, pyyaml, gdal, matplotlib, requests, nose, shapely, pyproj

There is a step-wise procedure to install SHyFt and get it up and running which is mentioned below.

1. Download the SHyFT package from: https://github.com/statkraft/shyft
2. Install GIT from https://git-scm.com
3. Start GIT and move it to the location where SHyFT will be unpacked (GIT is very important to clone SHyFT programs therefore, it is better that both GIT and cloned SHyFT be in the same location).
4. To take the SHyFT program to your desired location type "git clone https://github.com/statkraft/shyft.git" in the GIT window. If the desired location is C:/, then the directory should be C:/ in GIT. The files are then copied to your disk.
5. In the shyft-data: type "git clone https://github.com/statkraft/shyft-data.git" in the GIT window, again by using the same directory.
6. Go to https://github.com/statkraft/shyft/releases/tag/SK_2016_04_27 and download win_x64_common_bin.zip and win_x64_shyft_shyft_api_np110py34.zip and unzip the files from win_x64_shyft_shyft_api_np110py34.zip to shyft/shyft so that it removes the existing api directory. Also unzip win_x64_common_bin.zip into shyft/shyft/api and check if there is *.pyd files inside.
7. Install Anaconda 3.4, SWIG and PyCharm.
8. Install packages like NumPy, netcdf4, pyyaml, gdal, matplotlib, requests, nose, shapely, pyproj. From the command line write: conda install <name of package>, e.g. conda install numpy

Appendix 3.2: Snow melt model in SHyFT

HBV snow parameters in SHyFT

<table>
<tr><th colspan="2">Parameter class</th><th colspan="2">State class</th><th>Response class</th></tr>
<tr><td>S</td><td>snowfall redistribution vector</td><td rowspan="4">swe</td><td rowspan="4">snow water equivalent of the snowpack [mm]</td><td rowspan="7">set_outflow: set the value of the outflow [mm]</td></tr>
<tr><td>intervals</td><td>starting points for the quantiles</td></tr>
<tr><td>Lw</td><td>max liquid water content of the snow</td></tr>
<tr><td>Tx</td><td>threshold temperature determining if precipitation is rain or snow</td></tr>
<tr><td>Cx</td><td>temperature index, i.e.,
melt = cx(t - ts) in mm per °C</td><td rowspan="3">sca</td><td rowspan="3">fraction of area covered by the snowpack [0,1]</td></tr>
<tr><td>Ts</td><td>threshold temperature for melt onset</td></tr>
<tr><td>Cfr</td><td>refreeze coefficient
refreeze = cfr*cx*(ts - t)</td></tr>
</table>

Classes of Gamma snow routine in SHyFT

Parameter class	
glacier_fraction	Fraction of area covered by glacier
winter_end	Last day of accumulation season
initial_bare_ground_fraction	Bare ground fraction at melt onset
snow_cv	Spatial coefficient variation of fresh snowfall
Tx	Snow/rain threshold temperature [C]
wind_scale	Slope in turbulent wind function [m/s]
wind_const	Intercept in turbulent wind function
max_water	Maximum liquid water content
surface_magnitude	Surface layer magnitude
max_albedo	Maximum albedo value
min_albedo	Minimum albedo value
fast_albedo_decay_rate	Albedo decay rate during melt [days]
slow_albedo_decay_rate	Albedo decay rate in cold conditions [days]
snowfall_reset_depth	Snowfall required to reset albedo [mm]

glacier_albedo	Glacier ice fixed albedo
calculate_iso_pot_energy	Whether or not to calculate the potential energy flux
LandType	any of {LAKE, LAND}
State Class	
albedo	Broadband snow reflectivity fraction
Lwc	liquid water content [mm]
surface_heat	Snow surface cold content [J/m^2]
Alpha	Dynamic shape state in the SDC
sdc_melt_mean	Mean snow storage at melt onset [mm]
acc_melt	Accumulated melt depth [mm]
iso_pot_energy	Accumulated energy assuming isothermal snow surface [J/m^2]
temp_swe	Depth of temporary new snow layer during spring [mm]
Response class	
Sca	set snow covered area
storage	set snow storage [mm]
outflow	set water outflow [mm]

Appendix 3.3: Input format for SHyFT

Appendix 3.3.1: Cell Data

EPSG	**32644**							
Catchment_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	234126	3074030	489.33	2000000	0.00	0.50	0.00	0.00
0	235421	3074410	471.33	3000000	0.00	0.00	0.00	0.00
0	237025	3075000	701.80	1000000	0.00	0.00	0.00	0.00

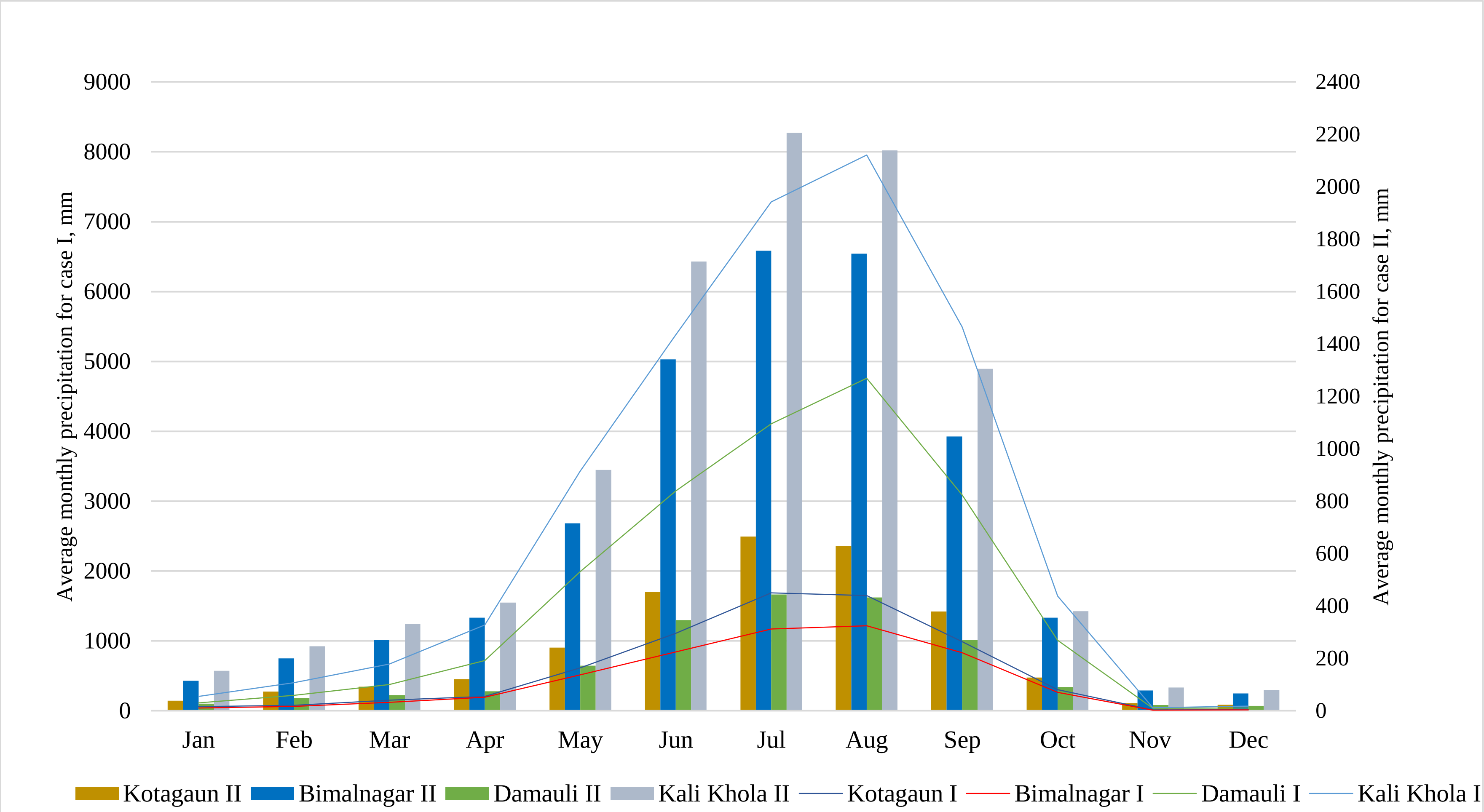
Appendix 3.3.2: Discharge data

EPSG	32644		
Missing_value	-999		
Unit	m3 s-1		
Station_name	Kotagaun	Damauli	Bimalnagar
X_coordinate [m]	756222	834562	824712
Y_coordinate [m]	3072217	3096321	3096051
Elevation [m]	198	290	354
Catchment_ID	0	1	2
2005.01.01 00:00	142.00	25.70	55.40
2005.01.02 00:00	121.00	26.40	55.40
2005.01.03 00:00	140.00	25.50	54.50

Appendix 3.3.3: Climate data

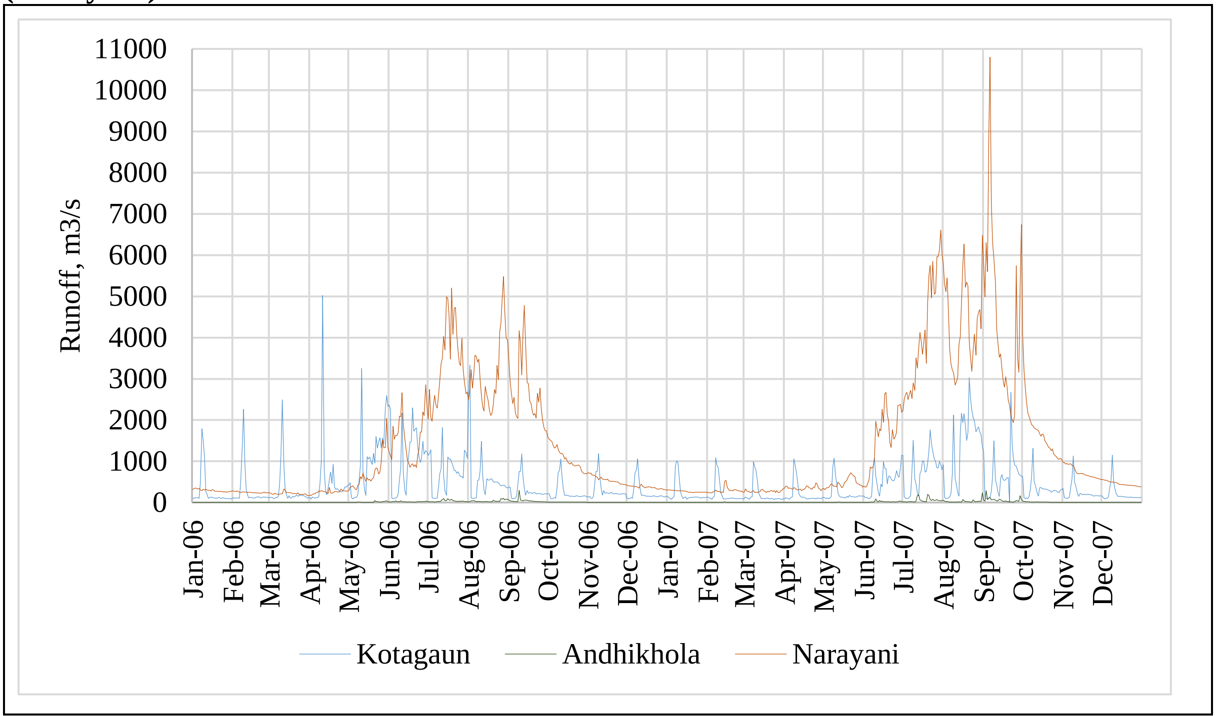
EPSG	32644		
Missing_value	-999		
Unit	mm hr-1		
Station_name	GURJA KHANI	DARBANG	SAMAR GAUN
X_coordinate [m]	716759	735172	771241
Y_coordinate [m]	3165678	3142010	3207485
Elevation [m]	2530	1160	3570
2005.01.01 00:00	0.00	0.00	0.00
2005.01.02 00:00	0.00	0.00	0.00
2005.01.03 00:00	0.00	0.00	0.42

Appendix 4: Areal precipitation of meteorological station and ERA-Interim dataset



Appendix 5: Comparison of runoff from stations

Appendix 5.1: Kotagaun with stations upstream (Andhikhola) and downstream (Narayani)



Appendix 5.2: East Rapti and station downstream (Narayani)